AQIS 2017 Programme Schedule [4 September – 8 September 2017]

Day 1 (Monday, 4 September 2017)

9:00am - 10:00am [Invited talk] Shalev Ben-David

The structure necessary for quantum speedups

10:30am - 11:00am [Long talk]

Experimental certification of millions of genuinely entangled atoms in a solid

Florian Fröwis, Peter Strassmann, Alexey Tiranov, Corentin Gut, Jonathan Lavoie, Nicolas Brunner, Félix Bussières, Mikael Afzelius and Nicolas Gisin

11:00am - 11:30am [Long talk]

Quantum non-malleability and authentication

Gorjan Alagic and Christian Majenz

11:30am - 12:00pm [Long talk]

Dining Philosophers, Leader Election and Ring Size problems, in the quantum setting

Maor Ganz, Dorit Aharonov and Loick Magnin

1:30pm - 2:45pm

[Parallel session] 1A

(1) *Characterizations of symmetrical and partial Boolean functions with exact quantum query complexity*

Daowen Qiu and Shenggen Zheng

(2) *Quantum algorithm for linear differential equations with exponentially improved dependence on precision*

Dominic Berry, Andrew Childs, Aaron Ostrander and Guoming Wang

(3) *Quantum centrality ranking via quantum walks and its experimental* realisation

Joshua Izaac, Xiang Zhan, Jian Li, Peng Xue, Paul Abbott, Xiaosong Ma and Jingbo Wang

[Parallel session] 1B

(1) *Resource Destroying Maps* with new applications

Zi-Wen Liu, Xueyuan Hu, Ryuji Takagi and Seth Lloyd

(2) *Logical paradoxes in deterministic quantum state injection*

Nadish de Silva

(3) *Convex geometry of quantum resource quantification: A general framework for measures of quantum resources*

Bartosz Regula

2:45pm - 3:35pm [Tutorial 1a] Peter Høyer

4:00pm - 4:50pm [Tutorial 1b] Peter Høyer

Title: Quantum walks

- **5:00pm - 6:30pm [Poster session 1]**

Day 2 (Tuesday, 5 September 2017)

9:00am - 10:00am [Invited talk] Mio Murao

Higher order quantum operations of unitaries and their implications

---

10:30am - 11:00am [Long talk]

Bell correlations in many-body systems

Nicolas Sangouard, Sebastian Wagner, Roman Schmied, Batiste Allard, Matteo Fadel Valerio Scarani, Philipp Treutlein and Jean-Daniel Bancal

11:00am - 11:30am [Long talk]

The information cost of quantum memoryless protocol

André Chailloux, Iordanis Kerenidis and Mathieu Lauriere

11:30am - 12:00pm [Long talk]

Compression for Quantum Population Coding

Yuxiang Yang, Ge Bai, Giulio Chiribella and Masahito Hayashi

1:30pm - 2:45pm

[Parallel session] 2A

(1) *Detecting metrologically useful asymmetry entanglement by few local measurements*
Chao Zhang, Benjamin Yadin, Zhi-Bo Hou, Huan Cao, Bi-Heng Liu, Yun-Feng Huang, Reevu Maity, Vlatko Vedral, Chuan-Feng Li, Guang-Can Guo and Davide Girolami

(2) *Past of a quantum particle: Common sense prevails*
Berge Englert, Kelvin Horia, Jibo Dai, Yink Loong Len and Hui Khoon Ng

(3) *Occam's Vorpal Quantum Razor: Memory reduction when simulating continuous-time stochastic processes with quantum devices*
Thomas Elliott and Mile Gu

[Parallel session] 2B

(1) *Semidefinite programming converse bounds for quantum communication*
Xin Wang, Kun Fang and Runyao Duan

(2) *Locality Preserving Logical Operators in Topological Stabiliser Codes*
Paul Webster and Stephen D. Bartlett

(3) *Self-testing of binary observables based on commutation*
Jedrzej Kaniewski

2:45pm - 3:35pm [Tutorial 2a] Charles Bennett

4:00pm - 4:50pm [Tutorial 2b] Charles Bennett

Title : Forging the culture of quantum information science

5:00pm - 6:30pm [Poster session 2]

Day 3 (Wednesday, 6 September 2017)

9:00am - 10:00am [Invited talk] Masahito Hayashi

Role of Hypothesis Testing in Quantum Information

10:30am - 12:35pm

[Parallel session] 3A

(1) *Practical round-robin-differential-phase-shift quantum key distribution*
Zhen-Qiang Yin, Shuang Wang, Wei Chen, Yunguang Han, Zheng-Fu Han and Guangcan Guo

(2) *Flow ambiguity: A path towards classically driven blind quantum computation*
Atul Mantri, Tommaso Demarie, Nicolas Menicucci and Joseph Fitzsimons

(3) *A Cost-Effective Approach for Satellite Based Quantum Key Distribution*
Alexander Lohrmann, Aitor Villar, Debashis Demunshi, Zhongkan Tang, Rakhitha Chandrasekara and Alexander Ling

(4) *Lorentz invariant entanglement distribution for the space-based quantum network*
Tim Byrnes, Batyr Ilyas, Louis Tessler, Masahiro Takeoka, Segar Jambulingam, Jonathan P. Dowling

(5) *Efficient classical verification of quantum computations*
Richard Jozsa and Sergii Strelchuk

[Parallel session] 3B

(1) *Quantum Sphere-Packing Bounds with Polynomial Prefactors*
Hao-Chung Cheng, Min-Hsiu Hsieh and Marco Tomamichel

(2) *Verifiable fault-tolerance in measurement-based quantum computation*
Keisuke Fujii and Masahito Hayashi

(3) *Bayesian Quantum Noise Spectroscopy*
Christopher Ferrie, Chris Grande, Gerardo Paz-Silva and Howard Wiseman

(4) *Moderate Deviations for Classical-Quantum Channels*
Hao-Chung Cheng and Min-Hsiu Hsieh

(5) *Analog quantum error correction with encoding a qubit into an oscillator*
Kosuke Fukui, Akihisa Tomita and Atsushi Okamoto

Day 4 (Thursday, 7 September 2017)

9:00am - 10:00am [Invited talk] Zhengfeng Ji

Entanglement in interactive proof systems

10:30am - 11:00am [Long talk]

Irreversibility of Asymptotic Entanglement Manipulation Under PPT-preserving Operations

Xin Wang and Runyao Duan

11:00am - 11:30am [Long talk]

Non-asymptotic entanglement distillation

Kun Fang, Xin Wang, Marco Tomamichel and Runyao Duan

11:30am - 12:00pm [Long talk]

Superadditivity of the classical capacity with limited entanglement assistance

Elton Yechao Zhu, Quntao Zhuang and Peter Shor

1:30pm - 2:45pm

[Parallel session] 4A

(1) *The quantum monad on relational structures*

Samson Abramsky, Rui Soares Barbosa, Nadish de Silva and Octavio Zapata

(2) *Converting multilevel nonclassicality into genuine multipartite entanglement*

Bartosz Regula, Marco Piani, Marco Cianciaruso, Thomas Bromley, Alexander Streltsov and Gerardo Adesso

(3) *Simultaneous hollowisation separability criterion in general multipartite systems*

Antoine Neven and Thierry Bastin

[Parallel session] 4B

(1) *Towards high-dimensional entanglement-based quantum communication in space*

Fabian Steinlechner, Sebastian Ecker, Matthias Fink, Bo Liu, Oliver Devries, Jessica Bavaresco, Marcus Huber, Erik Beckert, Thomas Scheidl and Rupert Ursin

(2) *Physical-depth architectural requirements for generating universal photonic cluster states*

Sam Morley-Short, Sara Bartolucci, Mercedes Gimeno-Segovia, Pete Shadbolt, Hugo Cable and Terry Rudolph

(3) *Quantum simulation of the quantum Rabi model in a single trapped ion*

Dingshun Lv, Shuoming An, Zhenyu Liu, Jingning Zhang, Julen Simon Pedernales, Lucas Lamata, Enrique Solano and Kihwan Kim

---

2:45pm - 3:45pm [Invited talk] Jian-Qiang You

Holonomic surface codes for fault-tolerant quantum computation

4:15pm - 5:15pm [Invited talk] William Slofstra

Group theory and non-local games

Day 5 (Friday, 8 September 2017)

9:00am - 10:00am [Invited talk] Henry Yuen

Classically testing the exponential nature of Hilbert space

10:30am- 11:00am [Long talk]

Fidelity of quantum strategies with applications to cryptography

Ansis Rosmanis, Jamie Sikora and Gus Gutoski

11:00am - 11:30am [Long talk]

Generalized entanglement entropies of quantum designs

Zi-Wen Liu, Seth Lloyd, Elton Yechao Zhu and Huangjun Zhu

11:30am - 12:00pm [Long talk]

No-Hypersignaling Principle

Michele Dall'Arno, Sarah Brandsen, Alessandro Tosini, Francesco Buscemi and Vlatko Vedral

---

1:30pm - 2:45pm

[Parallel session] 5A

(1) *A generalized quantum Slepian-Wolf*

Anurag Anshu, Rahul Jain and Naqueeb Ahmad Warsi

(2) *Fundamental rate-loss trade-off for the quantum internet*

Koji Azuma, Akihiro Mizutani and Hoi-Kwong Lo

(3) *Approximate broadcasting of quantum correlations*

Wei Xie, Kun Fang, Xin Wang and Runyao Duan

[Parallel session] 5B

(1) *Universal extensions of restricted classes of quantum operations*

Michal Oszmaniec and Zoltan Zimboras

(2) *Optimal quantum error correcting codes from absolutely maximally entangled states*

Zahra Raissi, Christian Gogolin, Arnau Riera and Antonio Acin

(3) *Efficient unitary designs with nearly time-independent Hamiltonian dynamics*

Yoshifumi Nakata, Christoph Hirche, Masato Koashi and Andreas Winter

---

2:45pm - 3:45pm [Invited talk] Stefanio Pironio

A semi-device-independent framework based on natural physical assumptions and its application to random number generation

The structure necessary for quantum speedups

Shalev Ben-David

University of Maryland, USA

Abstract: One of the central insights in quantum computing has been that quantum computation seems to provide exponential speedups over classical computation, but only for certain "structured" problems, such as factoring. For unstructured problems, like NP-complete problems, we do not expect an exponential quantum speedups. This raises the question: can we formalize this intuition? What types of structure suffice? In this talk, I will outline some of what we know about this problem, focusing primarily on the query complexity model due to its relative tractability. In the query complexity setting, we know that exponential speedups are not possible for total functions, but are sometimes possible when there is a promise on the input; I will describe what we know about the problem of characterizing the promises that allow exponential quantum speedups

Experimental certification of millions of genuinely entangled atoms in a solid

Florian Fröwis[1] [*] Peter C. Strassmann[1] [†] Alexey Tiranov[1] [‡] Corentin Gut[1] [§]
Jonathan Lavoie[1] [¶] Nicolas Brunner[1] [‖] Félix Bussières[1] [**] Mikael Afzelius[1] [††]
Nicolas Gisin[1] [‡‡]

[1] *Groupe de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland*

Abstract. Quantum theory predicts that entanglement can also persist in macroscopic physical systems, albeit difficulties to demonstrate it experimentally remain. Recently, genuine entanglement between up to 2900 atoms was reported. Here we demonstrate 16 million genuinely entangled atoms in a solid-state quantum memory prepared by the heralded absorption of a single photon. We develop an entanglement witness for quantifying the number of genuinely entangled particles based on the collective effect of directed emission combined with the nonclassical nature of the emitted light. The method is applicable to a wide range of physical systems and is effective even in situations with significant losses.

Keywords: Multipartite entanglement, entanglement depth, solid-state quantum memory

1 Introdution

A clear picture of large-scale entanglement with its complex structure is so far not developed. It is however important to understand the role of different facets of multipartite entanglement in nature and in technical applications [1, 2]. For example, the so-called Schrödinger cat states [3] are fundamentally different from a single photon coherently absorbed by a large atomic ensemble; even though both are instances of multipartite entanglement [4, chapter 16.5]. The theoretical study of large-scale entanglement has to be followed by an experimental demonstration, which consists of two basic steps: the preparation of an entangled system and a subsequent appropriate measurement verifying the presence of entanglement. In the context of entanglement in large systems, the preparation of entanglement is generally much simpler than its verification. For example, single-particle measurements are often not possible and collective measurements are typically restricted to certain types and are of finite resolution. These limitations call for new witnesses that allow one to certify entanglement based on accessible measurement data.

The concept of entanglement depth [5] was shown to be meaningful for and applicable to large quantum systems. It is defined as the smallest number of genuinely entangled particles that is compatible with the measured data. This allows one to witness at least one subgroup of genuinely entangled particles in a state-independent and scalable way. Large entanglement depth was successfully demonstrated with so-called spin-squeezed and oversqueezed states by measuring first and second moments of collective spin operators [6, 7, 8, 9]; lately up of 680 atoms [10]. Recently, a witness was proposed that is designed for the W state, which is a coherent superposition of a single excitation shared by many atoms [11]. Based on this witness, an entanglement depth of around 2900 was measured [12]. However, these witnesses do not detect entanglement when the vacuum component of the state is dominant [11], even though the W state is known to be quite robust against various sources of noise, in particular, against loss of particles and excitation [13]. Hence, much larger values for the entanglement depth could be expected.

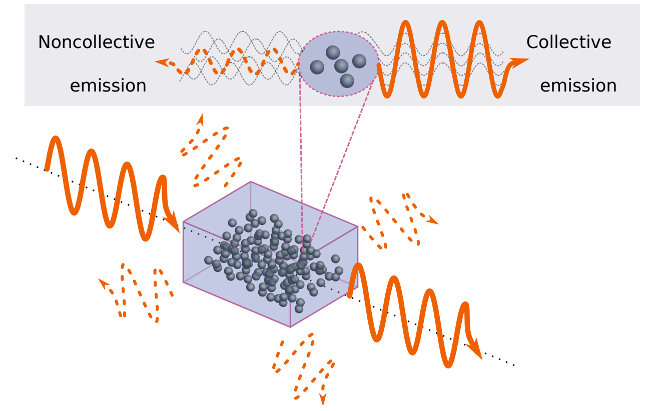


Figure 1: Basic intuition. When atoms spontaneously emit photons, phase coherence between the atoms leads to constructive interference and enhanced emission probability in a certain direction, measured by a single photon detector. Emission in any other direction is incoherent and hence not enhanced. If this phase coherence is generated by absorbing a single photon, the atoms are necessarily entangled.

Here, we present theoretical methods and experimental data that verify a large entanglement depth in a solid-state quantum memory. A rare-earth-ion-doped crystal

[*] florian.froewis@unige.ch
[†] peter.strassmann@unige.ch
[‡] alexey.tiranov@unige.ch
[§] corentin.gut@itp.uni-hannover.de
[¶] jlavoie@uoregon.edu
[‖] nicolas.brunner@unige.ch
[**] felix.bussieres@unige.ch
[††] mikael.afzelius@unige.ch
[‡‡] nicolas.gisin@unige.ch

spectrally shaped to an atomic frequency comb (AFC) is used to absorb and re-emit light at the single-photon level [14, 15, 16, 17], where at least 40 billion atoms collectively interact with the optical field. Using the measured photon number statistics of the re-emitted light we collect partial information about the quantum state of the atomic ensemble before emission. Then, we show that certain combinations of emission probabilities for one and two photons imply entanglement between a large number of atoms. With the measured data from our solid-state quantum memory we demonstrate inseparable groups of entangled particles containing at least 16 million atoms.

2 Results

Before discussing the experiment, we give an intuitive explanation for the appearance of large entanglement depth when a large atomic ensemble coherently interacts with a single photon. Suppose that N two-level atoms ($|g\rangle$ and $|e\rangle$ denote ground and excited state, respectively), couple to a light field. The quantised interaction in the dipole approximation is described by [18]

$$H_{\text{int}} = \sum_{j,\vec{k}} e^{-i\vec{k}\cdot\vec{r}_j} a_{\vec{k}} \sigma_+^{(j)} + e^{i\vec{k}\cdot\vec{r}_j} a_{\vec{k}}^\dagger \sigma_-^{(j)}, \qquad (1)$$

that is, a single photon with wave vector $\vec{k}$ is annihilated by exciting atom j via $\sigma_+ |g\rangle = |e\rangle$ and vice versa. The phase is given by the scalar product between $\vec{k}$ and the position $\vec{r}_j$ of the atom. When an incoming light field is absorbed via interaction (1), the imprinted phase relation between the atoms serves as a memory for the direction and the energy of the absorbed photons. Without this information, a spontaneous, directed re-emission is not possible. In other words, phase coherence between the atoms is necessary in order to a have well-controlled re-emission direction [19, 20]. Now, depending on the nature of the absorbed light, this coherence implies entanglement between the atoms or not. On the one hand, the absorption of a coherent state leads to a coherent atomic state, which is unentangled [4, chapter 16.7]. On the other hand, if a single photon $|1\rangle$ is absorbed, the quantisation of the field leads to a W state (or Dicke state with a single excitation) of the atomic state [4, chapter 16.5]

$$|1\rangle \to |D_1\rangle \propto \sum_j e^{-i\vec{k}\cdot\vec{r}_j} |g\dots ge_jg\dots g\rangle . \qquad (2)$$

Then, the ensemble is genuinely multipartite entangled [13]. These examples suggest a generic relation between directed emission, single-photon character of the emitted light and large entangled groups.

In our experiment, we use a neodymium-based solid-state quantum memory operating at a total read-write efficiency of 7%. This memory was demonstrated to be capable of storing different types of photonic states and preserving state properties such as the single-photon character [15, 17, 22, 23, 24]. A heralded single photon is produced via spontaneous parametric down conversion [25] and coupled to the atomic ensemble, which was prepared in the ground state $|D_0\rangle = |g\rangle^{\otimes N}$. After a 50 ns delay time, the coherent excitation is spontaneously re-emitted in forward direction and detected. In practice, this optical state is not exactly a single photon. Due to losses at different levels, the state contains a large vacuum component. Also higher photon components are present. However, since directed emission and non-classical photon number statistics are largely preserved, entanglement between large groups of atoms is expected.

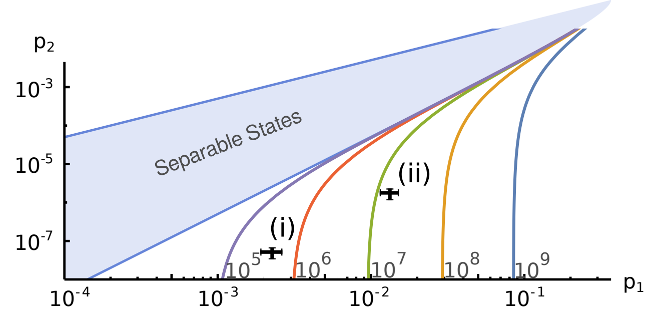


Figure 2: Lower bounds on the entanglement depth K for $N \approx 4.0(1) \times 10^{10}$ [21]. Entanglement is required to reach small p_2 while keeping p_1 constant. The number next to a colored line is the minimal K that is compatible with data points on this line. The two black crosses are data points from the experiment including one standard deviation: (i) raw data, (ii) taking detector noise into account (cp. table 1).

In order to certify this entanglement, we develop the following entanglement witness. Suppose a pure state that is subdivided into a product of M groups

$$|\psi\rangle = |\phi_1\rangle \otimes \dots \otimes |\phi_M\rangle , \qquad (3)$$

where the $|\phi_i\rangle$ are arbitrary. Phase coherence between the groups imply that each group has to carry some excitation. This necessarily amounts to an emission spectrum that also contains multi-photon components.

To be more specific, we consider the probabilities of the atoms emitting one and two photons, p_1 and p_2, respectively. In the low-excitation limit, these probabilities correspond to $p_1 = |\langle D_1|\,\psi\rangle|^2$ and $p_2 = |\langle D_2|\,\psi\rangle|^2$, where

$$|D_2\rangle \propto \sum_{j<l} e^{-i\vec{k}\cdot(\vec{r}_j+\vec{r}_l)} |g\dots ge_jg\dots ge_lg\dots g\rangle , \qquad (4)$$

that is, the phase-coherent superposition of two excitations. As shown in Ref. [21], it is possible to find the minimal p_2 for a given p_1 within the class (3) with fixed M. By varying p_1 and M one finds a lower bound on p_2 as a function of p_1 and M. Given the linearity of p_1 and p_2 when mixing states like in Eq. (3) (with arbitrary grouping but lower-bounded M), the extension of the bound to mixed states is straightforward. Comparing the lower bounds with experimental data in turn gives an upper bound on M and, by additionally measuring N, a lower bound on the entanglement depth, which simply reads $K = N/M$.

Experimentally, p_1 is obtained from the probability to measure a single re-emitted photon in the forward

Table 1: Results for entanglement depth K for (i) the raw data and (ii) the data after the detector noise has been subtracted out (cf. Fig. 2). By sampling p_1, p_2 and N around the measured values within the estimated uncertainties, we calculate the expected entanglement depth K. The values in the last columns are lower bounds on K with confidence $3\sigma = 99.7\%$.

Level of modeling	p_1	K	$K - 3\sigma$
(i) raw data	0.0023(3)	4.76×10^5	7.54×10^4
(ii) after re-mission	0.013(2)	1.64×10^7	3.72×10^6

mode heralded by the detection of the idler photon at the source. The value of p_2 corresponds to the two-photon statistics and is inferred from the measured autocorrelation function $g^{(2)}_{ss|i} = 2p_2/p_1^2$ and p_1. From the raw data, we find $p_1 = 2.3(3) \times 10^{-3}$ and $p_2 = 5(2) \times 10^{-8}$. The relatively small value of p_1 is a product of the efficiencies of the source, the memory and detectors. By taking the detectors inefficiencies into account, we can estimate the real values of p_1, p_2, which are higher.

A key element in the experiment is the high-precision measurement of N. The ratio of the coherent emission in the forward direction and the incoherent emission in the backward direction is a lower bound on the number of resonant atoms [20]. Since incoherent emission from single photons is much lower than detector dark counts, the single photon source is replaced by a bright coherent state for this measurement and we find $N \geq 4.0(1) \times 10^{10}$ [21]. The resulting K is presented in Fig. 2 and table 1 for the raw data and detector-noise-subtracted data.

3 Discussion

This work demonstrates that large entanglement depth is experimentally certifiable even with atomic ensembles beyond 10^{10} atoms and low detection and re-emission efficiencies. We prove that entanglement between many atoms is necessary for the functioning of quantum memories that are based on collective emission, because the combination of directed emission (i.e., high memory efficiency) and preservation of the single-photon character imply large entanglement depth.

Our results further illustrate the fundamental difference between various manifestations of large entanglement. The scales at which we observe entanglement depth seem to be completely out of reach for other types of large entanglement, such as Schrödinger-cat states [2, 26].

We report lower bounds on the minimal number of genuinely entangled atoms, which should not be confused with quantifying entanglement with an entanglement measure. Indeed, the nature of the target state, the W state $|D_1\rangle$, and the experimental challenges suggest that only a small amount of entanglement is present in the crystal during the storage.

References

[1] T. D. Ladd, *et al.*, Nature **464**, 45, 2010.

[2] L. Pezzè, *et al.*, *arXiv:1609.01609*, 2016.

[3] E. Schrödinger, Naturwissenschaften **23**, 807, 1935.

[4] L. Mandel and E. Wolf, *Optical Coherence And Quantum Optics*. Cambridge: Cambridge University Press, 1995.

[5] A. S. Sørensen and K. Mølmer, Phys Rev Lett **86**, 4431, 2001.

[6] M. F. Riedel, *et al.*, Nature **464**, 1170, 2010.

[7] C. Gross, *et al.*, Nature **464**, 1165, 2010.

[8] G. Vitagliano *et al.*, Phys Rev A **89**, 032307, 2014.

[9] B. Lücke *et al.*, Phys Rev Lett **112**, 155304, 2014.

[10] O. Hosten, *et al.*, Nature **529**, 505, 2016.

[11] F. Haas, *et al.*, Science **344**, 180, 2014.

[12] R. McConnell, *et al.*, Nature **519**, 439, 2015.

[13] J. K. Stockton, *et al.*, Phys Rev A **67**, 2, 022112, 2003.

[14] M. Afzelius, *et al.*, Phys Rev A **79**, 052329, 2009.

[15] C. Clausen, *et al.*, Nature **469**, 508, 2011.

[16] E. Saglamyurek, *et al.*, Nature, **469**, 512, 2011.

[17] C. Clausen, *et al.*, Phys Rev Lett **108**, 190503, 2012.

[18] R. Loudon, *The Quantum Theory of Light*. Oxford; New York: Oxford University Press, 2000.

[19] L.-M. Duan, *et al.*, Nature **414**, 413, 2001.

[20] M. O. Scully, *et al.*, Phys Rev Lett **96**, 010501, 2006.

[21] F. Fröwis, *et al.*, *arXiv:1703.04704*, 2017.

[22] F. Bussières, *et al.*, Nat Photon **8**, 775, 2014.

[23] A. Tiranov, *et al.*, Optica **2**, 279, 2015.

[24] A. Tiranov, *et al.*, Phys Rev Lett **117**, 240506, 2016.

[25] C. Clausen, *et al.*, New Journal of Physics **16**, 9, 093058, 2014.

[26] F. Fröwis, J Phys A: Math Theor **50**, 11, 114003, 2017.

Quantum non-malleability and authentication
(extended abstract)

Gorjan Alagic and Christian Majenz

QMATH, Department of Mathematical Sciences
University of Copenhagen

galagic@gmail.com majenz@math.ku.dk

1. Introduction.

Quantum cryptography has grown to be an important subfield of quantum information science. While the most well-known results (like QKD) concern the use of quantum information for classical cryptography, there has also been an increased interest in extending the framework of symmetric-key cryptography to the encryption of quantum data.

In its most basic form, encryption ensures the secrecy of transmissions against eavesdroppers. Besides secrecy, another desirable property is *non-malleability*, which guarantees that an adversary cannot meaningfully modify the plaintext by manipulating the ciphertext. In the classical setting, secrecy and non-malleability are independent: there are schemes which satisfy secrecy but are malleable, and schemes which are non-malleable but transmit the plaintext in the clear.

In the setting of quantum information, encryption is the task of transmitting quantum states over a completely insecure quantum channel. Information-theoretic secrecy for quantum encryption is well-understood. Moreover, significant progress has been made on more advanced constructions, such as authenticated encryption [5], quantum fully-homomorphic encryption [8], and many more. Despite this high-level progress, a basic aspect of quantum encryption remains largely unstudied. Indeed, *quantum non-malleability* was considered in only one previous work, by Ambainis, Bouda and Winter [4]. Their definition (which we call ABW-NM) requires secrecy, and that the "effective channel" $\mathsf{Dec} \circ \Lambda \circ \mathsf{Enc}$ of any adversary Λ is trivial[1].

Unlike non-malleability, the closely-related subject of quantum authentication (where decryption is allowed to reject) has received significant attention (see, e.g., [1, 5, 7, 9, 10].) In this setting, there are two definitions. The widely-adopted definition of Dupuis, Nielsen and Salvail (DNS-authentication) asks that, regardless of whether decryption rejects, the average effective channel of any adversary does not touch the plaintext [9]. A more recent definition of Garg, Yuen and Zhandry (GYZ-authentication) asks that, in the accept case, the adversary does not touch the plaintext with high probability over the key (rather than on average) [10].

2. Summary of results.

In this work, we devise a new definition of quantum non-malleability (denoted NM). We prove several new results about quantum non-malleability, quantum authentication, and the connections between these two concepts. A summary of our results is as follows; we will focus on the exact case, but all the definitions and results have appropriate relaxations to the approximate setting; see the full paper for details [2].

[1] More precisely, it is either the identity or replacement by a fixed state (or some combination of the two).

2.1. New definition. We give a new definition of quantum non-malleability (NM), which improves on ABW-NM in a number of ways:

1. it is expressed in terms of entropic quantities, generalizing classical definitions [11];
2. it can be alternatively characterized in terms of the effective attack (Theorem 5) for practical security guarantees
3. it prevents more powerful attacks, which make use of side information about the plaintext;
4. it is immune to a devastating "plaintext injection" attack, whereby an adversary against an ABW-NM scheme can send a plaintext of their choice to the receiver;
5. it does not require secrecy; instead, we show *quantum non-malleability implies quantum secrecy.*

The last point is analogous to the fact that quantum authentication implies encryption [5].

Informally, our definition states that any attack on the ciphertext will result in no information gain – except via a "trivial attack" which is always possible against any scheme. In this trivial attack, the adversary simply decides whether or not to destroy the ciphertext, and remembers that choice in their side information. A formal definition is as follows. The relevant quantum registers are: plaintext A, ciphertext C, user's reference R, and adversary's side information B.

Definition 1 *A scheme is non-malleable (NM) if for any ϱ_{ABR} and any attack $\Lambda_{CB\to C\tilde{B}}$, the effective channel $\tilde{\Lambda}_{AB\to A\tilde{B}} = \mathsf{Dec}\circ\Lambda\circ\mathsf{Enc}$ satisfies*

$$I(AR:\tilde{B})_{\tilde{\Lambda}(\varrho)} \leq I(AR:B)_{\varrho} + h(p_{=}(\Lambda,\varrho)).$$

The binary entropy term $h(p_{=}(\Lambda,\varrho))$ captures the information gain of the aforementioned trivial attack. Formally, $p_{=}(\Lambda,\varrho) = F\left(\mathrm{Tr}_{\tilde{B}}\Lambda((\cdot)_C\otimes\varrho_B)\right)^2$ is the squared entanglement fidelity of the attack map as it acts on the ciphertext register if ϱ_B is input in the side information register. Theorem 5 below provides an alternative way of defining NM, that is more suitable for practical security definitions. It can be understood in the ”real vs. ideal" framework: for NM schemes, the real effective attack is equal to the ideal, trivial, linear combination of the identity and a fixed constant channel.

2.2. New results on non-malleability. As mentioned above, our first result is that NM implies secrecy. Here, secrecy stands for one of a number of equivalent notions of security; one may for instance use analogues of IND [6] or SEM [3] for computationally unbounded adversaries.

Theorem 2 *If a quantum encryption scheme is non-malleable (NM), then it is also secret.*

We remark that this is a significant departure from the classical case, where secrecy and non-malleability are independent properties.

Next, we show that NM implies ABW-NM, and give a separation scheme which is secure under ABW-NM but insecure under NM. As described in [2], this scheme is susceptible to a powerful attack, whereby a simple adversary can freely choose the output of decryption.

Theorem 3 *If a quantum encryption scheme is NM, then it is also ABW-NM.*

On the other hand, if we restrict our attention to schemes where the encryption maps are unitary, then we are able to show the following.

Theorem 4 *Let Π be a quantum encryption scheme such that encryption E_k is unitary for all keys k. Then Π is NM if and only if $\{E_k\}_k$ is a two-design.*

Together with the results of [4], this implies that NM and $\mathsf{ABW\text{-}NM}$ are in fact equivalent for unitary schemes. Finally, we can also characterize NM schemes in the general case, as follows.

Theorem 5 *A scheme is* NM *if and only if, for any* $\Lambda_{CB\to C\tilde{B}}$, *there exist maps* $\Lambda'_{B\to\tilde{B}}$, $\Lambda''_{B\to\tilde{B}}$ *such that the effective attack* $\tilde{\Lambda}_{AB\to A\tilde{B}}$ *has the form*

$$\tilde{\Lambda} = \mathrm{id}_A \otimes \Lambda' + \frac{1}{|C|^2 - 1} \left(|C| \left\langle D_K(\mathbb{1}_C)\right\rangle - \mathrm{id}\right)_A \otimes \Lambda'' \,.$$

The maps Λ' and Λ'' are explicit functions of Λ [2]. This theorem shows that our notion provides *ciphertext non-malleability*: if the ciphertext is modified, the plaintext is replaced by $D_K(\mathbb{1}_C)$.

2.3. New results on quantum authentication. The techniques we developed for quantum non-malleability also yield several new results on quantum authentication, as follows. We note that our definitions of authentication deviate slightly from the original versions [9, 10], in that decryption outputs a reject symbol in place of the plaintext (rather than setting a flag to "reject.")

First, we show how to build authentication from non-malleability. Given an encryption scheme $\Pi = \{E_k\}$, we define Π_t^{tag} to be a new scheme whose encryption is $\varrho \mapsto E_k\big(\varrho_A \otimes |0\rangle\langle 0|_B^{\otimes t}\big)E_k^\dagger$, and whose decryption rejects unless B measures to $|0^t\rangle$.

Theorem 6 *If a scheme* $\Pi = \{E_k\}$ *satisfies* NM, *then* Π_t^{tag} *is* 2^{2-t}-DNS-*authenticating.*

If the starting NM scheme is encryption via the Clifford group, then the result is the well-known Clifford scheme for authentication [1].

Next, we show that GYZ-authentication implies DNS-authentication.

Theorem 7 *If a scheme is* ε-GYZ-*authenticating, then it is also* $O(\sqrt{\varepsilon})$-DNS-*authenticating.*

This result is technically non-trivial: on one hand, GYZ requires high probability of success while DNS only needs success-on-average; on the other hand, GYZ requires nothing in the reject case while DNS still makes rather stringent demands.

Finally, we show that GYZ-authentication can be satisfied by a scheme which "tags" plaintexts as before, and encrypts with a unitary 2-design. This is a significant improvement over the analysis of [10], which required eight-designs for the same construction.

Theorem 8 *Let* $\Pi = \{E_k\}_k$ *be a* 2^{-t}-*approximate 2-design scheme. Then* Π_t^{tag} *is* $2^{-\Omega(t)}$-GYZ-*authenticating.*

Given the conclusions of Theorem 4, we may state this as follows: if a unitary scheme Π is non-malleable, then Π_t^{tag} is GYZ-authenticating. We remark that the simulation of adversaries in this proof is efficient, in the sense of [7].

3. Conclusion and open problems.

In this work, we introduced a new definition of quantum non-malleability, a core concept in encryption. Our notion addresses a major vulnerability in the previous definition, and can serve as a primitive for constructing authentication schemes. When using unitary 2-designs (e.g., the Clifford group) for non-malleable encryption, the resulting authentication schemes are secure under the strongest known definitions [10]. We remark that our work is also a natural starting point for future research on quantum non-malleability in the setting of many messages and computational security assumptions.

Bibliography

[1] Dorit Aharonov, Michael Ben-Or, and Elad Eban. Interactive proofs for quantum computations. In *Innovations in Computer Science - ICS 2010, Tsinghua University, Beijing, China, January 5-7, 2010. Proceedings*, pages 453–469, 2010. URL http://conference.itcs.tsinghua.edu.cn/ICS2010/content/papers/35.html.

[2] Gorjan Alagic and Christian Majenz. Quantum non-malleability and authentication. *arXiv preprint arXiv:1610.04214, to appear in Advances in Cryptology - CRYPTO 2017*, 2016.

[3] Gorjan Alagic, Anne Broadbent, Bill Fefferman, Tommaso Gagliardoni, Christian Schaffner, and Michael StJules. Computational security for quantum encryption. In *9th International Conference on Information Theoretic Security (ITICS), to appear.*, 2016.

[4] Andris Ambainis, Jan Bouda, and Andreas Winter. Nonmalleable encryption of quantum information. *Journal of Mathematical Physics*, 50(4):042106, 2009.

[5] Howard Barnum, Claude Crépeau, Daniel Gottesman, Adam Smith, and Alain Tapp. Authentication of quantum messages. In *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*, pages 449–458. IEEE, 2002.

[6] Anne Broadbent and Stacey Jeffery. *Quantum Homomorphic Encryption for Circuits of Low T-gate Complexity*, pages 609–629. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. ISBN 978-3-662-48000-7. doi: 10.1007/978-3-662-48000-7_30. URL http://dx.doi.org/10.1007/978-3-662-48000-7_30.

[7] Anne Broadbent and Evelyn Wainewright. Efficient simulation for quantum message authentication. *arXiv preprint arXiv:1607.03075*, 2016.

[8] Y. Dulek, C. Schaffner, and F. Speelman. Quantum homomorphic encryption for polynomial-sized circuits. *ArXiv e-prints*, March 2016.

[9] Frédéric Dupuis, Jesper Buus Nielsen, and Louis Salvail. Actively secure two-party evaluation of any quantum operation. In *Advances in Cryptology–CRYPTO 2012*, pages 794–811. Springer, 2012.

[10] Sumegha Garg, Henry Yuen, and Mark Zhandry. New security notions and feasibility results for authentication of quantum data. *arXiv preprint arXiv:1607.07759*, 2016.

[11] Akinori Kawachi, Christopher Portmann, and Keisuke Tanaka. Characterization of the relations between information-theoretic non-malleability, secrecy, and authenticity. In *International Conference on Information Theoretic Security*, pages 6–24. Springer, 2011.

Dining Philosophers, Leader Election and Ring Size problems, in the quantum setting

Dorit Aharonov[1] * Maor Ganz[1] † Loïck Magnin[2] ‡

[1] *The Hebrew University*
[2] *Pure Storage*

Abstract. We provide the first quantum protocol for the Dining Philosophers problem (DP), a central problem in distributed algorithms, and use it to provide a new quantum protocol for the tightly related problem of exact leader election, improving significantly over Tani et al [TKM12]. These two problems are related to another important problem in distributed algorithms, the ring size problem, and interesting connections are discussed. The results raise several interesting open questions.

Keywords: AQIS, Dining philosophers, Leader election, Ring size problem, Symmetry breaking, Quantum algorithms, Distributive algorithms

Extended Abstract

Before stating the results, we provide background on the problems:

Dining philosophers problem The DP problem was first introduced by [Dij71] and is one of the central problems in distributed algorithms; it is heavily related to many synchronization problems, memory allocations and semaphores mechanisms. It is defined as follows. A group of n philosophers are sitting (and thinking) around a circular table in a Chinese restaurant. Between each pair of philosophers there is a chopstick (a total of n). As time passes by, a philosopher might get hungry. In order for a hungry philosopher to eat, he must hold both chopsticks (to his right, and to his left). A philosopher can only pick up one chopstick at a time, and obviously cannot pick up a chopstick which is already in the hand of a neighbor. The only communication allowed, is message sending between adjacent philosophers (neighbors). Our goal is to find an algorithm (each of the philosophers is identical (no ID) and runs the same algorithm - anonymous) s.t. every hungry philosopher will eventually eat, and to try and minimize the communication complexity, as well as the running time. In most of the paper we assume that n, the number of philosophers, is known in advance to all philosophers (or at least they know an upper bound for it). It was shown in [LR81] that no deterministic classical algorithm can solve the DP problem, even if we only want to ensure that *one* hungry philosopher will eventually eat. However, randomly, this is possible; and only $O(1)$ (classical) memory is required for each philosopher.

Fair leader election problem It turns out that the DP question is tightly related to the problem of leader election (LE). In this problem, we have a set of n identical parties, who want to elect a leader among themselves. Each party should have the same probability to be elected. Again each party is anonymous and runs the same algorithm.

A deterministic algorithm always ends in finite time and results in the election of a single leader. It is easy to show that LE, like the DP problem, cannot be done classically in a deterministic way, because LE implies DP. What about a probabilistic solution?

A randomized LE protocol is allowed to never end or to not elect a leader with some probability (preferably as small as possible), but is not allowed to elect more than one leader. When n is known to the players, or even when some bound is known, then indeed there exists a randomized solution to the LE problem [IR90].

It is based on the following idea: each player randomly chooses a number in some finite range which depends on n, and the one with the biggest number gets elected. The players use their knowledge about n to make sure that there is only a single biggest number.

In the quantum setting, a surprising result due to Tani et. al [TKM12] provides a quantum algorithm which solves the *exact* LE problem. Again, this assumes that some bound on n is known in advance.

The ring size problem The question of whether n is known to the players or not, plays an important role in the context of the LE and DP problems. A related question is thus the *ring size problem.* It is defined to be the problem of finding n, the size of the ring, under the same anonymous conditions. Classically, the ring size also cannot be deterministically determined even if we have an upper bound on n beforehand ([ASW88], [IR90]), but [IR90] showed that it can be found with small error probability if we know bounds on the ring size: $N \leq n < 2N$. [IR90] also proved that the ring size problem can be solved (even when no bound on n is known beforehand) with arbitrarily small error.

Unfortunately, the probabilistic algorithm for the ring size problem with error cannot be used to resolve the LE problem even with error, because due to the error, the possibility of choosing two leaders cannot be ruled out. Indeed, we know LE to be impossible even in the probabilistic setting, if no bound is known on n, but the

*dorit.aharonov@gmail.com
†maor.ganz@mail.huji.ac.il
‡loick@e-magnin.com

randomized ring size problem can be solved in this setting.

Our results To the best of our knowledge, neither the DP nor the ring size problem were investigates in the quantum setting before. Hence the current paper is the first time they are discussed in this context.

We first show:

Theorem 1 ***Existence of an exact DP quantum protocol.*** *(roughly) There exists a quantum protocol for the exact DP problem, in the setting in which n is known or at least a bound on it is known.*

This follows from the following lemma, which shows that one can derive a DP algorithm from a LE algorithm, by a simple classical reduction.

Lemma 2 ***LE-to-DP.*** *(roughly) The existence of a LE protocol implies that of a DP - inheriting its properties [exact / random], with an addition $O(n)$ time.*

We can thus use the Quantum algorithm for exact LE [TKM12] to give the above first result. Unfortunately, this solution for the DP problem inherits its parameters from the quantum solution. Moving forward, we can prove that a solution exists which is much more efficient in terms of memory, and is also just linear in time complexity and communication complexity. To this end, we reduce the DP problem, to the problem of breaking the symmetry, namely, dividing the parties to two non-trivial groups:

Lemma 3 ***(7)*** *(roughly) Given a symmetry breaking protocol, there exists a solution to the exact DP problem.*

The protocol of [TKM12] offers a way to break the symmetry; we find an improved way to achieve symmetry breaking. This leads to a more efficient algorithm for exact DP than our first algorithm:

Theorem 4 ***(3) Efficient exact DP quantum protocol.*** *(roughly) There exists a deterministic quantum protocol to the DP problem, when n or an upper bound on n is known, which uses $O(1)$ quantum memory and $O(\log n)$ classical memory per philosopher and $O(n^2)$ time complexity.*

Perhaps one can also prove an implication in the other direction, DP-to-LE? As we show, this does hold in the exact (quantum) case when n is known. Thus, the two problems are equivalent in the quantum setting, whereas interestingly, in the classical randomized setting they are not.

We thus make use of the following lemma:

Lemma 5 ***(4) DP-to-LE in the exact case.*** *Given a protocol that solves the DP problem deterministically, when n is known, one can solve the exact LE problem on a ring.*

The proof idea is to use rounds of DP as black box, advancing eating philosophers to higher rounds, while eliminating the others.

By plugging-in our own efficient protocol of DP, we actually get a significant improvement on the best previously known quantum algorithm for LE on the ring (again, when n or a bound on it is known), and with only $O(1)$ memory:

Theorem 6 ***(10) A new and more efficient quantum protocol for exact LE on a ring.*** *There exists a deterministic quantum LE algorithm on a ring of a known size n with $O(n^2 \log n)$ time, $O(1)$ quantum memory and $O(\log n)$ classical memory per philosopher, and total classical communication complexity of $O(n^2 \log n)$, and quantum communication complexity of $O(n \log n)$.*

If only a bound N on n is known, then the algorithm uses instead $O(N^2 \cdot n)$ time complexity, $O(1)$ quantum memory and $O(\log N)$ classical memory per philosopher, and total quantum bit communication complexity of $O(N^2)$ and classical bits communication complexity of $O(N^2 \cdot n)$.

Open questions What about when n is unknown? It is long known ([IR90], [AW04]) that LE is classically impossible in this case, even for a randomized algorithm. The idea is that one can build from a successful run on a ring with n parties, a run on a ring with $2n$ parties, by mirroring the parties. This yields two leaders, which is strictly forbidden. If there exists a quantum protocol for the exact ring size problem, this would lead to exact quantum algorithms for these problems. As we recall, a *randomized* solution for DP when n is unknown does exist; However like for LE, an exact protocol for DP is also not known.

An important related open question is whether these exact protocols can be achieved with $O(1)$ *total* (classical and quantum) memory. This is related to the following seemingly basic question: Is there a constant depth, translation invariant quantum circuit over constant dimensional particles on a circle, that can break symmetry exactly? We conjecture that the answer is no; Proving impossibility would be interesting, and possibly related to better understanding of quantum states emerging in translationally invariant quantum systems.

A Complete paper

A full version of the paper can be found online, quant-ph arXiv 1707.01187. The additional numbers in parentheses next to some of the lemmas and theorems are the numbers of these lemmas and theorems as they appear in the full version. We have added those in the cases in which these numbers do not match the numbers in this abstract, for convenience of reading.

References

[ASW88] Hagit Attiya, Marc Snir, and Manfred K. Warmuth, *Computing on an anonymous ring*, J. ACM **35** (1988), no. 4, 845–875.

[AW04] Hagit Attiya and Jennifer Welch, *Distributed computing: Fundamentals, simulations and advanced* *topics (2nd edition)*, John Wiley Interscience, March 2004.

[Dij71] Edsger W. Dijkstra, *Hierarchical ordering of sequential processes*, Acta Inf. **1** (1971), 115–138.

[IR90] Alon Itai and Michael Rodeh, *Symmetry breaking in distributed networks*, Inf. Comput. **88** (1990), no. 1, 60–87.

[LR81] Daniel J. Lehmann and Michael O. Rabin, *On the advantages of free choice: A symmetric and fully distributed solution to the dining philosophers problem*, Conference Record of the Eighth Annual ACM Symposium on Principles of Programming Languages, Williamsburg, Virginia, USA, January 1981 (John White, Richard J. Lipton, and Patricia C. Goldberg, eds.), ACM Press, 1981, pp. 133–138.

[TKM12] Seiichiro Tani, Hirotada Kobayashi, and Keiji Matsumoto, *Exact quantum algorithms for the leader election problem*, TOCT **4** (2012), no. 1, 1.

Characterizations of symmetrical and partial Boolean functions with exact quantum query complexity [*]

Daowen Qiu[1 †] Shenggen Zheng[1 ‡]

[1] *Institute of Computer Science Theory, School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510006, China*

Abstract. We give and prove an optimal exact quantum query algorithm with complexity $k+1$ for computing the promise problem (i.e., symmetric and partial Boolean function) DJ_n^k defined as: $DJ_n^k(x)=1$ for $|x|=n/2$, $DJ_n^k(x)=0$ for $|x|$ in the set $\{0,1,\ldots,k,n-k,n-k+1,\ldots,n\}$, and it is undefined for the rest cases, where n is even, $|x|$ is the Hamming weight of x. The case of $k=0$ is the well-known Deutsch-Jozsa problem. We outline all symmetric (and partial) Boolean functions with degrees 1 and 2, and prove their exact quantum query complexity. Then we prove that any symmetrical (and partial) Boolean function f has exact quantum 1-query complexity if and only if f can be computed by the Deutsch-Jozsa algorithm.

Keywords: exact quantum query algorithms, Deutsch-Jozsa problems, query complexity, symmetric Boolean functions, promise problems

1 General Description

The quantum query models are the quantum analog to the classical Boolean decision tree models, so they are also called *quantum decision tree* models [4] and are at least as powerful as the classical decision tree models. The implementation procedure of a quantum decision tree model is exactly a *quantum query algorithm*, and it can be roughly described as: it starts with a fixed starting state $|\psi_s\rangle$ of a Hilbert space $\mathcal{H}$ and performs a sequence of operations $U_0, O_x, U_1, \ldots, O_x, U_t$, where U_i's are unitary operators that do not depend on the input x but the query O_x does. This leads to the final state $|\psi_f\rangle = U_t O_x U_{t-1} \cdots U_1 O_x U_0 |\psi_s\rangle$. The result is obtained by measuring the final state $|\psi_f\rangle$.

A quantum query algorithm $\mathcal{A}$ *exactly computes* a Boolean function f if its output equals $f(x)$ with probability 1, for all inputs x. The exact quantum query algorithms for computing total Boolean functions also have been studied. The best known quantum speed-up was just by a factor of 2 for many years [6, 8]. In 2013, as a breakthrough result, Ambainis [2] has presented the first example of a Boolean function $f:\{0,1\}^n \to \{0,1\}$ for which $Q_E(f)=O(D(f)^{0.8675\cdots})$, where $D(f)$ denotes the minimum number of queries used by any classical deterministic query algorithm. The result was improved to nearly-quadratic separation by Ambainis *et al* [1, 3] in 2016.

However, for computing partial Boolean functions, there can be more than exponential separation [5, 9] between exact quantum and classical deterministic query complexity, and the first result was the well-known Deutsch-Jozsa algorithm [7]. Deutsch-Jozsa problem can be described as a partial Boolean function $\mathrm{DJ}_n^0 : \{0,1\}^n \to \{0,1\}$ defined as: n is even, and $\mathrm{DJ}_n^0(x)=1$ for $|x|=\frac{n}{2}$ and $\mathrm{DJ}_n^0(x)=0$ for $|x|=0$ or n, and the other cases are undefined, where $|x|$ is the Hamming weight of x.

Different from partially symmetric Boolean functions in literature, the functions DJ_n^0 and DJ_n^1 above are both symmetric and partial, so, they may be termed as *symmetrically partial Boolean functions* (i.e. promise problems) in this paper and the exact definition can be described as follows.

Definition 1 *Let $f:\{0,1\}^n \to \{0,1\}$ be a partial Boolean function, and let $D \subseteq \{0,1\}^n$ be its domain of definition. If for any $x \in D$ and for any $y \in \{0,1\}^n$ with $|x|=|y|$, it holds that $y \in D$ and $f(x)=f(y)$, then f is called a* symmetrically partial Boolean function. *When $D=\{0,1\}^n$, f is a symmetric function.*

So, a Boolean function is symmetrically partial if and only if it is symmetric and partial. In this paper, we give and prove an optimal exact quantum query algorithm with complexity $k+1$ for computing the symmetrically partial Boolean functions DJ_n^k defined as: $\mathrm{DJ}_n^k(x)=1$ for $|x|=n/2$, $\mathrm{DJ}_n^k(x)=0$ for $|x|$ in the set $\{0,1,\ldots,k,n-k,n-k+1,\ldots,n\}$, and it is undefined for the rest cases, where n is even, $|x|$ is the Hamming weight of x. The case of $k=0$ is the well-known Deutsch-Jozsa problem [7].

We outline all symmetrically partial Boolean functions with degrees 1 and 2, and prove their exact quantum query complexity. Then we prove that any symmetrically partial Boolean function f has exact quantum 1-query complexity if and only if f can be computed by the Deutsch-Jozsa algorithm.

We also discover the optimal exact quantum 2-query complexity for distinguishing between inputs of Hamming weight $\{\lfloor n/2 \rfloor, \lceil n/2 \rceil\}$ and Hamming weight in the set $\{0,n\}$ for all odd n. In addition, a method is provided to determine the degree of any symmetrically partial Boolean function.

[*] The full paper for this abstract can be found in arXiv:1603.06505. This work was supported by the National Natural Science Foundation of China (Nos. 61572532, 61272058, 61602532) and the Fundamental Research Funds for the Central Universities of China (Nos. 17lgjc24, 161gpy43).

[†] `issqdw@mail.sysu.edu.cn, Corresponding author`

[‡] `zhengshg@mail2.sysu.edu.cn, Corresponding and also co-first author`

2 A list of our main results

A general generalization of Deutsch-Jozsa problem is the following symmetrically partial function:

$$\mathrm{DJ}_n^k(x) = \begin{cases} 1 & \text{if } |x| = n/2, \\ 0 & \text{if } |x| \leq k \text{ or } |x| \geq n-k, \end{cases} \quad (1)$$

where n is even and $0 \leq k < n/2$. When $k = 0$, it is the Deutsch-Jozsa problem, and when $k = 1$, it equals the problem given by Montanaro *et al* [10]. Our first main result is as follows.

Theorem 2 *The exact quantum query complexity of* DJ_n^k *satisfies:*

$$Q_E(\mathrm{DJ}_n^k) = k + 1. \quad (2)$$

However, the classical deterministic query complexity for DJ_n^k *is:*

$$D(\mathrm{DJ}_n^k) = n/2 + k + 1. \quad (3)$$

A natural question is, what do Boolean functions with the same exact quantum query complexity have in common? Due to the importance and simplicity of symmetric functions, here we consider the case of exact quantum 1-query complexity for all symmetrical and partial functions.

Therefore, the question is what can be solved with exact quantum 1-query complexity? We can pose the question more precisely: if an exact quantum 1-query algorithm $\mathcal{A}$ computes a symmetrically partial function f, then, can any symmetrically partial function g with $Q_E(g) = 1$ be computed by $\mathcal{A}$? Our second main result answers this question as follows.

Theorem 3 *Any symmetrical and partial Boolean function f has $Q_E(f) = 1$ if and only if f can be computed by the Deutsch-Jozsa algorithm.*

To prove the above theorem, we prove the following three results.

Theorem 4 *Let $n > 1$ and let $f : \{0,1\}^n \to \{0,1\}$ be an n-bit symmetric and partial Boolean function. Then:*

(1) $\deg(f) = 1$ *if and only if f is isomorphic to the function* $f_{n,n}^{(1)}$;

(2) $\deg(f) = 2$ *if and only if f is isomorphic to one of the functions*

$$f_{n,k}^{(1)}(x) = \begin{cases} 0 & \text{if } |x| = 0, \\ 1 & \text{if } |x| = k, \end{cases} \quad (4)$$

$$f_{n,k}^{(2)}(x) = \begin{cases} 0 & \text{if } |x| = 0, \\ 1 & \text{if } |x| = k \text{ or } |x| = k+1, \end{cases} \quad (5)$$

$$f_{n,l}^{(3)}(x) = \begin{cases} 0 & \text{if } |x| = 0 \text{ or } |x| = n, \\ 1 & \text{if } |x| = l, \end{cases} \quad (6)$$

$$f_{n}^{(4)}(x) = \begin{cases} 0 & \text{if } |x| = 0 \text{ or } |x| = n, \\ 1 & \text{if } |x| = \lfloor n/2 \rfloor \text{ or } |x| = \lceil n/2 \rceil, \end{cases} \quad (7)$$

where $n - 1 \geq k \geq \lfloor n/2 \rfloor$, *and* $\lceil n/2 \rceil \geq l \geq \lfloor n/2 \rfloor$.

Theorem 5 *Let n be even and let $f : \{0,1\}^n \to \{0,1\}$ be an n-bit symmetric and partial function. Then $Q_E(f) = 1$ if and only if f is isomorphic to one of these functions:* $f_{n,k}^{(1)}$ *and* $f_{n,n/2}^{(3)}$, *where* $k \geq n/2$.

Theorem 6 *Let n be odd and let $f : \{0,1\}^n \to \{0,1\}$ be an n-bit symmetric and partial function. Then $Q_E(f) = 1$ if and only if f is isomorphic to one of the functions* $f_{n,k}^{(1)}$, *where* $k \geq \lceil n/2 \rceil$.

References

[1] S. Aaronson, S. Ben-David, R. Kothari, *Separations in query complexity using cheat sheets*, In Proceedings of the 48th STOC(2016) 863–876.

[2] A. Ambainis, *Superlinear advantage for exact quantum algorithms*, SIAM J. Comput. **45** (2016) 617–631. Earlier version in STOC'13.

[3] A. Ambainis, K. Balodis, A. Belovs, T. Lee, M. Santha, J. Smotrovs, *Separations in Query Complexity Based on Pointer Functions*, In Proceedings of 48th STOC (2016) 800–813.

[4] H. Buhrman, R. de Wolf, *Complexity measures and decision tree complexity: a survey*, Theoretical Computer Science **288** (2002) 1–43.

[5] G. Brassard, P. Høyer, *An exact quantum polynomial-time algorithm for Simon's problem*, In Proceedings of the Fifth Israeli Symposium on Theory of Computing and Systems, 1997, pp. 12–23. Also arXiv:9704027.

[6] R. Cleve, A. Eckert, C. Macchiavello, M. Mosca, *Quantum algorithms revisited*, In Proceedings of the Royal Society of London, 454A (1998): 339–354. Also arXiv:9708016.

[7] D. Deutsch, R. Jozsa, *Rapid solution of problems by quantum computation*, In Proceedings of the Royal Society of London, 439A (1992): 553–558.

[8] E. Farhi, J. Goldstone, S. Gutmann, M. Sipser, *A limit on the speed of quantum computation in determining parity*, Physical Review Letters **81** (1998) 5442–5444. Also arXiv:9802045.

[9] J. Gruska, D.W. Qiu, S.G. Zheng, *Generalizations of the distributed Deutsch-Jozsa promise problem*, Mathematical Structures in Computer Science **27** (2017) 311–331. Also arXiv:1402.7254.

[10] A. Montanaro, R. Jozsa, G. Mitchison, *On exact quantum query complexity*, Algorithmica **419** (2015) 775–796. Also arXiv:1111.0475.

Quantum algorithm for linear differential equations with exponentially improved dependence on precision

Dominic W. Berry[1] * Andrew M. Childs[2 3] † Aaron Ostrander[4] Guoming Wang[4]

[1] *Department of Physics and Astronomy, Macquarie University*
[2] *Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland*
[3] *Joint Center for Quantum Information and Computer Science, University of Maryland*
[4] *Department of Physics, University of Maryland*

Abstract. We present a quantum algorithm for systems of (possibly inhomogeneous) linear ordinary differential equations with constant coefficients. The algorithm produces a quantum state that is proportional to the solution at a desired final time. The complexity of the algorithm is polynomial in the logarithm of the inverse error, an exponential improvement over previous quantum algorithms for this problem. Our result builds upon recent advances in quantum linear systems algorithms by encoding the simulation into a sparse, well-conditioned linear system that approximates evolution according to the propagator using a Taylor series. Unlike with finite difference methods, our approach does not require additional hypotheses to ensure numerical stability. The full version of this work is given in Ref. [1].

Keywords: quantum algorithm, complexity, differential equations

1 Introduction

One of the original motivations for developing a quantum computer was to efficiently simulate Hamiltonian dynamics, i.e., differential equations of the form $\frac{d\vec{x}}{dt} = A\vec{x}$ where A is anti-Hermitian. Given a suitable description of A, a copy of the initial quantum state $|x(0)\rangle$, and an evolution time T, the goal is to produce a quantum state that is ε-close to the final state $|x(T)\rangle$. The first algorithms for this problem had complexity polynomial in $1/\varepsilon$ [2, 3, 4, 5]. Subsequent work gave an algorithm with complexity $\mathrm{poly}(\log(1/\varepsilon))$—an exponential improvement—which is optimal in a black-box model [6]. More recent work has streamlined these algorithms and improved their dependence on other parameters [7, 8, 9, 10, 11, 12].

While Hamiltonian simulation has been a focus of quantum algorithms research, the more general problem of simulating linear differential equations of the form $\frac{d\vec{x}}{dt} = A\vec{x} + \vec{b}$ for arbitrary A is less well studied. Reference [13] solves this problem using a quantum linear systems algorithm (QLSA) [14] to implement linear multistep methods, which represent the differential equations with a system of linear equations by discretizing time. The complexity of this approach is $\mathrm{poly}(1/\varepsilon)$. Considering the recent improvements to the complexity of Hamiltonian simulation, it is natural to ask whether linear differential equations can be solved more efficiently as a function of ε.

Hamiltonian simulation is a central component of the QLSA, and the techniques underlying $\mathrm{poly}(\log(1/\varepsilon))$ Hamiltonian simulation have been adapted to give a QLSA with complexity $\mathrm{poly}(\log(1/\varepsilon))$ [15]. However, even if this improved QLSA is used to implement the algorithm of Ref. [13], the overall complexity is still $\mathrm{poly}(1/\varepsilon)$, since the multistep method is a source of error.

In this work, we circumvent these limitations and present a new quantum algorithm for linear differential equations with complexity $\mathrm{poly}(\log(1/\varepsilon))$, an exponential improvement over Ref. [13]. We use the new QLSA in Ref. [15], but encode a truncation of the Taylor series of $\exp(At)$, the propagator for the differential equation, into a linear system, instead of using a linear multistep method.

2 Contribution

2.1 Our Result

More formally, we consider the following problem:

ODE Simulation Problem: The $N \times N$ matrix $A = VDV^{-1}$ is diagonalizable, s-sparse, and has eigenvalues with non-positive real parts. In addition, $\vec{b}$ and $\vec{x}_{\mathrm{in}}$ are N-dimensional vectors with known norms. We have an oracle that computes entries of A, as well as oracles that prepare states proportional to $\vec{b}$ and $\vec{x}_{\mathrm{in}}$. Produce a quantum state ε-close (in ℓ^2 norm) to $\vec{x}(T)/\|\vec{x}(T)\|$ for $T > 0$, where $\vec{x}$ has the initial condition $\vec{x}(0) = \vec{x}_{\mathrm{in}}$ and evolves according to

$$\frac{d\vec{x}}{dt} = A\vec{x} + \vec{b}. \tag{1}$$

Our algorithm for solving this problem achieves the following (see Theorem 9 in [1]):

Main Result: Let $g := \max_{t\in[0,T]} \|\vec{x}(t)\|/\|\vec{x}(T)\|$, $\beta := (\|\vec{x}_{\mathrm{in}}\| + T\||b\rangle\|)/\|\vec{x}(T)\|$ and $\kappa_V = \|V\| \cdot \|V\|^{-1}$. There is a quantum algorithm that solves the ODE Simulation Problem with constant probability which has query and gate complexities that are $\mathrm{poly}(\log(\beta/\varepsilon))$ and linear (up to logarithmic factors) in $\|A\|$, T, s, g, and κ_V.

In addition to scaling well with the simulation error, our algorithm has favourable performance as a function of other parameters. The complexity is nearly linear in the evolution time, which is a quadratic improvement over Ref. [13] and is nearly optimal [5]. The complexity

*`dominic.berry@mq.edu.au`

†`amchilds@umd.edu`

is also nearly linear in the sparsity of A and in g, which characterizes the decay of the solution vector. The latter dependence is necessary since producing a normalized version of a subnormalized solution vector is equivalent to postselection, which is computationally intractable [16]. Along similar lines, we assume that the eigenvalues of A have non-positive real parts, since it is intractable to simulate exponentially growing solutions. This improves upon Ref. [13], where the eigenvalues λ of A must satisfy $|\arg(-\lambda)| \leq \alpha$ for some constant α depending on the stability of the multistep method.

2.2 High-level Overview

Our new method is based on approximating the matrix $\exp(Ah)$ by the sum $T_k(Ah) = \sum_{j=0}^{k}(Ah)^j/j!$ and then using that approximation to evolve the system forward. The exact solution of the equation is $\vec{x}(t) = \exp(At)\vec{x}(0) + [\exp(At) - \mathbb{I}]A^{-1}\vec{b}$ which we can approximate for small times h as $\vec{x}(h) = T_k(Ah)\vec{x}(0) + [T_k(Ah) - \mathbb{I}]A^{-1}\vec{b} = T_k(Ah)\vec{x}(0) + \sum_{j=1}^{k}[(Ah)^{j-1}/j!]h\vec{b}$. To see how we implement this approximation, consider the following linear system for $k = 4$.

$$\begin{pmatrix} \mathbb{I} & 0 & 0 & 0 & 0 & 0 \\ -Ah & \mathbb{I} & 0 & 0 & 0 & 0 \\ 0 & -Ah/2 & \mathbb{I} & 0 & 0 & 0 \\ 0 & 0 & -Ah/3 & \mathbb{I} & 0 & 0 \\ 0 & 0 & 0 & -Ah/4 & \mathbb{I} & 0 \\ -\mathbb{I} & -\mathbb{I} & -\mathbb{I} & -\mathbb{I} & -\mathbb{I} & \mathbb{I} \end{pmatrix} \begin{pmatrix} \vec{z}_0 \\ \vec{z}_1 \\ \vec{z}_2 \\ \vec{z}_3 \\ \vec{z}_4 \\ \vec{z}_5 \end{pmatrix} = \begin{pmatrix} \vec{x}(0) \\ \vec{b}h \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{2}$$

The vector on the right-hand side only contains information about the initial state and the inhomogeneity; however, solving this system gives a vector that includes information about $\vec{x}(h)$.

If we used a QLSA to solve this system, then we would have a state proportional to $\sum_{j=0}^{5} \vec{z}_j \otimes |j\rangle$. Conditioned on measuring 5 in the last register, we would have a state proportional to $\vec{z}_5 = T_4(Ah)\vec{x}(0) + [T_4(Ah) - \mathbf{1}]A^{-1}\vec{b} \approx \vec{x}(h)$ up to some error. For $j < 5$ the $\vec{z}_j$ do not give useful information about the evolution of the state.

Using $T_k(Ah)$ to evolve the system is advantageous because the coefficients for the Taylor series decay factorially, so we obtain a good approximation of $\exp(Ah)$ even with small k (i.e. low order of truncation). This should be contrasted with Berry's algorithm [13] which uses linear multistep methods whose errors do not decay as rapidly when going to higher orders. Berry's algorithm also required additional hypotheses to guarantee the numerical stability of the multistep method, whereas this is not a problem for the Taylor series approach.

The system above only evolves forward for a short time h. To evolve for the total time T, we simply construct a matrix containing T/h copies of the matrix above (we would choose h so T/h is an integer). In addition, at the end we use lines with $\begin{pmatrix} \cdots & 0 & -\mathbb{I} & \mathbb{I} & 0 & \cdots \end{pmatrix}$ in the matrix, to obtain a solution where many of the $\vec{z}_j$ are equal to the final value, which approximates $\vec{x}(T)$. Success is obtained when we obtain one of the $\vec{z}_j$ that approximates $\vec{x}(T)$, so by this process we can boost the success probability to a constant. See Section 2 of Ref. [1] for details of this construction.

3 Conclusion

Compared to the previously proposed quantum algorithm for linear differential equations [13], our algorithm shows an exponential improvement with respect to the error. Our result adds to the growing literature of quantum algorithms for differential equations. Linear systems are ubiquitous in the mathematical and natural sciences, so our algorithm could be utilized in real-world applications of quantum computing.

References

[1] Dominic W Berry, Andrew M Childs, Aaron Ostrander, and Guoming Wang. Quantum algorithm for linear differential equations with exponentially improved dependence on precision. arXiv:1701.03684, 2017.

[2] Seth Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, 1996.

[3] Dorit Aharonov and Amnon Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *Proc. of the 35th ACM STOC*, pages 20–29, 2003.

[4] Andrew M. Childs. *Quantum information processing in continuous time.* PhD thesis, Massachusetts Institute of Technology, 2004.

[5] Dominic W. Berry, Graeme Ahokas, Richard Cleve, and Barry C. Sanders. Efficient quantum algorithms for simulating sparse Hamiltonians. *Comm. Math. Phys.*, 270(2):359–371, 2007.

[6] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. Exponential improvement in precision for simulating sparse Hamiltonians. In *Proc. of the 46th ACM STOC*, pages 283–292, 2014.

[7] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. Simulating Hamiltonian dynamics with a truncated Taylor series. *Phys. Rev. Lett.*, 114(9):090502, 2015.

[8] Dominic W. Berry, Andrew M. Childs, and Robin Kothari. Hamiltonian simulation with nearly optimal dependence on all parameters. In *Proc. of the 56th Annual Symposium on Foundations of Computer Science*, pages 792–809, 2015.

[9] Guang Hao Low and Isaac L. Chuang. Optimal Hamiltonian simulation by quantum signal processing. *Phys. Rev. Lett.*, 118(9):010501, 2017.

[10] Dominic W. Berry and Leonardo Novo. Corrected quantum walk for optimal Hamiltonian simulation. *Quantum Information and Computation*, 16(15-16):1295, 2016.

[11] Guang Hao Low and Isaac L. Chuang. Hamiltonian simulation by qubitization. arXiv:1610.06546, 2016.

[12] Leonardo Novo and Dominic W. Berry. Improved Hamiltonian simulation via a truncated Taylor series and corrections. *Quantum Information and Computation*, 17:0623, 2017.

[13] Dominic W. Berry. High-order quantum algorithm for solving linear differential equations. *J. Phys. A*, 47(10):105301, 2014.

[14] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.*, 103(15):150502, 2009.

[15] Andrew M. Childs, Robin Kothari, and Rolando D. Somma. Quantum linear systems algorithm with exponentially improved dependence on precision. arXiv:1511.02306, 2015.

[16] Scott Aaronson. Quantum computing, postselection, and probabilistic polynomial-time. *P. Roy. Soc. Lond. A*, 461(2063):3473–3482, 2005.

Quantum centrality ranking via quantum walks and its experimental realisation

Josh Izaac[1] * Xiang Zhan[2] Jian Li[2] Peng Xue[2] Paul Abbott[1] Xiaosong Ma[3] Jingbo Wang[1] †

[1] *School of Physics, The University of Western Australia, Crawley WA 6009*
[2] *Department of Physics, Southeast University, Nanjing 211189, China*
[3] *School of Physics, Nanjing University, Nanjing 210093, China*

Abstract. Network centrality has important implications well beyond its role in physical and information transport analysis; as such, various quantum-walk-based algorithms have been proposed for measuring network vertex centrality. In this work, we propose a continuous-time quantum walk algorithm for determining vertex centrality, and show that it generalizes to arbitrary graphs via a statistical analysis of randomly generated scale-free and Erds-Rnyi networks. As a proof of concept, the algorithm is detailed on a four-vertex star graph and physically implemented via linear optics, using spatial and polarization degrees of freedoms of single photons; the first successful physical demonstration of a quantum centrality algorithm. Finally, we extend our quantum centrality to directed non-Hermitian graph structures using PT-symmetric quantum walks; a statistical analysis shows strong agreement with the classical PageRank measure on directed acyclic graphs.

Keywords: network centrality, quantum information, quantum algorithms, non-Hermitian dynamics

Quantum walks are an important tool in the field of quantum information theory. Indeed as a method of universal quantum computation [4], they have motivated the creation of quantum algorithms that are faster and more efficient than their classical analogues [3], and provided a vital link between quantum computation and modeling complex quantum dynamical systems (for example, photosynthesis [7]). Moreover, in providing a method of modelling network structures that doubles as a universal system of quantum computation, the quantum walk is uniquely placed in the quest to find quantum analogues of classical network algorithms. As such, one potential application of the quantum walk is in providing an efficient quantum algorithm for vertex centrality ranking in network analysis. Network centrality has important implications well beyond its role in physical and information transport analysis; as such, various quantum-walk-based algorithms have been proposed for measuring network vertex centrality. These include algorithms built on the standard discrete-time quantum walk [3], the Szegedy discrete-time quantum walk [9, 10, 6], or the continuous-time quantum stochastic walk (QSW) [6, 11, 5]. However, whilst comparing well to classical centrality measures, these have the distinct disadvantage of requiring expanded Hilbert spaces (up to N^2 dimensions for a graph of N vertices), or in the case of the QSW, muting the quantum behaviour due to decoherence. As such, these quantum algorithms require increased quantum resources, and as a result, for large networks are difficult to implement on current experimental setups.

We instead propose a continuous-time quantum walk algorithm for determining vertex centrality, allowing us to preserve the full quantum behaviour of the walker, whilst limiting the dimension of the Hilbert space to N. Further, we show that this scheme generalizes to arbitrary graphs via a statistical analysis of randomly generated scale-free and Erdős-Rényi networks; the analysis indicates that the proposed measure is highly correlated with the classical eigenvector centrality, and we suggest that it provides an extension of the eigenvector centrality to the quantum realm. As a proof of concept, the algorithm is detailed on a four-vertex star graph and physically implemented via linear optics, using spatial and polarization degrees of freedoms of single photons. This talk reports the first successful physical demonstration of a quantum centrality algorithm.

Unfortunately, one disadvantage of the quantum walk as utilised above is the imposition of unitarity, due to the quantum nature of the walkers. As such, the conventional quantum walk is unable to model or analyze directed network structures, without either a) resulting in non-unitary dynamics, or b) modifying the framework. This serves as a particular hindrance in extending established quantum algorithms (e.g. quantum search, centrality measures, graph isomorphism) and quantum dynamical models to systems with direction/biased potentials (such as transport of electrons or excitons).

One such solution to this problem lies in the field of PT-symmetry, which offers the capability to perform quantum walks on directed graphs with non-Hermitian Hamiltonians whilst preserving the norm [1, 2, 8]. Thus, in the second part of this presentation, we formalize a rigorous framework for continuous-time quantum walkers on pseudo-Hermitian directed graph structures. This is then extended to the cases of multi-particle quantum walks and interdependent networks, before being applied to measure vertex centrality in ensembles of randomly generated PT-symmetric scale-free and Erdős-Rényi networks – resulting in strong agreement with the classical PageRank algorithm, and enabling the proposed quantum 'eigenvector-like' centrality algorithm to accurately

*josh.izaac@uwa.edu.au

†jingbo.wang@uwa.edu.au

extract the most central vertex in directed acyclic networks which fail under the classical eigenvector centrality. Furthermore, we show that this formalism is equivalent to considering an undirected, yet weighted, complete graph with self-loops, providing a structural interpretation that may lead to simple experimental implementation.

Whilst prior work on quantum centrality measures has generally focused on discrete-time quantum walk-based algorithms, our work shows that those based on continuous-time quantum walks should not be so readily discounted. With a reduced Hilbert space, they have the ability to rank vertices in a manner suggestive of the classical eigenvector centrality, providing an extension of eigenvector centrality to the quantum realm. Moreover, through the PT-symmetric framework, this proposed centrality scheme can be extended to directed graph structures, and provide centrality ranking on graphs which the classical eigenvector centrality ranking is inconclusive. Finally, we have demonstrated a successful experimental implementation of a quantum centrality scheme. By exploring the capability of continuous-time quantum walks in network centrality analysis, the work presented here may, in future, lead to easily-implementable and efficient quantum centrality algorithms that take advantage of the potential speedup provided by quantum computation.

References

[1] Carl M. Bender and Stefan Boettcher. Real spectra in Non-Hermitian hamiltonians having PT symmetry. *Physical Review Letters*, 80(24):5243–5246, June 1998.

[2] Carl M. Bender and Philip D. Mannheim. symmetry and necessary and sufficient conditions for the reality of energy eigenvalues. *Physics Letters A*, 374(15–16):1616–1620, April 2010.

[3] Scott Berry and Jingbo Wang. Quantum-walk-based search and centrality. *Physical Review A*, 82(4):042333, October 2010.

[4] Andrew M. Childs. Universal computation by quantum walk. *Physical Review Letters*, 102(18):180501, May 2009.

[5] Peter E. Falloon, Jeremy Rodriguez, and Jingbo B. Wang. QSWalk: a mathematica package for quantum stochastic walks on arbitrary graphs. *arXiv:1606.04974 [quant-ph]*, May 2016.

[6] T. Loke, J. W. Tang, J. Rodriguez, M. Small, and J. B. Wang. Comparing classical and quantum PageRanks. *Quantum Information Processing*, 16(1):25, January 2017.

[7] Masoud Mohseni, Patrick Rebentrost, Seth Lloyd, and Alán Aspuru-Guzik. Environment-assisted quantum walks in photosynthetic energy transfer. *The Journal of Chemical Physics*, 129(17):174106–174106–9, November 2008.

[8] Ali Mostafazadeh. Pseudo-Hermiticity versus PT symmetry: The necessary condition for the reality of the spectrum of a non-Hermitian hamiltonian. *Journal of Mathematical Physics*, 43(1):205–214, January 2002.

[9] G. D. Paparo and M. A. Martin-Delgado. Google in a quantum network. *Scientific Reports*, 2:00444, June 2012.

[10] Giuseppe Davide Paparo, Markus Müller, Francesc Comellas, and Miguel Angel Martin-Delgado. Quantum google in a complex network. *Scientific Reports*, 3:02773, October 2013.

[11] Ilya Sinayskiy and Francesco Petruccione. Open quantum walks: a short introduction. *Journal of Physics: Conference Series*, 442(1):012003, 2013.

Based on the following manuscripts

[1] J. A. Izaac, J. B. Wang, P. C. Abbott, and X. S. Ma. Quantum centrality testing on directed graphs via PT-symmetric quantum walks. *arXiv:1607.02673 [quant-ph]*, July 2017.

[2] Josh A. Izaac, Xiang Zhan, Zhihao Bian, Kunkun Wang, Jian Li, Jingbo B. Wang, and Peng Xue. Centrality measure based on continuous-time quantum walks and experimental realization. *Physical Review A*, 95(3):032318, March 2017.

Resource destroying maps, with new applications

[Phys. Rev. Lett. **118**, 060502 (2017) and an upcoming work]

Zi-Wen Liu,[1, *] Xueyuan Hu,[2] Ryuji Takagi,[1] and Seth Lloyd[3]

[1] *Center for Theoretical Physics and Department of Physics,*
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[2] *School of Information Science and Engineering, and Shandong Provincial Key Laboratory*
of Laser Technology and Application, Shandong University, Jinan 250100, China
[3] *Department of Mechanical Engineering, Massachusetts*
Institute of Technology, Cambridge, Massachusetts 02139, USA

Resource theory is a widely applicable framework for analyzing the physical resources required for given tasks, such as computation, communication, and energy extraction. We propose a general scheme for analyzing resource theories based on resource destroying maps, which leave resource-free states unchanged but erase the resource stored in all other states. We introduce a group of general conditions that determine whether a quantum operation exhibits typical resource-free properties in relation to a given resource destroying map. Our theory reveals fundamental connections among basic elements of resource theories, in particular, free states, free operations, and resource measures. In particular, we define a class of simple resource measures that can be calculated without optimization, and that are monotone nonincreasing under operations that commute with the resource destroying map. We apply our theory to the resources of coherence and quantum correlations (e.g., discord), two prominent features of nonclassicality.

In an upcoming work, we use the theory of resource destroying map to show a rather surprising result that a simple discord measure called diagonal discord is generically monotone under local discord non-generating operations.

Resource theory originates from the observation that certain properties of physical systems become valuable resources, when the operations that can be performed are restricted so that such properties are hard to create. The framework of resource theory has been applied to various other concepts in quantum information, such as purity [1], magic states [2] and coherence [3, 4], and to broader areas, such as asymmetry [5] and thermodynamics [6]. In recent years, considerable effort has been devoted to developing a unified framework of resource theories [7–9]. In particular, Ref. [7] studies the general case where the set of free operations is maximal, i.e., all (asymptotically) resource non-generating operations are allowed, when the resource satisfies several postulates (e.g., the set of free states is convex). Some key aspects of resource theories are not addressed by existing frameworks, however. For example, characterizing a proper set of free operations is frequently a major difficulty in establishing a resource theory, and we do not yet have general principles and understandings for nonmaximal theories. Indeed, a successful resource theory is usually specified by physical restrictions on the set of allowed operations: LOCC and thermal operations [6, 10, 11] are prominent examples. But such restrictions are often stronger than merely nongenerating, and may lead to mathematical difficulties in characterizing and calculating monotones. Moreover, existing results do not apply to some resources, such as discord, where the set of free states is nonconvex.

In this work, we propose a mathematically friendly framework of resource-free operations based on the notion of resource destroying maps. Our framework applies to all resources including those with nonconvex sets of free states.

Resource destroying maps. This is the key concept of our theory. Let F be the set of free states for a certain theory. For all input states ρ, a resource destroying map λ satisfies the following requirements: (i) resource destroying: if $\rho \notin F$, $\lambda(\rho) \in F$; (ii) nonresource fixing: if $\rho \in F$, $\lambda(\rho) = \rho$. In other words, a resource destroying map outputs a free state if the input is not free, and leaves the input unchanged otherwise. The resource destroying map characterizes the resource-free space: F consists precisely of the fixed points of λ. It is easy to see that resource destroying maps are idempotent and surjective. It is helpful to draw an analogy with fiber bundles: λ defines a bundle projection onto F. Call a non-free state a *parent state* of its image free state. Then each free state defines a *family* consisting of corresponding parent states (the fiber) and the free state itself. Note that many important properties of our framework sharply contrast linear resource destroying maps with nonlinear ones (which we discuss in the long version).

Note that a resource destroying map does not have to be completely positive or linear, and can be highly nonuniform. However, we are mostly interested in the physically motivated ones, usually with simple descriptions that work universally for all inputs. For example, the simplest case is when the resource destroying map can be represented by a quantum channel. It can be shown that λ cannot be a linear map (thus not a channel) when F is nonconvex.

Resource-free conditions. The definition of resource destroying maps allows us to write down a group of general conditions that determine whether an operation exhibits various key resource-free properties. Consider a theory with resource destroying map λ. The following duo of conditions

$$\mathcal{E} \circ \lambda = \lambda \circ \mathcal{E} \circ \lambda \tag{1}$$

$$\lambda \circ \mathcal{E} = \lambda \circ \mathcal{E} \circ \lambda \tag{2}$$

respectively determines if a quantum operation $\mathcal{E}$ is resource non-generating/non-activating. The free set F is closed under *resource non-generating operations* [Eq. (1)]: they never map a free state to a resourceful one. This is a necessary constraint on free operations, since any other operation can create resource, thus making the theory trivial. As studied in e.g. [7, 8], such maximal theories (under some assumptions e.g. convexity) possess a common structure: they are reversible, and the regularized relative entropy is the unique monotone (asymptotically). On the other hand, the *non-activating condition* [Eq. (2)] has not been well studied to our knowledge. Think of the output of λ as the free part of an input state. This condition means that $\mathcal{E}$ cannot make use of the resource stored in any input to affect the free part. An alternative interpretation is that resource non-activating operations do not mix up fibers in the fiber bundle. More colloquially, such operations never break apart a family: members of the same family cannot be mapped to different families. The combination of these two constraints gives us the *commuting condition*:

$$\lambda \circ \mathcal{E} = \mathcal{E} \circ \lambda. \tag{3}$$

It is also meaningful to consider the "selective" versions of each condition: in practice, one may want to require that the above conditions be satisfied even when considering selective measurements, i.e., the outcome of the measurement is accessible. This leads to the following variation of each condition: there is a Kraus decomposition $\mathcal{E}(\cdot) = \sum_\mu K_\mu \cdot K_\mu^\dagger$ such that all Kraus arms $\mathcal{E}_\mu(\cdot) \equiv K_\mu \cdot K_\mu^\dagger$ satisfies the condition.

The sets of free operations derived from the above resource destruction framework exhibit several general features. We show that certain free operations can be constructed by composing arbitrary operations and the resource destroying map in certain sequences. Moreover, all conditions hold for convex combinations when λ is linear, and for compositions. In particular, we demonstrate that the commuting condition plays a key role in the quantification of resources. In particular, the distance (as measured by any contractive distance measure) between a state and its resource-free version is monotone nonincreasing under commuting operations. More formally, for any state ρ and contractive distance D, define $\tilde{\mathfrak{D}}(\rho) := D(\rho, \lambda(\rho))$, then, for any Γ that commutes with λ,

$$\tilde{\mathfrak{D}}(\rho) \geq D(\Gamma(\rho), \Gamma(\lambda(\rho))) = D(\Gamma(\rho), \lambda(\Gamma(\rho))) \equiv \tilde{\mathfrak{D}}(\Gamma(\rho)), \tag{4}$$

where the inequality follows from the contractivity of D. Therefore, for any resource theory with free operations satisfying the commuting condition, we have a class of computationally easy monotones which avoid optimizations (given that λ is suitably defined). We should note that $\tilde{\mathfrak{D}}$ is not necessarily continuous everywhere when λ is nonlinear, which requires more careful analysis in application (as will be demonstrated for discord).

Applications. Coherence and quantum correlations are prominent features of nonclassical systems, which are under active study in recent years. We explicitly apply our framework to these two theories in this work.

The resource theory treatment of coherence has drawn a lot of attention in recent years. See [12] for a recent review. For coherence, the natural resource destroying map is simply a measurement in the preferred basis. We show that a class of coherence-free operations with different physical correspondences can be derived from our theory. Notably, several recent proposals, namely Incoherent Operations [3, 4], Strictly Incoherenct Operations [13], Dephasing-covariant Operations (appeared during the preparation of this work) [14, 15], naturally emerge from the theory of coherence destroying map. The results of coherence can therefore be extended to other theories through our framework. See the long version for more details.

On the other hand, discord-type quantum correlation has been refusing satisfactory treatments within the framework of resource theories. Unlike separable states for entanglement, the sets of classically correlated states (discord-free states) are nonconvex, and various key aspects of discord such as free operations are poorly understood. Nevertheless, discord fits well into our theory, which provides new insights into discord-free operations and discord monotones. The canonical discord destroying map is simply a local measurement in the eigenbasis, which we call π: classically correlated states are fixed points of π (on the classical side); and a local projective measurement always erases discord. Local operations that do not create discord have been studied in Refs. [16, 17], but other classes have not been considered before to our knowledge. We find that some of the simplest quantum operations exhibit typical but different behaviors in this theory. In this paper and an upcoming work [18], we show that all local isotropic channels (including unitary, antiunitary, depolarizing channels and their combinations) are π-commuting. Local rank-one projective measurements, however, are discord non-generating but activating. We also define a measure-and-prepare protocol depending on the input that can generate but cannot activate discord, but it is unclear at the moment whether there are channels in this class. Contractive distances between any ρ_{AB} and $\pi_A(\rho_{AB})$, e.g., a physically motivated simple measure of discord called diagonal discord [19], are monotone under $X(\pi)$ (so all isotropic channels). Together with some numerical evidence [18], we find that, rather surprisingly, diagonal discord is likely monotone under all local discord non-generating channels. We should point out that diagonal discord may suffer from discontinuities [20, 21], however it can be shown that they do not occur at full rank non-degenerate states [18].

Concluding remarks. In this work, we propose a simple and widely applicable theory of resource theories based on the notion of resource destroying maps. Our theory provides a general scheme for understanding the power of quantum operations in relation to certain quantum resources. The theory shows how to extend results that have been previously derived for specific resources to a more general class of resource theories. In particular, our framework may lead to conceptual advances in understanding nonconvex theories such as discord. It would also be interesting to apply the framework of resource destroying maps to other important resource theories, such as those of entanglement, magic states, asymmetry and thermodynamics.

* zwliu@mit.edu

[1] M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. A **67**, 062104 (2003).

[2] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, New J. Phys. **16**, 013009 (2014).

[3] T. Baumgratz, M. Cramer, and M. B. Plenio, Phys. Rev. Lett. **113**, 140401 (2014).

[4] A. Winter and D. Yang, Phys. Rev. Lett. **116**, 120404 (2016).

[5] I. Marvian and R. W. Spekkens, Nat. Commun. **5** (2014).

[6] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Phys. Rev. Lett. **111**, 250404 (2013).

[7] F. G. S. L. Brandão and G. Gour, Phys. Rev. Lett. **115**, 070503 (2015).

[8] M. Horodecki and J. Oppenheim, Int. J. Mod. Phys. B **27**, 1345019 (2013).

[9] B. Coecke, T. Fritz, and R. W. Spekkens, arXiv:1409.5531.

[10] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and T. Beth, Int. J. Theor. Phys. **39**, 2717 (2000).

[11] M. Horodecki and J. Oppenheim, Nat. Comm. **4** (2013).

[12] A. Streltsov, G. Adesso, and M. B. Plenio, ArXiv e-prints (2016), arXiv:1609.02439 [quant-ph].

[13] B. Yadin, J. Ma, D. Girolami, M. Gu, and V. Vedral, arXiv:1512.02085.

[14] I. Marvian and R. W. Spekkens, arXiv:1602.08049.

[15] E. Chitambar and G. Gour, Phys. Rev. Lett. **117**, 030401 (2016).

[16] X. Hu, H. Fan, D. L. Zhou, and W.-M. Liu, Phys. Rev. A **85**, 032102 (2012).

[17] A. Streltsov, H. Kampermann, and D. Bruß, Phys. Rev. Lett. **107**, 170502 (2011).

[18] Z.-W. Liu, unpublished.

[19] S. Lloyd, V. Chiloyan, Y. Hu, S. Huberman, Z.-W. Liu, and G. Chen, arXiv:1510.05035.

[20] S. Wu, U. V. Poulsen, and K. Mølmer, Phys. Rev. A **80**, 032319 (2009).

[21] A. Brodutch and K. Modi, Quantum Info. Comput. **12**, 721 (2012).

Logical paradoxes in deterministic quantum state injection

Nadish de Silva

Department of Computer Science,
University College London,
WC1E 6BT London, United Kingdom

While quantum computers are expected to yield considerable advantages over classical devices, the precise features of quantum theory accounting for these advantages remain unclear. Contextuality—the denial of a notion of classical *physical reality*—has emerged as a promising hypothesis.

Howard *et al.* showed that single-qudit magic states, resources critical to achieving quantum universality, exhibit a standard form of contextuality known to facilitate *probabilistic* advantages in computational and communicational tasks. Strong contextuality is a logical form of contextuality describing systems, e.g. the GHZ state, that exhibit paradoxical behaviour: the true statements describing their response to measurement are inconsistent.

We consider the role of paradoxes in *deterministically* achieving quantum universality. We present large families of strongly contextual multiqudit magic states admitting deterministic implementation of gates from the third level of the Clifford hierarchy. Our results contribute to the computational resource theory of contextuality by applying logical tools towards a structural understanding of quantum information theory.

Despite decades of research, identification of the precise physical and logical features accounting for quantum advantages over classical devices, and the mechanisms by which they do so, remains a pressing open problem in quantum computation. Further refinement of quantum computational resource theories will clarify for which computational problems quantum computers offer advantages and facilitate concrete design improvements in the architecture of quantum devices.

Contextuality is a concept from the foundations of quantum mechanics first articulated by Bell-Kochen-Specker [1, 2]. Their theorem denies the possibility of a classical explanation for the statistical predictions of quantum theory in terms of *hidden variables*. Quantum measurements cannot be straightfowardly modelled as revealing properties of a pre-existing classical reality. Contextuality subsumes nonlocality as a special case. Contextuality and nonlocality have emerged as promising hypotheses as essential quantum resources needed for achieving advantages in computation and communication. In particular, a recent, seminal result of Howard *et al.* [3] demonstrates the necessity of contextuality in fault-tolerantly achieving quantum universal computation in the experimentally tractable setting of magic state distillation.

Our motivation is to refine the computational resource theory of contextuality by considering the role of logical paradoxes realized by quantum resources. These paradoxes are relevant in a diverse variety of settings: e.g. nonlocal games [4, 5], measurement-based quantum computation [6], zero-error information theory [7], etc. We consider here, for the first time, their role in achieving (fault-tolerant) universal quantum computation [8].

I. STRONG CONTEXTUALITY AND GHZ-TYPE PARADOXES

We make essential use of Abramsky-Brandenburger's notion of *strong contextuality* [9], which generalizes the notion of *maximal nonlocality* introduced by Elitzur-Popescu-Rohrlich [10] and Barrett *et al.* [11]; it captures the paraxodical nature of systems that is key to achieving deterministic advantages. Well known examples of strong contextuality include the GHZ state [12] and the PR box [13]. Whereas standard contextuality is witnessed by the violation of a probabilistic inequality

[14–16], strong contextuality is witnessed by a logical paradox [17]. Correlations exhibiting standard probabilistic contextuality can be seen as stochastic mixtures of classical and paradoxical systems [18]. Thus, the probabilistic advantage conferred by contextual resources can be understood as a consequence of probabilistic access to paradoxical resources (and their attendant conferment of deterministic advantages).

II. CONTEXTUALITY AS A RESOURCE

Contextuality (and the special case of nonlocality) has been shown to be critical to achieving quantum advantage in a variety of information theoretic tasks. The pattern of strongly contextual resources conferring deterministic advantages is seen in a diverse variety of computational and communicational tasks: e.g. communication complexity [19], measurement-based quantum computation [20], nonlocal games [4], zero-error classical channel capacity [7], etc.

Sharing an unlimited number of PR boxes between two parties renders all communication complexity problems trivial [19]. Kochen-Specker configurations play an essential role in boosting the zero-error (i.e. deterministic) capacity of certain classical channels [7]. Perfect strategies for nonlocal games require strongly contextual resources [18].

Anders and Browne [21] proved that a linear computer with access to a GHZ state can deterministically compute the OR of two bits (a non-linear function). The input bits determine which measurements performed on the three qubits; linear post-processing of the measurement outcomes yields the desired output bit. Raussendorf [20] generalised this to linear measurement-based quantum computers. A computer restricted to performing linear arithmetic with access to measurements on an empirical model can, with non-zero probability, compute a nonlinear function only if the empirical model is contextual. It can *deterministically* compute a nonlinear function only if the resource is strongly contextual. The success probability of the computation is bounded by the product of a measure of the function's nonlinearity and the contextual fraction of the resource state [18].

III. RESULTS

As in Howard *et al.*, we consider stabilizer quantum mechanics [22]. Circuits built within this scheme, i.e. from Clifford gates, are relatively easy to implement in a fault-tolerant way. By the Gottesman-Knill theorem [23], however, these circuits are efficiently classically simulable and thus offer none of the power of quantum computation. The ability to perform a non-Clifford gate promotes this scheme to universal quantum computation. This can be achieved by state injection [24]: a variation on the quantum teleportation protocol that consumes a resource magic state and implements a non-Clifford gate. Howard *et al.* showed that single-qudit magic states, resources critical to achieving quantum universality, exhibit standard contextuality with respect to two-qudit stabilizer operations once paired with an ancilla qudit. In this work, we consider the contextuality properties of two-qudit resources needed for deterministic state injection.

Here, we establish a link between strong contextuality and the Clifford hierarchy [22], a distinguished class of gates admitting deterministic state injection. In contrast, a general magic state must undergo a stochastic distillation and injection process (due to Bravyi and Kitaev [8]) before being converted into a non-Clifford gate. We present large families of strongly contextual multiqudit magic states admitting deterministic implementation of gates from the third level of the Clifford hierarchy.

Unlike Howard *et al.*, we consider genuinely multiqudit states without the need for ancilla qudits. We also make no use of negativity or Wigner functions [25]; instead, we exhibit GHZ-type paradoxical behaviour via elementary number theory.

The structure of our results is as follows:

- We use Cui-Gottesman-Krishna's recent classification of the diagonal gates of the Clifford hierarchy [26] to define *strong* magic states (Definition 1) as those capable of deterministically injecting gates from the third level with the reasonable restriction of having vanishing *local terms*.

- We prove a surprising lemma about the class of possible hidden variable models for multiqudit stabilizer quantum mechanics. Lemma 1 indicates that such hidden variable models must preserve the vector space structure of the Gross phase space [30] for stabilizer theory.

- We provide a number-theoretic criterion (Lemma 2) for the possibility vs. impossibility of witnessing a given joint outcome upon stabilizer measurements on magic states.

- By choosing appropriate measurements to witness our paradoxes (Table 1), and exploiting Dickson's 1896 classification of *permutation polynomials* of low degree [31] (Theorem 2), we are able to prove that all strong magic states exhibit generalised GHZ-type paradoxes. These paradoxes exist in infinitely many dimensions.

IV. SIGNIFICANCE AND OUTLOOK

We have presented large families of genuinely multiqudit magic states exhibiting GHZ-type paradoxes that are intimately connected with the Clifford hierarchy. The families exist in arbitrarily large dimensions. Our results solidify the intuition that the logical paradoxes described by strong contextuality (and nonlocality) play an important role in facilitating deterministic advantages over classical devices. Most analyses of contextuality as a resource have considered standard contextuality and probabilistic advantage; probabilistically contextual resources can be seen as mixtures of strongly contextual and classical resources.

A closer analysis of how these states facilitate the deterministic conversion of magic states into third level gates is warranted. A natural further question is to classify strong contextuality in stabilizer mechanics and to fully delineate its relationship with the Clifford hierarchy. (It is noteworthy that Lemma 2 holds even when Φ is not a polynomial. That the diagonal subset of the Clifford hierarchy is defined using polynomials provides a strong hint that of their underlying logical structure. Cubics are particularly suited for been proven to be permutation polynomials.) This may require extension the elegant characterisation of Cui, Gottesman, and Krishna, from the diagonal to the general case.

V. ACKNOWLEDGEMENTS

We thank Samson Abramsky, Rui Soares Barbosa, Eric Cavalcanti and Mark Howard for helpful discussions. Financial support from the EPSRC via grant EP/N017935/1 (Contextuality as a resource in quantum computation) is gratefully acknowledged. This work was partially conducted during the Logical Structures in Computation programme at the Simons Institute for the Theory of Computing at the University of California, Berkeley.

In these appendices we:

- Prove background material on contextuality, with formal mathematical definitions.
- Prove the background material on multiqudit stabilizer mechanics as presented by Gross [30] as well as the Cui-Gottesman-Krisha [26] classification of the diagonal Clifford hierarchy.
- Give our definitions and results and outline the structure of the proof of our main theorem.
- Give the details of the proofs of our lemmas and main theorem.

VI. APPENDIX I: CONTEXTUALITY, PROBABILISTIC AND PARADOXICAL

Here, we follow the Abramsky-Brandenburger framework for contextuality and nonlocality [9]. An experiment wherein it is not assumed that all measurements can be performed simultaneously is formally described by a *measurement scenario*: a triple $(\mathcal{M}, \mathcal{C}, \mathcal{O})$ where $\mathcal{M}$ is a set of measurement labels and $\mathcal{C}$ is the set of *contexts*. A context is a maximal set $C \subset \mathcal{M}$ of compatible measurements. Measurement of each individual $m \in \mathcal{M}$ yields a value from the outcome set $\mathcal{O}$. For example, the standard Bell scenario is captured by $\mathcal{M} = \{A_0, A_1, B_0, B_1\}$, contexts $\mathcal{C} = \{\{A_x, B_y\} : x, y \in \{0,1\}\}$, and outcome set $\mathcal{O} = \{0,1\}$. Empirical data from performing an experiment on a system in a fixed state is given by a context-indexed family of conditional distributions $\mathcal{E}_C : \mathcal{O}^C \to [0,1]$ on joint outcomes. Data satisfying physically reasonable conditions of generalised nonsignalling—the marginal distributions $\mathcal{E}_C|_{C \cap C'}$ and $\mathcal{E}_{C'}|_{C \cap C'}$ agree for all C and C'—constitute an *empirical model* $\mathcal{E}$.

An empirical model $\mathcal{E}$ is noncontextual when its predictions can be accounted for by a *noncontextual hidden variable model.* Such a hidden variable model can, without loss of generality [9, 27], be assumed to have a canonical form: the hidden variable space is $\Lambda = \mathcal{O}^{\mathcal{M}}$ with hidden variables being functions $\lambda : \mathcal{M} \to \mathcal{O}$ together with a distribution $\mu : \Lambda \to [0,1]$ yielding the $\mathcal{E}_C$ as marginal distributions. Data arising from a hidden variable model will satisfy all Bell-type inequalities; thus, contextuality is witnessed by violation of an inequality.

A hidden variable $\lambda : \mathcal{M} \to \mathcal{O}$ and an empirical model $\mathcal{E}$ are *consistent* when, for each context, the joint outcome λ prescribes to the measurements in M has nonzero probability: $\mathcal{E}_C(\lambda|_C) > 0$. An empirical model is *strongly contextual* when it is inconsistent with all hidden variables. Well-known examples of strongly contextual empirical models include GHZ states and PR boxes. Whereas the standard probabilistic form of contextuality is witnessed by a violation of an equality, strong contextuality is witnessed by a logical paradox.

To each measurement $M \in \mathcal{M}$ and outcome $o \in \mathcal{O}$, the symbol $M_i \to o$, is interpreted as "m is measured resulting in the outcome o". For measurements $M_1, ..., M_n \in C$ in a common context C, one can construct sentences with symbols $M_i \to o$ and the connectives AND, OR, and NOT. The *theory* of an empirical model is the subset of sentences that are true with certainty upon measurement, no matter the outcome. A state is strongly contextual if and only if its theory is logically inconsistent.

An empirical model $\mathcal{E}$ can be expressed as a convex mixture of a noncontextual part $\mathcal{E}^{NC}$ and a strongly contextual part $\mathcal{E}^{SC}$ as $\mathcal{E} = \mathrm{CF}(\mathcal{E})\mathcal{E}^{SC} + [1 - \mathrm{CF}(\mathcal{E})]\mathcal{E}^{NC}$. Here, $\mathrm{CF}(\mathcal{E})$ is a measure of contextuality known as the *contextual fraction* [9]; it generalises the nonlocal fraction [11]. An empirical model is strongly contextual if and only if $\mathrm{CF}(\mathcal{E})$ is 1. Thus, the contextual fraction measures the degree to which a model provides probabilistic access to paradoxical data. Strongly contextual models have a geometric interpretation: they are precisely the convex sums of the nonsignalling polytope's contextual vertices [28, 29].

VII. APPENDIX III: STABILIZER QUANTUM MECHANICS

Here, we follow the presentation of Gross [30]. In what follows, d is an odd prime number. The d-dimensional single-qudit Pauli spin matrices are defined by $X(q)\,|j\rangle = |j+q\rangle$ and $Z(p)\,|j\rangle = \omega^{pj}\,|j\rangle$ where ω is the phase factor $e^{2\pi i/d}$ and addition is modulo d. A Weyl operator, represented by *phase point* coordinates in $\mathbb{Z}_d^2$, is defined by $W(p,q) = \omega^{-2^{-1}pq}Z(p)X(q)$ where 2^{-1} denotes the multiplicative inverse of 2 in $\mathbb{Z}_d$. An n-particle Weyl operator, represented by coordinates in $\mathbb{Z}_d^{2n}$, is defined as $W(p_1,q_1,...,p_n,q_n) = W(p_1,q_1)\otimes ... \otimes W(p_n,q_n)$. For notational convenience, we denote a point $(p_1,q_1,...,p_n,q_n)$ in an n-particle phase space as $(\mathbf{p},\mathbf{q})$. The *symplectic inner product* of two phase space points is defined by: $[(\mathbf{p},\mathbf{q})] = \sum_{i=1}^n p_iq_i' - p_i'q_i$. Weyl operators obey a composition law: $W(\mathbf{p},\mathbf{q})W(\mathbf{p'},\mathbf{q'}) = \omega^{2^{-1}[(\mathbf{p},\mathbf{q}),(\mathbf{p'},\mathbf{q'})]}W(\mathbf{p}+\mathbf{p'},\mathbf{q}+\mathbf{q'})$. Therefore, $W(\mathbf{p},\mathbf{q})W(\mathbf{p'},\mathbf{q'}) = W(\mathbf{p'},\mathbf{q'})W(\mathbf{p},\mathbf{q}) = W(\mathbf{p}+\mathbf{p'},\mathbf{q}+\mathbf{q'})$ if and only if $[(\mathbf{p},\mathbf{q}),(\mathbf{p'},\mathbf{q'})] = 0$.

The n-Weyl operators with arbitrary phase form the n-Pauli group $\mathcal{C}_1^n$. The Clifford gates are those unitaries preserving the Pauli group: $\mathcal{C}_2^n = \{U : UPU^* \in \mathcal{C}_1^n \text{ for all } P \in \mathcal{C}_1^n\}$. The Clifford hierarchy [22] is defined similarly: $\mathcal{C}_k^n = \{U : UPU^* \in \mathcal{C}_{k-1}^n \text{ for all } P \in \mathcal{C}_1^n\}$.

The magic states we will consider arise from the third level of the Clifford hierarchy. Cui, Gottesman, and Krishna [26] give an explicit description of all diagonal gates in the k^{th} level. For $d > 3$, diagonal gates of the third level have the form:

$$U_\Phi = \sum_{\mathbf{j}\in\mathbb{Z}_d^n} \omega^{\Phi(\mathbf{j})}\,|\mathbf{j}\rangle\,\langle\mathbf{j}|$$

where Φ is a multivariable polynomial of degree 3. Every such gate yields a magic state $|\Phi\rangle = U_\Phi\,|+\rangle^{\otimes n} = \sum_{\mathbf{j}\in\mathbb{Z}_d^n} \omega^{\Phi(\mathbf{j})}\,|\mathbf{j}\rangle$. The magic states arising in this way from gates $U_\Phi \in \mathcal{C}_3^n \setminus \mathcal{C}_2^n$ in the third level of the Clifford hierarchy are especially useful achieving quantum universality in that they admit deterministic protocols for injecting the gate U_Φ. Gates from outside the Clifford hierarchy may be implemented via state injection; however, this is a stochastic process that requires randomly many attempts.

VIII. APPENDIX VI: PARADOXES OF STRONG MAGIC STATES

Stabilizer quantum mechanics contains, as available measurements, the set of Weyl operators as indexed by phase points: $\mathcal{M} = \mathbb{Z}_d^{2n}$. The contexts are given by maximal commuting subsets of $\mathcal{M}$. We now present large families of GHZ-type paradoxes arising from magic states $|\Phi\rangle$ with *no local terms*, i.e. the coefficients for the j^3, k^3 terms vanish.

Definition 1. *A two-qudit magic state $|\Phi\rangle = d^{-1}\sum_{j,k\in\mathbb{Z}_d^2} \omega^{\Phi(j,k)}\,|j\rangle\,|k\rangle$ is* strong *if, with either ϕ_1 or $\phi_2 \not\equiv 0$,*

$$\Phi(j,k) = \phi_1 j^2k + \phi_2 jk^2 + \phi_3 j^2 + \phi_4 k^2 + \phi_5 jk + \phi_6 j + \phi_7 k + \phi_8$$

Theorem 1. *Suppose that the dimension $d \not\equiv 1$ (mod 3). All strong magic states $|\Phi\rangle$ are strongly contextual with respect to stabilizer measurements. The states $C\,|\Phi\rangle$, where C is any Clifford gate, are also strongly contextual.*

We prove that for any strong magic state $|\Phi\rangle$ and any hidden variable $\lambda : \mathbb{Z}_d^4 \to \mathbb{Z}_d$ assumed, for contradiction, to be consistent with $|\Phi\rangle$, λ must predict the occurrence of an impossible event for one of the following measurements.

Type	**Operators, up to phase**	**Phase points**
I_α	$Z \otimes \mathbb{I}$ and $\mathbb{I} \otimes Z^\alpha X$	$(1,0,0,0)$ and $(0,0,\alpha,1)$
II_α	$\mathbb{I} \otimes Z$ and $Z^\alpha X \otimes \mathbb{I}$	$(0,0,1,0)$ and $(\alpha,1,0,0)$
$\mathrm{III}_{\alpha,\beta}$	$Z \otimes Z^\beta$ and $X \otimes Z^\alpha X^{-\beta^{-1}}$	$(1,0,\beta,0)$ and $(0,1,\alpha,-\beta^{-1})$

TABLE I. The three families of contexts needed for our argument. Here, $\alpha, \beta \in \mathbb{Z}_d$ and $\beta \neq 0$.

The following, surprising lemma allows us to dramatically reduce (from exponential to polynomial) the number of hidden variables $\lambda : \mathbb{Z}_d^{2n} \to \mathbb{Z}_d$ we must consider. The hidden variable space Λ can be taken to be the dual vector space of the phase space $\mathbb{Z}_d^{2n}$.

Lemma 1. *Suppose that $n \geq 2$ and that $\lambda : \mathbb{Z}_d^{2n} \to \mathbb{Z}_d$ is a hidden variable with respect to n-particle stabilizer quantum mechanics that is consistent with a quantum state. Then, $\lambda(\boldsymbol{p}, \boldsymbol{q}) = \boldsymbol{\lambda} \cdot (\boldsymbol{p}, \boldsymbol{q})$ for some $\boldsymbol{\lambda} \in \mathbb{Z}_d^{2n}$.*

We then establish a number-theoretic criterion for joint outcomes being in the support of a state $|\Phi\rangle$.

Lemma 2. *The joint outcome (A, B) is impossible for the measurement of $U = W(\boldsymbol{p}, \boldsymbol{q})$ and $V = W(\boldsymbol{p'}, \boldsymbol{q'})$ with $[U, V] = 0$ on the state $|\Phi\rangle$ if and only if $\Psi(m, n)$ is a* permutation polynomial *for all $j, k \in \mathbb{Z}_d$ where*

$$\Psi(m,n) = -mA - nB - 2^{-1}((mp_1 + np'_1)(mq_1 + nq'_1) + (mp_2 + np'_2)(mq_2 + nq'_2)) + \\ j(mp_1 + np'_1) + k(mp_2 + np'_2) + \Phi(j - (mq_1 + nq'_1), k - (mq_2 + nq'_2))$$

A permutation polynomial is one that takes each value in its range equally many times. We will require elements of Dickson's classification of one-variable permutation polynomials of low degree [31].

Theorem 2 (Dickson, 1896)**.** *Suppose $d \not\equiv 1 \ (mod\ 3)$ and $f(x) : \mathbb{Z}_d \to \mathbb{Z}_d$ has degree at most 3. Then, f is a permutation polynomial if and only if $f(x) = ag(x + b) + c$ where $a \neq 0$ and $g(x) = x$ or x^3.*

By applying a Clifford gate to $|\Phi\rangle$, which preserves contextuality properties, we may assume that ϕ_i vanishes for $i \geq 3$. Noting that, by Lemma 1, λ prescribes to $W(\mathbf{p}, \mathbf{q})$ the outcome $\lambda_1 p_1 + \lambda_2 q_2 + \lambda_3 p_2 + \lambda_4 q_2$, the consistency of λ with $|\Phi\rangle$ implies that none of the polynomials in any of the following families are permutation polynomials.

$$\begin{aligned}
\Psi_{\mathrm{I}_\alpha}(m,n) &= m(j - \lambda_2) + n^2(j\phi_2 - 2^{-1}\alpha) + n(\alpha(k - \lambda_4) - \lambda_3 - j^2\phi_1 - 2jk\phi_2) \\
\Psi_{\mathrm{II}_\alpha}(m,n) &= m(k - \lambda_4) + n^2(k\phi_1 - 2^{-1}\alpha) + n(\alpha(j - \lambda_2) - \lambda_1 - k^2\phi_2 - 2jk\phi_1) \\
\Psi_{\mathrm{III}_{\alpha,\beta}}(m,n) &= m(j - \lambda_2 + \beta(k - \lambda_4)) + n^3(\beta^{-1}(\phi_1 - \beta^{-1}\phi_2) + n^2(\beta^{-1}(2^{-1}\alpha - 2j\phi_1 - 2k\phi_2 + \\
&\qquad \beta^{-1}j\phi_2) + k\phi_1) + n(\alpha(k - \lambda_4) + \beta^{-1}(\lambda_3 + j^2\phi_1 + 2jk\phi_2) - \lambda_1 - 2jk\phi_1 - k^2\phi_2)
\end{aligned}$$

We prove that this cannot be the case and, thus, that no hidden variable is consistent with $|\Phi\rangle$.

IX. APPENDIX V: PROOFS

Here, we provide explicit details of the proofs of our results. First, we establish that the algebraic relations between commuting Weyl operators enforce a strong condition on hidden variables: they must be group homomorphisms from phase space to outcomes. This dramatically reduces (from $d^{d^{2n}}$ functions to d^{2n} homomorphisms) the number of hidden variables we need to consider. The assumption of multiple qudits is crucial here.

Lemma 1. *Suppose that $n \geq 2$ and that $\lambda : \mathbb{Z}_d^{2n} \to \mathbb{Z}_d$ is a hidden variable with respect to n-particle stabilizer quantum mechanics that is consistent with a quantum state. Then, $\lambda(\boldsymbol{p}, \boldsymbol{q}) = \boldsymbol{\lambda} \cdot (\boldsymbol{p}, \boldsymbol{q})$ for some $\boldsymbol{\lambda} \in \mathbb{Z}_d^{2n}$.*

Proof. The hidden variable λ prescribes the outcome $\lambda(\mathbf{p}, \mathbf{q})$ to $W(\mathbf{p}, \mathbf{q})$. If λ is consistent with a quantum state, these outcomes must respect the algebraic relations between commuting Weyl operators: $\lambda(\mathbf{p}, \mathbf{q}) + \lambda(\mathbf{p'}, \mathbf{q'}) = \lambda(\mathbf{p{+}p'}, \mathbf{q{+}q'})$ whenever $[(\mathbf{p}, \mathbf{q}), (\mathbf{p'}, \mathbf{q'})] = 0$. First, suppose that $n = 2$. Thus, $\lambda(p_1, q_1, p_2, q_2) = \lambda(p_1, 0, p_2, 0) + \lambda(0, q_1, 0, q_2)$ whenever $p_1 q_1 = -p_2 q_2$. These observations justify the following manipulations.

$$\begin{aligned}
\lambda(1,k,0,0) &= \lambda(1,k,0,-2^{-1}k) + \lambda(0,0,0,2^{-1}k) \\
&= \lambda(1,2^{-1}k,1,-2^{-1}k) + \lambda(0,2^{-1}k,-1,0) + \lambda(0,0,0,2^{-1}k) \\
&= \lambda(1,0,1,0) + \lambda(0,2^{-1}k,0,-2^{-1}k) + \lambda(0,2^{-1}k,0,0) + \lambda(0,0,-1,0) + \lambda(0,0,0,2^{-1}k) \\
&= \lambda(1,0,0,0) + \lambda(0,0,1,0) + \lambda(0,2^{-1}k,0,0) + \lambda(0,2^{-1}k,0,0) + \lambda(0,0,-1,0) \\
&= \lambda(1,0,0,0) + \lambda(0,k,0,0)
\end{aligned}$$

By a similar argument, $\lambda(0,0,1,k) = \lambda(0,0,1,0) + \lambda(0,0,0,k)$. Since $\lambda(p_1, q_1, p_2, q_2) = \lambda(p_1, q_1, 0, 0) + \lambda(0, 0, p_2, q_2)$ it follows that λ is linear. For $n > 2$, a similar argument holds. □

Next, we establish a master equation governing the possibility vs. impossibility of observing a given outcome upon measurement of a pair of commuting Weyl measurements on a state $|\Phi\rangle$. The equation is easily extended to more qudits/measurements and to polynomials Φ of any degree and holds for any dimension d.

Lemma 2. *The joint outcome (A, B) is impossible for the measurement of $U = W(\boldsymbol{p}, \boldsymbol{q})$ and $V = W(\boldsymbol{p'}, \boldsymbol{q'})$ with $[U, V] = 0$ on the state $|\Phi\rangle$ if and only if $\Psi(m, n)$ is a* permutation polynomial *for all $j, k \in \mathbb{Z}_d$ where*

$$\begin{aligned}
\Psi(m,n) = -mA - nB - 2^{-1}((mp_1 + np_1')(mq_1 + nq_1') + (mp_2 + np_2')(mq_2 + nq_2')) + \\
j(mp_1 + np_1') + k(mp_2 + np_2') + \Phi(j - (mq_1 + nq_1'), k - (mq_2 + nq_2'))
\end{aligned}$$

Proof. The eigenvalues of a Weyl operator W are ω^k for $k \in \mathbb{Z}_d$. The projection onto the 1-eigenspace of W is given by $S_0 = d^{-1} \sum_{m \in \mathbb{Z}_d} W^m$; this can be seen by noting that $W^d = \mathbb{I}$, $W^* = W^{-1}$. Therefore, the projection onto the ω^k-eigenspace of W is given by $S_k = d^{-1} \sum_{m \in \mathbb{Z}_d} \omega^{-mk} W^m$.

Measuring $U = W(\mathbf{p}, \mathbf{q})$ and $V = W(\mathbf{p'}, \mathbf{q'})$ with $[U, V] = 0$ and obtaining outcome $(A, B) \in \mathbb{Z}_d^2$ corresponds to the projection:

$$\begin{aligned}
\Pi(A, B | U, V) &= (\sum_{m \in \mathbb{Z}_d} \omega^{-mA} U^m)(\sum_{n \in \mathbb{Z}_d} \omega^{-nB} V^n) \\
&= \sum_{m,n \in \mathbb{Z}_d} \omega^{-mA-nB} U^m V^n \\
&= \sum_{m,n \in \mathbb{Z}_d} \omega^{-mA-nB} W(\mathbf{P}, \mathbf{Q}) \\
&= \sum_{m,n \in \mathbb{Z}_d} \omega^{-mA-nB-2^{-1}(\mathbf{P} \cdot \mathbf{Q})} Z(P_1) X(Q_1) \otimes \ldots \otimes Z(P_n) X(Q_n)
\end{aligned}$$

where $(\mathbf{P}, \mathbf{Q}) = m(\mathbf{p}, \mathbf{q}) + n(\mathbf{p'}, \mathbf{q'})$.

Applying this to a two-qudit magic state $|\Phi\rangle = d^{-1}\sum_{j,k\in\mathbb{Z}_d}\omega^{\Phi(j,k)}|j\rangle|k\rangle$, we obtain:

$$\begin{aligned}\Pi(A,B|U,V)|\Phi\rangle &= d^{-1}\sum_{m,n\in\mathbb{Z}_d}\sum_{j,k\in\mathbb{Z}_d}\omega^{-mA-nB-2^{-1}(P_1Q_1+P_2Q_2)+\Phi(j,k)}Z(P_1)X(Q_1)\otimes Z(P_2)X(Q_2)|j\rangle|k\rangle\\ &= d^{-1}\sum_{m,n\in\mathbb{Z}_d}\sum_{j,k\in\mathbb{Z}_d}\omega^{-mA-nB-2^{-1}(P_1Q_1+P_2Q_2)+\Phi(j,k)}Z(P_1)\otimes Z(P_2)|j+Q_1\rangle|k+Q_2\rangle\\ &= d^{-1}\sum_{m,n\in\mathbb{Z}_d}\sum_{j,k\in\mathbb{Z}_d}\omega^{-mA-nB-2^{-1}(P_1Q_1+P_2Q_2)+\Phi(j-Q_1,k-Q_2)}Z(P_1)\otimes Z(P_2)|j\rangle|k\rangle\\ &= d^{-1}\sum_{j,k\in\mathbb{Z}_d}\sum_{m,n\in\mathbb{Z}_d}\omega^{-mA-nB-2^{-1}(P_1Q_1+P_2Q_2)+jP_1+kP_2+\Phi(j-Q_1,k-Q_2)}|j\rangle|k\rangle\end{aligned}$$

So, observing the outcome (A,B) for the measurements U,V on the state $|\Phi\rangle$ is impossible precisely when the terms

$$\sum_{m,n\in\mathbb{Z}_d}\omega^{-mA-nB-2^{-1}(P_1Q_1+P_2Q_2)+jP_1+kP_2+\Phi(j-Q_1,k-Q_2)}$$

vanish for all $j,k\in\mathbb{Z}_d$. Such a term is a sum of d^2 many primitive d^{th}-roots of unity and vanishes if and only if each d^{th}-root appears d many times. Thus, impossibility of the measurement outcome is equivalent to

$$\begin{aligned}\Psi(m,n) &= -mA-nB-2^{-1}(P_1Q_1+P_2Q_2)+jP_1+kP_2+\Phi(j-Q_1,k-Q_2)\\ &= -mA-nB-2^{-1}((mp_1+np_1')(mq_1+nq_1')+(mp_2+np_2')(mq_2+nq_2'))+\\ &\qquad j(mp_1+np_1')+k(mp_2+np_2')+\Phi(j-(mq_1+nq_1'),k-(mq_2+nq_2'))\end{aligned}$$

being a *permutation polynomial* in m,n for all $j,k\in\mathbb{Z}_d$. □

We are now ready to prove our main theorem.

Theorem 1. *Suppose that the dimension $d\not\equiv 1$ (mod 3). All strong magic states $|\Phi\rangle$ are strongly contextual with respect to stabilizer measurements. The states $C|\Phi\rangle$, where C is any Clifford gate, are also strongly contextual.*

Proof. Suppose $|\Phi\rangle$ is a strong magic state and assume that there is exists a hidden variable $\lambda:\mathbb{Z}_d^{2n}\to\mathbb{Z}_d$ consistent with $|\Phi\rangle$. We will prove that the value prescribed by λ to one of the following measurements is, in fact, impossible to observe of a system $|\Phi\rangle$, contradicting the consistency of λ with $|\Phi\rangle$.

Type	**Operators, up to phase**	**Phase points**
I_α	$Z\otimes\mathbb{I}$ and $\mathbb{I}\otimes Z^\alpha X$	$(1,0,0,0)$ and $(0,0,\alpha,1)$
II_α	$\mathbb{I}\otimes Z$ and $Z^\alpha X\otimes\mathbb{I}$	$(0,0,1,0)$ and $(\alpha,1,0,0)$
$\text{III}_{\alpha,\beta}$	$Z\otimes Z^\beta$ and $X\otimes Z^\alpha X^{-\beta^{-1}}$	$(1,0,\beta,0)$ and $(0,1,\alpha,-\beta^{-1})$

TABLE I. The three families of contexts needed for our argument. Here, $\alpha,\beta\in\mathbb{Z}_d$ and $\beta\neq 0$.

As, by Lemma 1, λ prescribes to $W(p_1,q_1,p_2,q_2)$ the outcome $\lambda_1p_1+\lambda_2q_2+\lambda_3p_2+\lambda_4q_2$, we find that consistency of λ with $|\Phi\rangle$ implies that none of the following are permutation polynomials for all $j,k\in\mathbb{Z}_d$:

$$\begin{aligned}
\Psi_{\mathrm{I}_\alpha}(m,n) &= m(j-\lambda_2) + n^2(j\phi_2 - 2^{-1}\alpha) + n(\alpha(k-\lambda_4) - \lambda_3 - j^2\phi_1 - 2jk\phi_2) \\
\Psi_{\mathrm{II}_\alpha}(m,n) &= m(k-\lambda_4) + n^2(k\phi_1 - 2^{-1}\alpha) + n(\alpha(j-\lambda_2) - \lambda_1 - k^2\phi_2 - 2jk\phi_1) \\
\Psi_{\mathrm{III}_{\alpha,\beta}}(m,n) &= m(j-\lambda_2+\beta(k-\lambda_4)) + n^3(\beta^{-1}(\phi_1 - \beta^{-1}\phi_2) + n^2(\beta^{-1}(2^{-1}\alpha - 2j\phi_1 - 2k\phi_2 + \\
&\qquad \beta^{-1}j\phi_2) + k\phi_1) + n(\alpha(k-\lambda_4) + \beta^{-1}(\lambda_3 + j^2\phi_1 + 2jk\phi_2) - \lambda_1 - 2jk\phi_1 - k^2\phi_2)
\end{aligned}$$

We have dropped all terms constant in m, n. Our proof proceeds by closely analysing these polynomials and repeatedly applying Theorem 2. First, consider $\Psi_{\mathrm{I}_\alpha}(m,n)$. This expression is linear in m. When $j \neq \lambda_2$, this is a permutation polynomial in n for each fixed value of m, so we need only be concerned with the $j = \lambda_2$ case. In this case, Ψ_{I_α} is quadratic in n. By choosing $\alpha = 2\lambda_2\phi_2$, the degree 2 term vanishes and the result is a permutation polynomial only if the linear coefficient is nonzero. Thus, g is inconsistent with $|\Phi\rangle$ unless $\lambda_3 = -\lambda_2(2\lambda_4\phi_2 + \lambda_2\phi_1)$.

By a similar analysis of $\Psi_{\mathrm{II}_\alpha}(m,n)$, we find that by choosing $\alpha = 2\lambda_4\phi_1$, g is inconsistent with $|\Phi\rangle$ unless $\lambda_1 = -\lambda_4(2\lambda_2\phi_1 + \lambda_4\phi_2)$.

Finally, we make the substitutions for λ_1, λ_3 in $\Psi_{\mathrm{III}_{\alpha,\beta}}(m,n)$ and note that, as they too are linear in m, we need only consider the pairs $(j,k) = (-\beta k + \beta\lambda_4 + \lambda_2, k)$. After multiplying through by β, we have:

$$\begin{aligned}
&n^3(\phi_1 - \beta^{-1}\phi_2) + n^2(2^{-1}\alpha + 3k(\beta\phi_1 - \phi_2) + \beta^{-1}\lambda_2\phi_2 + \lambda_4\phi_2 - 2\beta\lambda_4\phi_1 - 2\lambda_2\phi_1) + \\
&n(3\beta k^2(\beta\phi_1 - \phi_2) + k(\beta(\alpha - 4\lambda_2\phi_1 + 2\lambda_4\phi_2) - 4\beta^2\lambda_4\phi_1 + 2\lambda_2\phi_2) + \beta(-\alpha\lambda_4 + \lambda_4^2\phi_2 + \\
&4\lambda_2\lambda_4\phi_1) + \beta^2\lambda_4^2\phi_1 - 2\lambda_2\lambda_4\phi_2)
\end{aligned}$$

We consider two cases. (Note that we may assume that $\phi_1 \not\equiv 0$ by swapping qudits.) First, if $\phi_2 \equiv -1$, then, by choosing $\alpha = 6(\lambda_2\phi_1 - \lambda_4)$ and $\beta = \phi_1^{-1}$, the resulting polynomial factors as:

$$2(n + \phi_1^{-1}(k - \lambda_4))^3.$$

However, if $\phi_2 \not\equiv -1$, we may choose $\alpha = 2(\phi_2+1)^{-1}(\lambda_2\phi_1(\phi_2+2) + \lambda_4(\phi_2^2 - 1))$ and $\beta = \phi_1^{-1}(\phi_2+1)$; the resulting polynomial factors as:

$$(n + \phi_1^{-1}(k - \lambda_4)(\phi_2+1))^3.$$

□

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).
[3] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Nature **510**, 351 (2014).
[4] R. Cleve, P. Hoyer, B. Toner, and J. Watrous, in *Computational Complexity, 2004. Proceedings. 19th IEEE Annual Conference on* (2004) pp. 236–249.
[5] J. Briët, H. Buhrman, T. Lee, and T. Vidick, Quantum Information and Computation **13**, 334 (2013).
[6] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).
[7] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, Phys. Rev. Lett. **104**, 230503 (2010).
[8] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).
[9] S. Abramsky and A. Brandenburger, New J. Phys. **13**, 113036 (2011).
[10] A. C. Elitzur, S. Popescu, and D. Rohrlich, Phys. Lett. A **162**, 25 (1992).
[11] J. Barrett, A. Kent, and S. Pironio, Phys. Rev. Lett. **97**, 170409 (2006).
[12] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, Am. J. Phys **58**, 1131 (1990).

[13] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).

[14] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).

[15] A. A. Klyachko, M. A. Can, S. Binicioğlu, and A. S. Shumovsky, Phys. Rev. Lett. **101**, 020403 (2008).

[16] A. Cabello, S. Severini, and A. Winter, Phys. Rev. Lett. **112**, 040401 (2014).

[17] S. Abramsky and L. Hardy, Phys. Rev. A **85**, 062114 (2012).

[18] S. Abramsky, R. S. Barbosa, and S. Mansfield, arXiv preprint arXiv:1705.07918 (2017).

[19] W. van Dam, *Nonlocality & communication complexity*, Ph.D. thesis, Faculty of Physical Sciences, University of Oxford (1999).

[20] R. Raussendorf, Phys. Rev. A **88**, 022322 (2013).

[21] J. Anders and D. E. Browne, Phys. Rev. Lett. **102**, 050502 (2009).

[22] D. Gottesman, *Stabilizer codes and quantum error correction*, Ph.D. thesis, Caltech (1997).

[23] D. Gottesman, Phys. Rev. A (1998).

[24] D. Gottesman and I. Chuang, Nature **402**, 390 (1999).

[25] E. F. Galvão, Phys. Rev. A **71**, 042302 (2005).

[26] S. X. Cui, D. Gottesman, and A. Krishna, Physical Review A **95**, 012329 (2017).

[27] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).

[28] S. Abramsky, R. S. Barbosa, K. Kishida, L. R., and S. Mansfield, "Possibilities determine the combinatorial structure of probability polytopes," (2016), quant-ph/1603.07735.

[29] A. Acín, T. Fritz, A. Leverrier, and A. B. Sainz, Comm. Math. Phys. **334**, 533 (2015).

[30] D. Gross, J. Math. Phys. **47**, 122107 (2006).

[31] L. E. Dickson, Annals of Mathematics **11**, 65 (1896).

Convex geometry of quantum resource quantification

A general framework for measures of quantum resources

Bartosz Regula[1] *

[1]*School of Mathematical Sciences, University of Nottingham,*
University Park, Nottingham NG7 2RD, United Kingdom

Abstract. We introduce a framework unifying the mathematical characterisation of different measures of general quantum resources and allowing for a systematic way to define a variety of faithful quantifiers for any given convex quantum resource theory. The approach allows us to describe many commonly used measures such as matrix norm–based quantifiers, robustness measures, convex roof–based measures, and witness-based quantifiers together in a common formalism based on the convex geometry of the underlying sets of resource-free states. We provide a detailed characterisation of the measures as well as derive various bounds and relations between them and their duals, generalising and in many cases simplifying results found in the resource theories of quantum entanglement and coherence. We present an explicit application of the results to the resource theories of multi-level coherence, entanglement of Schmidt number k, multipartite entanglement, as well as magic states.

Introduction

Many physical phenomena in quantum information science have gone from being of purely theoretical interest to enjoying a variety of uses as *resources* in quantum information processing tasks. The developments sparked an investigation into the mathematical formulation of such resource theories, aiming to characterise the quantum states and operations that one can use to perform the physical tasks. In particular, it is crucial to be able to *quantify* the given resource, allowing us to discriminate which quantum states are the most useful in the given physical task. Throughout the development of the resource theory of entanglement, various measures were established [1, 21], many of which have been adapted to other resource theories recently [7, 8, 10, 17, 20, 22, 23]. However, defining and characterising the measures of a given quantum resource is usually cumbersome — the investigation of such functions typically has to be approached in a resource-dependent way, and properties such as faithfulness and monotonicity of the quantifiers have to be explicitly verified. Moreover, although some connections between the various quantities are known, there are very few known results which provide a common framework relating them and their features together.

In this work, we introduce a unifying formalism based on the gauge functions of convex sets which significantly simplifies the construction and characterisation of quantifiers of general quantum resources. Gauge functions, a fundamental tool in functional and convex analysis [24, 25], have recently seen a surge of popularity in optimisation research after a framework for linear inverse problems based on the so-called atomic gauge functions was introduced [26]. We apply a similar formalism to the quantification of quantum resources, establishing a consolidated view of many resource quantifiers. In particular, we show that many commonly used and well-known quantifiers — such as ones based on matrix norms, measures built through the convex roof, the so-called robustness measures, as well as various witness-based quantifiers — are all examples of such atomic gauge functions, allowing us to relate them in a common geometric framework. This allows us to establish an extensive family of quantifiers for any given quantum resource, introduce easily verifiable criteria for a measure to satisfy desirable properties such as faithfulness and strong monotonicity under relevant free operations, and generalise known measures to new quantum resources very easily. Further, we show that many relations and bounds between the measures, some of which known in the resource theories of entanglement and coherence, are in fact universal among quantum resources, and the proofs of such properties can be simplified in the present framework.

The formalism presented in this work applies to general finite-dimensional resource theories with a convex set of resource-free states, which is a common and intuitive assumption [16, 17]. A particularly useful case of such resources, and one that we will focus on, is when the set of free states is obtained as the convex hull of free pure states. One can readily apply our results to any given resource theory constructed in this way. In our examples, we consider some representative examples of such theories — quantum entanglement, quantum coherence, and magic states — obtaining new results in the quantification of the resources. In addition to the characterisation of quantifiers already defined in the literature, we introduce several new measures, such as: a measure of multi-level quantum coherence which generalises the ℓ_1 norm of coherence [7], faithful quantifiers of magic [13], measures of bipartite entanglement of Schmidt number k and k-partite entanglement which generalise the convex roof–extended negativity [27], as well as a class of norms which generalise the greatest cross norm [28] to the hierarchy of k-partite entanglement, with computable formulas for genuine multipartite entanglement. We additionally show that many proofs and properties of such measures are significantly simplified in this formalism, deriving novel results for quantifiers such as robustness of Schmidt rank k entanglement [29] and robustness of k-coherence [30].

*bartosz.regula@gmail.com

General framework

The basic tool of our approach are the so-called *gauge functions* [25]. Given a convex set $\mathcal{C}$, the gauge function $\gamma_{\mathcal{C}}$ is defined as

$$\gamma_{\mathcal{C}}(\rho) = \inf_{\lambda \geq 0} \ \rho \in \lambda \mathcal{C}, \tag{1}$$

that is, the least amount that the set $\mathcal{C}$ has to "grow" in order to contain ρ inside it. We will in particular deal with gauge functions of sets which are defined as the convex hull of some non-convex set $\mathcal{S}$ (for example, a subset of pure quantum states). We then define the *atomic gauge function* [26] as the gauge function of the convex hull:

$$A_{\mathcal{S}}(\rho) = \gamma_{\mathrm{conv}(\mathcal{S})}(\rho). \tag{2}$$

Such functions satisfy many useful properties, as we will show.

We define a general convex resource theory as follows: we begin with a set of free pure states $\mathcal{V} \subseteq \mathbb{C}^d$, which satisfy some property defining the given resource theory. We then define the set of free pure-state density matrices: $\mathcal{S}_+ = \left\{ |\psi\rangle \langle\psi| \;\middle|\; |\psi\rangle \in \mathcal{V} \right\}$ such that any free (mixed) state is in the convex hull of $\mathcal{S}_+$. We can now define a variety of quantifiers by taking the atomic gauges corresponding to different sets, and in fact we can show that many well-known monotones belong to the gauge function formalism. We will introduce several examples to demonstrate the versatility of our approach.

Let us begin with the simplest case, that is, an atomic gauge for the set of pure states: $A_{\mathcal{V}}$. This quantity has the advantage that it is, in general, much easier to compute than gauges defined for density matrices. It corresponds to well-known quantifiers of pure-state quantum resources — in the resource theory of entanglement, it is equal to the sum of Schmidt coefficients [33], and in the resource theory of coherence it corresponds to the ℓ_1 norm [7].

To define atomic gauge functions for mixed density matrices, we start by taking the set $\mathcal{S}_+ \cup (-\mathcal{S}_+)$, that is, $\mathcal{S}_+$ symmetrised around the origin. The gauge function $A_{\mathcal{S}_+ \cup (-\mathcal{S}_+)}$, up to a constant, is nothing but the **robustness** $R_{\mathcal{S}_+}$ — a fundamental measure defined first in the theory of entanglement [34]. Similarly, if we consider the atomic gauge of the set $\mathcal{S}_+ \cup (-\mathbb{D})$, where $\mathbb{D}$ is the set of all density matrices, we obtain the **generalised robustness** $R^G_{\mathcal{S}_+}$ — an extension of the robustness which was found to have useful operational interpretations not only in the resource theories of entanglement [35] and coherence [22, 36], but also in generalised frameworks for resource theories [17, 37]. We further show that many other quantifiers in this framework, such as experimentally-friendly families of witness-based measures [38–40], can be defined simply by taking the gauge function $A_{\mathcal{S}_+ \cup \mathcal{X}}$ with different choices of the set $\mathcal{X}$.

Additionally, one can define resource quantifiers through **matrix norms** (or more general matrix gauge functions). Examples of such measures include the greatest cross norm for bipartite entanglement [28, 41], the ℓ_1 norm of coherence [7], or the Schmidt operator norm [42]. We show that these measures are also gauge functions, and they can in fact be straightforwardly defined for *any* resource theory as the atomic gauge of the set $\mathcal{S} = \left\{ |\alpha\rangle \langle\beta| \;\middle|\; |\alpha\rangle, |\beta\rangle \in \mathcal{V} \right\}$. We will denote such norm-based quantifiers as $A_{\mathcal{S}}$.

Finally, an important concept in quantum resource quantification is the **convex roof**. Many common measures are defined using this concept — they include quantifiers such as the entanglement of formation (and concurrence) [43], convex-roof extended negativity [27], coherence of formation [44], or coherence concurrence [45, 46]. If we have a gauge function $A_{\mathcal{V}}$ which is defined for pure states only, we can extend it to all mixed states by minimising over all pure-state decompositions:

$$A^{\cup}_{\mathcal{S}_+}(\rho) = \inf_{\{p_i, |\psi_i\rangle\}} \sum_i p_i A_{\mathcal{V}}(|\psi_i\rangle)^2 \tag{3}$$

where the minimisation is over all ensembles such that $\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i|$ with $\sum_i p_i = 1$. In fact, we show the convex roof to also be a gauge function, allowing us to relate it to the other quantifiers easily.

Properties and results

We prove many useful properties of the measures defined in the above formalism. In the following, we summarise the main results of our work. Precise statements of the theorems, including some assumptions that we omit here for brevity, can be found in the manuscript [47].

Theorem 1 *For any resource theory, the quantifiers are faithful, that is, we have that* $\rho \in \mathrm{conv}(\mathcal{S}_+)$ *if and only if* $A^{\cup}_{\mathcal{S}_+}(\rho) - 1 = R_{\mathcal{S}_+}(\rho) = A_{\mathcal{S}}(\rho) - 1 = R^G_{\mathcal{S}_+}(\rho) = 0.$

Theorem 2 *For any resource theory, the quantifiers satisfy strong monotonicity under free operations: for a given quantum channel* $\Gamma(\rho) = \sum_i \Lambda_i(\rho)$, *where each* Λ_i *is a relevant free operation (subchannel) in the given resource theory, we have that*

$$\sum_i \mathrm{Tr}\left(\Lambda_i(\rho)\right) A\left(\frac{\Lambda_i(\rho)}{\mathrm{Tr}\left(\Lambda_i(\rho)\right)}\right) \leq A(\rho) \tag{4}$$

where A *is any one of the gauge-based quantifiers.*

The two results above show that the atomic gauge functions defined in this formalism are all valid measures of the given resource. This is remarkable, given that proving the monotonicity and faithfulness of measures defined in a more ad-hoc way is often extremely cumbersome — here, the result follows "for free" from the gauge function formalism, and can be applied to any given resource theory. In particular, it establishes a family of measures which can faithfully and reliably detect and quantify quantum resources.

Theorem 3 *For any resource theory, the quantifiers bound each other as*

$$R_{\mathcal{S}_+}(\rho) \geq A_{\mathcal{S}}(\rho) - 1 \geq R^G_{\mathcal{S}_+}(\rho), \quad A^{\cup}_{\mathcal{S}_+}(\rho) \geq A_{\mathcal{S}}(\rho) \tag{5}$$

This result immediately allows us to relate the introduced quantifiers with each other, and it generalises some known bounds for the robustness of coherence and the ℓ_1 norm of coherence [22] as well as quantitative relations between measures such as the robustness of entanglement and the greatest cross norm [28, 41].

Theorem 4 *For any resource theory, the quantifiers reduce to the vector atomic gauge $A_{\mathcal{V}}$ for pure states: $A^{\cup}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) - 1 = A_{\mathcal{S}}(|\psi\rangle\langle\psi|) - 1 = R^{G}_{\mathcal{S}_+}(|\psi\rangle\langle\psi|) = A_{\mathcal{V}}(|\psi\rangle)^2 - 1$.*

This result explicitly shows that the quantifiers are all closely related to each other, and the bounds obtained in Theorem 3 are in fact tight. Additionally, it shows that the quantification of pure-state resources is always simplified, which is frequently a non-trivial fact to show for a given resource theory. Note that the gauge function $A_{\mathcal{V}}$ is often significantly easier to compute that the general forms of the quantifiers, in many cases leading to an analytical characterisation of pure-state resources.

In addition to the above results, in the paper we relate the introduced quantifiers to other types of measures, obtaining bounds and relations between gauge functions and quantifiers such as geometric resource measures based on quantum fidelity [48] as well as distance measures based on relative entropy. Further, we show that many classes of witness-based measures, experimentally-friendly quantification methods based on optimising witness operators [38], correspond in fact to gauge functions and therefore share many of the simplified properties investigated herein. These results, again, generalise many quantitative results to arbitrary resource theories, and allow for a direct comparison between gauge-based measures and other common quantifiers.

It is also important to note that each atomic gauge function A has an associated *dual* gauge function A°. These functions can be used to characterise the witnesses of the given resource — for example, the witnesses of Schmidt rank k entanglement are called k-block positive operators and have a close relation with k-positive maps [42, 49, 50]. We obtain many quantitative results regarding the dual gauge functions, and in particular we show that for any positive semidefinite matrix X, the dual quantity $A^{\circ}(X)$ is the same regardless of which of the quantifiers $A^{\cup}_{\mathcal{S}_+}$, $A_{\mathcal{S}}$, $R_{\mathcal{S}_+}$, $R^{G}_{\mathcal{S}_+}$ we choose, therefore establishing a further dual relation between the measures. The dual quantifiers can be quantitatively related to the aforementioned geometric measures [48] as well.

We remark that the choice of quantifiers that we investigated — $A^{\cup}_{\mathcal{S}_+}$, $R_{\mathcal{S}_+}$, $A_{\mathcal{S}}$, and $R^{G}_{\mathcal{S}_+}$ — is by no means unique. In fact, the formalism allows one to define atomic gauges for other chosen sets, and the tools that we introduced can be used to investigate such measures straightforwardly, allowing us to establish easily verifiable criteria for any gauge-based measure to satisfy desirable properties such as faithfulness and strong monotonicity under relevant free operations.

Hierarchies of entanglement and coherence

We have discussed the applications of our quantifiers to the resource theories of entanglement and coherence. However, in some physical tasks, not every entangled state (or coherent state) is useful as a resource — one might require entanglement of a particular Schmidt rank, or entanglement between many parties in a multipartite system, or coherence between multiple levels of a quantum system [10, 30, 42, 51–55]. We can then define a hierarchy of free states: for example, let $\mathcal{S}^k$ be the set of pure states which have at most k non-zero Schmidt coefficients (Schmidt rank k). All separable states then form the convex hull of $\mathcal{S}^1$, all density matrices form the convex hull of $\mathcal{S}^d$, and in general we have $\mathcal{S}^1 \subset \mathcal{S}^2 \subset \cdots \subset \mathcal{S}^d$ [56]. A very similar hierarchy can be defined for coherence, where the coherence rank is defined to be the number of non-zero coefficients of a state $|\psi\rangle$ in a given basis, and the convex hull $\mathcal{C}^k$ of states with a given coherence rank gives us $\mathcal{C}^1 \subset \cdots \subset \mathcal{C}^d$. Finally, an important hierarchy of the same kind is formed by the sets of k-partite entangled states [57, 58].

An advantage of our approach based on convex geometry is that it easily generalises to such hierarchies of resources. The definitions of many measures straightforwardly extend to each level of the hierarchy, making it easy to quantify a specific "rank" of a given quantum resource.

In particular, the application of results that we outlined in the previous section immediately allows us to obtain many quantitative relations. We establish exact formulas for quantities such as the robustness of k coherence, robustness of Schmidt rank k entanglement [29], and norm-based measures of the resources. We introduce novel quantifiers, such as a convex roof measure of Schmidt rank k entanglement which generalises the convex roof-extended negativity, a measure of k-coherence which generalises the ℓ_1 norm of coherence, and a hierarchy of norms which are faithful quantifiers of k-partite entanglement generalising the greatest cross norm. Further, we show that the quantification of gauge-based measures of genuine multipartite entanglement on pure states in fact reduces to the so-called genuine multipartite negativity [59].

Further, we show an application of the gauge function framework to the resource theory of magic states [13, 14, 23, 60–64], providing computable measures of the resource and establishing quantitative relations between them.

Conclusions

We have introduced a general formalism which allows one to define a variety of quantifiers for any general convex quantum resource. We have shown that the quantifiers defined in this way are closely related to each other, establishing quantitative relations between different types of measures. Additionally, quantifiers in the framework are guaranteed to satisfy desirable criteria such as monotonicity and faithfulness, and exhibit useful properties such as simplified formulas for pure states.

The universality of the formalism makes it applicable to many different resource theories — since the quantitative results can be easily adapted to any chosen convex resource theory, it provides insight into the general structure of such theories. We therefore believe that the framework will complement the recent efforts to establish a mathematical formalism of general quantum resource theories [15, 17, 19, 20].

References

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

[2] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Phys. Rev. Lett. **111**, 250404 (2013).

[3] G. Gour and R. W. Spekkens, New J. Phys. **10**, 033023 (2008).

[4] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, Phys. Rev. X **5**, 021001 (2015).

[5] M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. A **67**, 062104 (2003).

[6] J. Aberg, (2006), arXiv:quant-ph/0612146 .

[7] T. Baumgratz, M. Cramer, and M. B. Plenio, Phys. Rev. Lett. **113**, 140401 (2014).

[8] A. Streltsov, G. Adesso, and M. B. Plenio, (2016), arXiv:1609.02439 .

[9] W. Vogel and J. Sperling, Phys. Rev. A **89**, 052302 (2014).

[10] T. Theurer, N. Killoran, D. Egloff, and M. B. Plenio, (2017), arXiv:1703.10943 .

[11] R. Gallego and L. Aolita, Phys. Rev. X **5**, 041008 (2015).

[12] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, P. Joshi, W. Kłobus, and A. Wójcik, Phys. Rev. Lett. **112**, 120401 (2014).

[13] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, New J. Phys. **16**, 013009 (2014).

[14] M. Ahmadi, H. B. Dang, G. Gour, and B. C. Sanders, (2017), arXiv:1706.03828 .

[15] M. Horodecki and J. Oppenheim, Int. J. Mod. Phys. B **27**, 1345019 (2012).

[16] J. Sperling and W. Vogel, Phys. Scr. **90**, 074024 (2015).

[17] F. G. S. L. Brandão and G. Gour, Phys Rev Lett **115**, 070503 (2015).

[18] L. del Rio, L. Kraemer, and R. Renner, (2015), arXiv:1511.08818 .

[19] B. Coecke, T. Fritz, and R. W. Spekkens, Information and Computation Quantum Physics and Logic, **250**, 59 (2016).

[20] G. Gour, Phys. Rev. A **95**, 062314 (2017).

[21] M. B. Plenio and S. Virmani, Quantum Info. Comput. **7**, 1 (2007).

[22] M. Piani, M. Cianciaruso, T. R. Bromley, C. Napoli, N. Johnston, and G. Adesso, Phys. Rev. A **93**, 042107 (2016).

[23] M. Howard and E. Campbell, Phys. Rev. Lett. **118**, 090501 (2017).

[24] W. Rudin, *Functional Analysis*, 2nd ed. (McGraw-Hill, New York, 1991).

[25] R. T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, 1970).

[26] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, Found Comput Math **12**, 805 (2012).

[27] S. Lee, D. P. Chi, S. D. Oh, and J. Kim, Phys. Rev. A **68**, 062304 (2003).

[28] O. Rudolph, Journal of Mathematical Physics **42**, 5306 (2001).

[29] N. Johnston and D. W. Kribs, Houst. J. Math. , 831 (2015), arXiv:1304.2328 .

[30] M. Ringbauer, T. R. Bromley, M. Cianciaruso, S. Lau, G. Adesso, A. G. White, A. Fedrizzi, and M. Piani, (2017), arXiv:1707.05282 .

[31] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275 (1997).

[32] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, 2007).

[33] G. Vidal and R. F. Werner, Phys. Rev. A **65**, 032314 (2002).

[34] G. Vidal and R. Tarrach, Phys. Rev. A **59**, 141 (1999).

[35] M. Steiner, Phys. Rev. A **67**, 054305 (2003).

[36] C. Napoli, T. R. Bromley, M. Cianciaruso, M. Piani, N. Johnston, and G. Adesso, Phys. Rev. Lett. **116**, 150502 (2016).

[37] N. Datta, Int. J. Quantum Inform. **07**, 475 (2009).

[38] F. G. S. L. Brandão, Phys. Rev. A **72**, 022310 (2005).

[39] O. Gühne, M. Reimpell, and R. F. Werner, Phys. Rev. Lett. **98**, 110502 (2007).

[40] J. Eisert, F. G. S. L. Brandão, and K. M. R. Audenaert, New J. Phys. **9**, 46 (2007).

[41] O. Rudolph, J. Phys. A: Math. Gen. **33**, 3951 (2000).

[42] N. Johnston and D. W. Kribs, Journal of Mathematical Physics **51**, 082202 (2010).

[43] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).

[44] A. Winter and D. Yang, Phys. Rev. Lett. **116**, 120404 (2016).

[45] X. Yuan, H. Zhou, Z. Cao, and X. Ma, Phys. Rev. A **92**, 022124 (2015).

[46] X. Qi, T. Gao, and F. Yan, J. Phys. A: Math. Theor. **50**, 285301 (2017).

[47] B. Regula, (2017), arXiv:1707.06298 .

[48] T.-C. Wei and P. M. Goldbart, Phys. Rev. A **68**, 042307 (2003).

[49] D. Chruściński and A. Kossakowski, Commun. Math. Phys. **290**, 1051 (2009).

[50] Ł. Skowronek, E. Størmer, and K. Życzkowski, Journal of Mathematical Physics **50**, 062106 (2009).

[51] B. M. Terhal and P. Horodecki, Phys. Rev. A **61**, 040301 (2000).

[52] N. Killoran, F. E. S. Steinhoff, and M. B. Plenio, Phys. Rev. Lett. **116**, 080402 (2016).

[53] S. Chin, (2017), arXiv:1702.03219 .

[54] S. Chin, (2017), arXiv:1702.06061 .

[55] B. Regula, M. Piani, M. Cianciaruso, T. R. Bromley, A. Streltsov, and G. Adesso, (2017), arXiv:1704.04153 .

[56] A. Sanpera, D. Bruß, and M. Lewenstein, Phys. Rev. A **63**, 050301 (2001).

[57] O. Gühne, G. Tóth, and H. J. Briegel, New J. Phys. **7**, 229 (2005).

[58] M. Walter, D. Gross, and J. Eisert, (2016), arXiv:1612.02437 .

[59] M. Hofmann, T. Moroder, and O. Gühne, J. Phys. A: Math. Theor. **47**, 155301 (2014).

[60] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, New J. Phys. **14**, 113011 (2012).

[61] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Nature **510**, 351 (2014).

[62] N. Delfosse, P. Allard Guerin, J. Bian, and R. Raussendorf, Phys. Rev. X **5**, 021003 (2015).

[63] S. Bravyi, G. Smith, and J. A. Smolin, Phys. Rev. X **6**, 021043 (2016).

[64] R. Raussendorf, D. E. Browne, N. Delfosse, C. Okay, and J. Bermejo-Vega, Phys. Rev. A **95**, 052334 (2017).

Quantum walks

Peter Høyer

Calgary, Canada

Abstract: We give an introduction to quantum walks. No prior knowledge is assumed. We discuss the construction of quantum walks and their correspondance to and with random walks. We cover the main concepts and ideas when working with quantum walks. We explain how quantum walks relate to quantum search. We discuss applications in algorithmics and communication complexity, and state some of the main open questions.

Higher order quantum operations of unitaries and their implications

Mio Murao

Tokyo, Japan

Abstract: A supermap is a transformation from a map to a map. Transforming a unitary to its inversed, transposed, complex conjugated and controlled unitary are examples of supermaps. We consider the case where the input and output maps are quantum operations and perform the output map directly by applying the input map given by a quantum black box, a quantum system implementing an unknown quantum operation, together with a sequence of input-independent fixed quantum operations called a quantum comb. We regard such direct implementations of supermaps for quantum operations in quantum mechanics as higher order quantum operations. General properties required for achieving higher order quantum operations can be formulated by the framework for quantum networks based on quantum combs proposed by Chiribella et al. There are several known no-go theorems for higher order quantum operations with a single use of the black box. If infinite uses of the black box is allowed, the full classical description of the input is obtained, therefore it is possible to achieve higher order quantum operations by implementing the output map calculated by applying the supermap on the classical description of the input map. However, it is not well known which supermaps are achievable with finite uses of the black box. In this talk, we present go-results for higher order quantum operations of unitaries with finite uses of the black box. We mainly focus on a universal quantum algorithm for performing a complex conjugate unitary and present its implications representing new characteristic features of quantum mechanics exhibited in higher order quantum operations of unitaries.

Bell correlations in many-body systems

Nicolas Sangouard[1] * Sebastian Wagner[1] Roman Schmied[2] Batiste Allard[2]
Matteo Fadel[2] Valerio Scarani[3 4] Philipp Treutlein[2] Jean-Daniel Bancal[1]

[1] *Quantum Optics Theory Group, Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel*
[2] *Quantum Atom Optics Lab, Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel*
[3] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543*
[4] *Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542*

Abstract. How can one characterize the quantum correlations between the constituent particles of a many-body system? Although entanglement is routinely observed in many systems, we will focus on the detection of a subset of quantum correlations – namely Bell correlations. We will derive Bell correlation witnesses from many-particle Bell inequalities involving only one- and two-body correlation functions. We will address the question of the statistics required to witness Bell correlated states in practice and we will show first experimental results successfully reporting on the violation of a Bell correlation witness between hundreds of spins.

Keywords: Theory of quantum entanglement and nonlocality, many-body systems

In 1964, John Bell proposed an experimental test in which two black-boxes receiving classical inputs and producing classical outputs only, can certify that the correlations between the outputs cannot be explained by classical means [1]. While these results are fundamentally appealing to test the limits of classical physics as a complete description of Nature, it has been realized in 1992 that Bell test can be used to certify that the state on which the black boxes operate is a well identified entangled state [2, 3]. As entanglement is at the core of secure communication, the use of a Bell test is nowadays seen as an appealing technique to certify the security of communication tools independently of the details and imperfections of the actual implementations [4, 5].

Although Bell tests occupy a privileged position in physics at the interface between fundamental and applied physics, Bell inequalities have been tested in small systems only. Whereas new forms of correlations are known to arise in presence of a larger number of parties [6], testing a Bell inequality on many parties is technically challenging. Indeed, a Bell test requires addressing of individual particles, which is seldom possible when dealing with more than a few tens of particles. The number of measurements that need to be performed also increases rapidly with the number of parties, and multipartite Bell inequalities typically involve many-body correlations functions, which are difficult to evaluate on systems involving many particles.

Inspired by recent results [7], we consider here the situation in which well-characterized collective measurements are performed on an ensemble of particles. Using few-body correlator inequalities, we construct witness operators for Bell correlated quantum states, i.e. states violating a Bell inequality. These witnesses only involve up to the second moment of collective measurements and are thus suitable for experimental tests on large systems. We address the question of the statistics required to witness Bell correlated states in many-body systems [8] and we show first experimental results successfully reporting on the first violation of a Bell correlation witness in a spin-squeezed Bose-Einstein condensate [9].

References

[1] J.S. Bell, Physics 1, 195 (1964)

[2] L. Braunstein, A. Mann, M. Revzen, Phys. Rev. Lett. 68, 3259 (1992)

[3] S. Popescu and D. Rohrlich, Phys. Lett. A 169, 411 (1992)

[4] J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. 95, 050103 (2005)

[5] A. Acin et al. Phys. Rev. Lett. 98, 230501 (2007)

[6] G. Svetlichny, Phys. Rev. D 35, 3066 (1987)

[7] J. Tura, R. Augusiak, A. B. Sainz, T. Vertesi, M. Lewenstein, A. Acin, Science 344, 1256 (2014)

[8] S. Wagner, R. Schmied, M. Fadel, P. Treutlein, N. Sangouard, and J.-D. Bancal arXiv:1604.07810

[9] R. Schmied, J.-D. Bancal, B. Allard, M. Fadel, V. Scarani, P. Treutlein and N. Sangouard, Science 352, 441 (2016)

*nicolas.sangouard@unibas.ch

The information cost of quantum memoryless protocols*

André Chailloux[1] [†] Iordanis Kerenidis[2] [‡] Mathieu Laurière[3] [§]

[1] *Inria, Paris*
[2] *CNRS IRIF, Université Paris 7*
[3] *NYU-ECNU Institute of Mathematical Sciences at NYU Shanghai*

Abstract. We consider memoryless quantum communication protocols, where the two parties do not possess any memory besides their classical input and they take turns performing unitary operations on a pure quantum state that they exchange between them. Most known quantum protocols are of this type and recently a deep connection between memoryless protocols and Bell inequality violations has been explored in [8]. We study the information cost of such protocols by looking at a canonical problem: bounded-round quantum communication for the one-bit AND function. We prove directly a tight lower bound of $\Theta(\frac{\log k}{k})$ for the information cost of AND for k-round memoryless quantum protocols and for the input distribution needed for the Disjointness function.

Keywords: communication complexity, information complexity, quantum protocols

1 Context

In the model of communication complexity, two players, Alice and Bob, receive inputs and would like to solve some distributed task that depends on these inputs, while minimizing the number of bits they exchange. This model has deep connections to many areas of computer science, including data structures, circuit lower bounds and streaming algorithms [17]. Recently, a lot of attention has been given to a different measure of complexity for communication protocols, namely the amount of information that is leaked about the players inputs during the protocol. The information cost of a protocol is always lower than the communication cost, since one communicated bit can carry at most one bit of information. It has proved to be one of the strongest techniques we have to lower bound the communication complexity of functions [2, 5, 3, 14].

One can also define the notions of communication and information complexity in the quantum setting, where the two players exchange quantum messages. While it is straightforward to define the communication cost of a quantum protocol as the number of qubits that the two players exchange, one has to be careful when defining the information cost of a quantum protocol. Besides some application-specific definitions [13, 12], recently two main definitions have been put forward. Touchette [21] has defined a notion of quantum information cost (QIC) and has proved that it has a number of important properties, including that for any function the quantum information complexity, namely the information cost of the optimal quantum protocol that solves the function, equals the amortized communication complexity of the function. Kerenidis et al. [15] proposed a different notion, the classical input information cost (CIC), that is more intuitively related to the information leakage of the protocol, but is smaller than the QIC notion (hence it is a weaker lower bound on communication complexity). Very recently, [19] clarified the relation between the two notions showing that while CIC measures how much information each player learns about the other's input during the protocol, the QIC measures, on top of this, the information the players forget during the protocol. While our understanding of the flow of information during a quantum protocol has deepened, these notions remain difficult to use in practice. The main reason is that mathematically they both involve a quantum conditional mutual information, where the conditioning is on a quantum variable. This quantity is notoriously difficult to handle, even though there has been some recent breakthrough work on it [9].

Here, we try to overcome this difficulty by looking at a rich subclass of protocols that we call *memoryless* protocols. In these protocols, the two parties take turns performing unitary operations on a pure quantum state that they exchange between them. They do not possess any memory and hence when they send a message they do not keep anything in their private space apart from their classical input. There are many reasons why it is interesting to look at such protocols. First, almost all quantum protocols we know are memoryless. This includes all protocols in the simultaneous message passing model, eg. fingerprints for Equality [6], and in the one-way model, e.g. Hidden Matching [1, 10], but also the two-way protocol for Disjointness in [7] and for Vector in Subspace [20, 16]. Second, there is a deep connection between memoryless protocols and Bell inequality violations that has been explored in [8, 18]. Third, moving towards implementations of quantum communication protocols and the realization of quantum networks, memoryless protocols can be much easier to implement as it has already been shown [22, 11]. Last but not least, it may be easier to understand the flow of quantum information in memoryless protocols. For example, it is easy to see that the relation between CIC and QIC is in this case clear: for any memoryless protocol, QIC is exactly two times CIC, since the players forget exactly as much as they learn. Note that forgetting is not necessarily a drawback of quantum protocols: forgetting is necessary

[†] andre.chailloux@inria.fr

[‡] jkeren@liafa.univ-paris-diderot.fr

[§] mathieu.lauriere@nyu.edu

in order to obtain quantum communication speed-ups for some problems [19].

2 Contributions

In this work, we initiate the study of the information cost of memoryless quantum protocols by looking at a canonical problem: bounded-round quantum communication for the AND function on two bits. One of the main reasons to study the AND function is its close relation to the Disjointness problem (DISJ), where the players receive one set each and their goal is to decide whether these two sets are disjoint. One can see DISJ as a function that takes as inputs two n-bit strings x, y and returns the OR of the coordinate-wise AND of these strings, i.e. $\mathrm{DISJ}(x, y) = \mathrm{OR}(\mathrm{AND}(x_1, y_1), \ldots, \mathrm{AND}(x_n, y_n))$. In the classical world, a very elegant lower bound for Disjointness using information-theoretic tools was given by Bar-Yossef et al. [2] and its proof consists of the following two steps. First, one reduces DISJ to AND: given a protocol for DISJ on inputs of size n, the players construct a protocol for AND as follows: they embed their one-bit inputs for AND in some random coordinate for DISJ, use their private coins to pick the remaining $(n-1)$ inputs uniformly from $\{(0,0), (0,1), (1,0)\}$, and run the DISJ protocol. The output of DISJ for such inputs is the same as the output of the AND function. One can show this way that if the information cost of the DISJ protocol is I, then the information cost of the new protocol for AND is I/n. This implies the information complexity of DISJ is at least n times the information complexity of AND for the above input distribution. The second stage of the proof involves computing directly the information complexity of the AND function, and showing to be at least a constant, the tight $\Omega(n)$ lower bound for DISJ is obtained.

In the quantum world, things are considerably more complicated. The first attempt to provide an information-theoretic proof of the bounded-round quantum communication complexity of DISJ was by Jain et al. [13]. In their work, they introduced a different information-theoretic notion from QIC and CIC and used it to reduce the DISJ problem to the AND problem. By directly lower bounding this quantity for the AND function they managed to show that any k-round protocol for DISJ has communication cost $\Omega(n/k^2)$. There is no clear way to improve this lower bound using their information-theoretic notion and this bound falls short of the optimal bound of $\Omega(n/k)$. Very recently, [4] provided a proof which gives a bound of $\widetilde{\Omega}(n/k)$ for k-round protocols for DISJ by reducing DISJ to AND and then using the already known lower bound for DISJ to lower bound the complexity of AND. This proof does not provide a direct proof for the information complexity of AND.

Here we focus on the subclass of memoryless protocols and prove directly a tight lower bound for the information complexity of AND for memoryless protocols and for the input distribution needed for DISJ. More precisely, considering the input distribution $\mathcal{U}_0$ defined by $\mathcal{U}_0(x, y) = \frac{1}{3}$ for $(x, y) \neq (1, 1)$ and $\mathcal{U}_0(1, 1) = 0$, we show the following, where $\mathrm{CIC}^{\mathrm{ML}}_{\mathcal{U}_0,\varepsilon,k}(\mathrm{AND})$ is the minimum CIC achieved by a k-round memoryless quantum protocol computing AND with error at most ε on input distribution $\mathcal{U}_0$.

Theorem 1 *For any $\varepsilon \in (0, 1/2)$ and any integer k, $\mathrm{CIC}^{\mathrm{ML}}_{\mathcal{U}_0,\varepsilon,k}(AND) = \Theta_\varepsilon\left(\frac{\log(k)}{k}\right)$.*

The upper bound in Theorem 1 comes from a protocol described in [4] credited to Jain, Radhakrishnan and Sen. Note also that from [13], we could obtain a non-optimal bound of $\mathrm{CIC}^{\mathrm{ML}}_{\mathcal{U}_0,\varepsilon,k}(\mathrm{AND}) = \Omega_\varepsilon(1/k)$, since the information-theoretic notion used in [13] becomes equivalent to CIC for memoryless protocols.

The question is then whether we can lift the lower bound of Theorem 1 to memoryless quantum protocols for DISJ. The obvious way to try and do it is to start with a memoryless quantum protocol for DISJ and use it in order to construct a memoryless protocol for AND. However, there is an issue: to solve AND, the players are given one-bit inputs, say x and y. But if they want to use a protocol solving DISJ over n bits, they need to create $n-1$ inputs for each party distributed in a way such that the protocol for DISJ will actually compute $\mathrm{AND}(x, y)$. In the classical case, the players use private coins to choose the remaining inputs for DISJ, when we embed the AND function to it. In [13], the players used a superposition of coins in order to choose these inputs. Now, if the players keep these superpositions in their workspace, then we lose the memoryless property of our protocols. On the other hand, if they send these superpositions to the other player, the information cost of the protocol might considerably increase.

Since it is not obvious how to reduce DISJ to AND while retaining the memoryless property, similarly to the classical case where we do not know how to perform the reduction without the use of private coins, we slightly enhance our model. More precisely, we look at the model where the players do not possess any memory and hence they do not keep anything in their private space apart from their classical input *and* some classical private coins. Note that one can also assume the players share public coins without changing the model. In the classical case, we do allow for private coins when we define the information cost of a protocol. In the quantum case, we cannot unitarily create classical coins. Allowing classical coins seems like a minimal addition to the model. One can see that the communication complexity in this new model is not different from the communication complexity in the model without coins. Indeed, any protocol with coins can be simulated by a protocol where the coins are created in superposition by the players without changing the communication cost. But what about the information complexity? On the one hand, the information complexity cannot increase, since we can always ignore the coins. Surprisingly, we show that it becomes as small as it can possibly be, namely, it equals the information revealed just by the value of the function. So any function can be computed privately. In fact, we show that every quantum protocol can be turned into a quantum protocol with coins that has the same input-output behaviour as the original protocol and that is perfectly private, i.e. the players only learn the value of the function the protocol computes and nothing more.

Theorem 2 *For every quantum communication protocol* Π*, there exists a memoryless quantum protocol* Π' *with private classical coins such that: on every input pair* (x, y)*,* Π' *has the same output distribution as* Π*, and the information cost of* Π' *is only the information gained by Bob's output* Π_{out} *in* Π*. This means that for every input distribution* μ*, we have* $\mathrm{CIC}_\mu(\Pi') = I\left(\Pi_{out}(X, Y) : X|Y\right)$*, where* (X, Y) *is a random variable distributed according to* μ*,* I *denotes the (classical) conditional mutual information and* $\Pi_{out}(x, y)$ *is the (classical) random variable corresponding to Bob's output in* Π *on input* (x, y)*.*

Although we call the protocol Π' private, note that we are not considering a cryptographic scenario where the players might deviate from the protocol: we are interested in studying the information of fixed protocols. In high level, in protocol Π' Alice and Bob follow Π but they use private coins to encrypt their messages. At the end of the protocol, if their coins were the same, Bob is able to output as in Π and knows nothing else than this value. However, if their coins were different, our construction prevents them from getting any information about each other's input, in which case they just restart the process until they get the same coins. This construction yields a private protocol at the expense of a very high communication cost. Our results imply that given any function f, if we take for Π the protocol where Alice just sends over x to Bob who computes and output $f(x, y)$, we obtain a protocol Π' that can perfectly compute f with CIC only the information gained from $f(x, y)$. Let $\mathrm{CIC}^{\mathrm{ML,C}}_{\mu,k}(f)$ denote the minimum CIC achieved by a k-round memoryless quantum protocol with private classical coins to compute f exactly on input distribution μ.

Corollary 3 *For every input distribution* μ*, and every positive integer* k*,* $\mathrm{CIC}^{\mathrm{ML,C}}_{\mu,k}(f) = I\left(f(X, Y) : X|Y\right)$*, where* (X, Y) *is a random variable distributed according to* μ*, and* I *denotes the (classical) conditional mutual information.*

Note that, in the case of AND, on distribution $\mathcal{U}_0$ the output of $\mathrm{AND}(x, y) = x \wedge y$ is always 0. Hence, by the above result, $\mathrm{CIC}^{\mathrm{ML,C}}_{\mathcal{U}_0,k}(\mathrm{AND}) = 0$ for every integer k. There are two sides to this result. On the one hand, adding classical coins to quantum protocols allows for perfectly private protocols. This is impossible in the classical world and shows how quantum communication can offer advantages over classical communication. On the other hand, allowing the players to use private coins without restrictions weakens the power of information complexity as a lower bound for quantum communication complexity.

In order to try and salvage the notion of information complexity as a strong lower bound while allowing the players to use private coins, we consider an intermediate model where the players are allowed to use what we call *one-shot coins.* These are private coins that can be used only once during the protocol. In the classical setting, this assumption is not restrictive and does not change the communication complexity nor the information complexity [19]. We show that this is not the case in the quantum setting: allowing general coins or one-shot coins can lead to very different information complexities. In fact, allowing one-shot coins do not necessarily decrease the information complexity in the model without coin by much.

Theorem 4 *For every* k*-round memoryless quantum protocol* Π *with one-shot coins, we can construct a* k*-round memoryless protocol* Π' *without coins, which outputs as* Π *and such that* $\mathrm{CIC}_{\mathcal{U}_0}(\Pi') = O\big(\mathrm{CIC}_{\mathcal{U}_0}(\Pi) \cdot \big(\log(k) + |\log \mathrm{CIC}_{\mathcal{U}_0}(\Pi)|\big)\big)$*.*

Informally, the proof goes as follows. The transformation from Π to Π' is informally the following: (1) quantize the coins from Π *i.e.*, put them in quantum superposition in quantum registers; (2) at each odd (or even) round, Alice (or Bob) applies the same transformation as in Π. Then, Alice (or Bob) would like to send all their quantum registers, including the coin registers, to the other player. Before doing that, Alice (or Bob) applies a compensation unitary that will limit the information cost increase that occurs because of the sending of all the quantum registers. This result, combined with Theorem 1, implies in particular that $\mathrm{CIC}^{\mathrm{ML,C_1}}_{\mathcal{U}_0,\varepsilon,k}(AND) = \Theta_\varepsilon\left(\frac{1}{k}\right)$, where $\mathrm{CIC}^{\mathrm{ML,C_1}}_{\mu,\epsilon,k}(f)$ denotes the minimum CIC achieved by a k-round memoryless quantum protocol with private one-shot coins that computes f with error ϵ on input distribution μ. We see that while private coins allow for private protocols, one-shot coins not always do. The main open question is whether one-shot coins can be useful to reduce DISJ to AND or more generally prove some direct sum property for quantum information complexity.

References

[1] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis. Exponential separation of quantum and classical one-way communication complexity. *SIAM J. Comput.*, 38(1):366–384, 2008.

[2] Z. Bar-Yossef, T. S. Jayram, R. Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. *J. Comput. System Sci.*, 68(4):702–732, 2004.

[3] M. Braverman. Interactive information complexity. *SIAM J. Comput.*, 44(6):1698–1739, 2015.

[4] M. Braverman, A. Garg, Y. K. Ko, J. Mao, and D. Touchette. Near-optimal bounds on bounded-round quantum communication complexity of disjointness. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science—FOCS 2015*, pages 773–791, 2015.

[5] M. Braverman and A. Rao. Information equals amortized communication. In *52nd Annual Symposium on Foundations of Computer Science—FOCS 2011*, pages 748–757, 2011.

[6] H. Buhrman, R. Cleve, J. Watrous, and R. De Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16):167902, 2001.

[7] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs classical communication and computation. In *Proceedings of the 30th Annual Symposium on Foundations of Computer Science*, FOCS '98, pages 63–68, 1998.

[8] H. Buhrman, Ł. Czekaj, A. Grudka, M. Horodecki, P. Horodecki, M. Markiewicz, F. Speelman, and S. Strelchuk. Quantum communication complexity advantage implies violation of a Bell inequality. *Proceedings of the National Academy of Sciences*, 113(12):3191–3196, 2016.

[9] O. Fawzi and R. Renner. Quantum conditional mutual information and approximate Markov chains. *Comm. Math. Phys.*, 340(2):575–611, 2015.

[10] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential separation for one-way quantum communication complexity, with applications to cryptography. *SIAM J. Comput.*, 38(5):1695–1708, 2008/09.

[11] J.-Y. Guan, F. Xu, H.-L. Yin, Y. Li, W.-J. Zhang, S.-J. Chen, X.-Y. Yang, L. Li, L.-X. You, T.-Y. Chen, Z. Wang, Q. Zhang, and J.-W. Pan. Observation of quantum fingerprinting beating the classical limit. *Phys. Rev. Lett.*, 116:240502, Jun 2016.

[12] R. Jain and A. Nayak. The space complexity of recognizing well-parenthesized expressions in the streaming model: the index function revisited. *IEEE Transactions on Information Theory*, 66(10):1–23, 2014.

[13] R. Jain, J. Radhakrishnan, and P. Sen. A lower bound for the bounded round quantum communication complexity of set disjointness. In *Proceedings of the 44th IEEE symposium on Foundations of Computer Science*, STOC '03, pages 220–229, 2003.

[14] I. Kerenidis, S. Laplante, V. Lerays, J. Roland, and D. Xiao. Lower bounds on information complexity via zero-communication protocols and applications. In *Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, FOCS '12, pages 500–509, 2012.

[15] I. Kerenidis, M. Laurière, F. Le Gall, and M. Rennela. Information cost of quantum communication protocols. *Quantum Inf. Comput.*, 16(3-4):181–196, 2016.

[16] B. Klartag and O. Regev. Quantum one-way communication can be exponentially stronger than classical communication. In *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*, STOC '11, pages 31–40, 2011.

[17] E. Kushilevitz and N. Nisan. *Communication complexity*. Cambridge University Press, Cambridge, 1997.

[18] S. Laplante, M. Laurière, A. Nolin, J. Roland, and G. Senno. Robust bell inequalities from communication complexity. In *11th Conference on the Theory of Quantum Computation, Communication and Cryptography*, 2016.

[19] M. Laurière and D. Touchette. The flow of information in interactive quantum protocols: the cost of forgetting. *To appear in Innovations in Theoretical Computer Science (ITCS)*, 2017.

[20] R. Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 358–367. ACM, 1999.

[21] D. Touchette. Quantum information complexity. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, pages 317–326. ACM, 2015.

[22] F. Xu, J. Arrazola, K. Wei, W. Wang, P. Palacios-Avila, C. Feng, S. Sajeed, N. Lutkenhaus, and H.-K. Lo. Experimental quantum fingerprinting with weak coherent pulses. *Nature communications*, 6, October 2015.

Compression for Quantum Population Coding

Yuxiang Yang[1] * Ge Bai[1] † Giulio Chiribella[2] ‡ Masahito Hayashi[3] §

[1] *Department of Computer Science, The University of Hong Kong*
[2] *Department of Computer Science, The University of Hong Kong,*
Canadian Institute for Advanced Research, CIFAR Program in Quantum Information Science
[3] *Graduate School of Mathematics, Nagoya University,*
CQT, National University of Singapore

Abstract. We study the compression of arbitrary parametric families of n identically prepared finite-dimensional quantum states, in a setting that can be regarded as a quantum analogue of population coding. For a family with f free parameters, we propose an asymptotically faithful protocol that requires a memory of overall size $(f/2)\log n$. Our construction uses a quantum version of local asymptotic normality and, as an intermediate step, solves the problem of the optimal compression of n identically prepared displaced thermal states. Our protocol achieves the ultimate bound predicted by quantum Shannon theory. In addition, we explore the minimum requirement for quantum memory: On the one hand, the amount of quantum memory used by our protocol can be made arbitrarily small compared to the overall memory cost; on the other hand, any protocol using only classical memory cannot be faithful.

Keywords: Population coding; Compression; Quantum system; Local asymptotic normality; Identically prepared states

1 Introduction

Many problems in quantum information theory involve a source that prepares multiple copies of the same quantum state. This is the case, for example, of quantum tomography [1], quantum cloning [2, 3], and quantum state discrimination [4]. The state prepared by the source is generally unknown to the agent who has to carry out the task. Instead, the agent knows that the state belongs to some parametric family of density matrices $\{\rho_\theta\}_{\theta\in\Theta}$, with the parameter θ varying in the set Θ. Also, it is promised that all the particles emitted by the source are independently prepared in the same quantum state ρ_θ: when the source is used n times, it generates n quantum particles in the tensor product state $\rho_\theta^{\otimes n}$.

How much information is contained in the n-particle state $\rho_\theta^{\otimes n}$? One way to address this question is to quantify the minimum amount of memory needed to store the state. It is important to stress that the problem of storing the n-copy states $\{\rho_\theta^{\otimes n}, \theta \in \Theta\}$ in a quantum memory is different from the standard problem of quantum data compression [5, 6, 7]. In our scenario, the mixed state ρ_θ is not regarded as the average state of an information source, but, instead, as a physical encoding of the parameter θ. The goal of compression is to preserve the encoding of the parameter θ, by storing the state $\rho_\theta^{\otimes n}$ into a memory and retrieving it with high fidelity for all possible values of θ. To stress the difference with standard quantum compression, we refer to our scenario as *compression for quantum population coding.* The expression "quantum population coding" refers to the encoding of the parameter θ into the many-particle state $\rho_\theta^{\otimes n}$. We choose this expression in analogy with (classical) population coding, whereby a parameter is encoded into the population of n individuals [8]. The typical example of population coding arises in computational neuroscience, where the population consists of neurons and the parameter represents an external stimulus.

The compression for quantum population coding has been first studied in the case of pure qubit states [9, 10], and then extended to mixed states and higher dimensions [11]. Later, a new protocol that reaches the ultimate information-theoretic bound was found for qubit states [12]. The classical version of the problem was addressed in [13]. However, finding the optimal protocol for arbitrary parametric families of quantum states has remained as an open problem so far.

In this paper, we provide the general theory for the compression of n-tensor product state in a quantum parametric state family. We consider two categories of state families: families of finite-dimensional states and families of displaced thermal states in infinite dimensions. These two categories of state families turn out to be connected by the quantum version of local asymptotic normality (Q-LAN)[14, 15, 16, 17], which reduces n-tensor product of a finite-dimensional state locally to a displaced thermal state. As the first step, we discuss this kind of compression for the thermal states family, which can be regarded as the quantum extension of the Gaussian distribution. In the next step, employing Q-LAN, we reduce the problem of compressing generic finite-dimensional states to the case of displaced thermal states. Unlike previous works, our protocol does not require any assumption on the symmetry of the state family. In addition, an intriguing feature of our compression protocol is that the ratio between the size of quantum memory and the size of classical memory can be made arbitrarily close to but not equal to zero. This feature is not an accident: for identically prepared displaced thermal states and qudit states, we show that any compression protocol using only classical memory must have non-vanishing error.

*yangyx09@gmail.com
†bg95@163.com
‡giulio@hku.hk
§masahito@math.nagoya-u.ac.jp

The extended version of this paper can be found on arXiv [18].

2 Main result

The main result of our work is the optimal compression of identically prepared quantum states. We consider two major categories of states: finite dimensional (i.e. qudit) states and displaced thermal states. The key problem is to find the minimum amount of memory needed to encode these states, in a way that they can be recovered with an error vanishing in the number of input copies. A compression protocol for a parametric (sub)family $\{\rho_\theta\}_{\theta\in\mathbf{\Theta}}$ consists of two components: the encoder and the decoder, characterized by a couple of quantum channels (completely positive trace-preserving linear maps) $\mathscr{E}$ and $\mathscr{D}$ respectively.

The memory cost essentially depends on the (sub)family from which the states are drawn. For instance, the memory cost for states diagonalized in the same basis (i.e. classical probability distributions) should be less than the cost for general qudit states. As a consequence, we need to specify the state subfamily being considered before stating the main result.

We begin by introducing the parameterization for qudit and displaced thermal states. The parameters are categorized into two classes: *classical parameters* and *quantum parameters*. Roughly speaking, a classical parameter controls the eigenvalues of the density matrix, while a quantum parameter determines the eigenstates.

Any non-degenerate qudit state can be generated by rotating a fixed diagonal state $\rho_0(\mu)$ with spectrum μ, i.e. $\rho_\theta = U_\xi \rho_0(\mu) U_\xi^\dagger$, where $\theta = (\mu, \xi) \in \mathbb{R}^{d^2-1}$, with $\mu \in \mathbb{R}^{d-1}$ being the spectrum and $\xi \in \mathbb{R}^{d(d-1)}$ characterizing the rotation. The explicit form can be found in the extended version of this paper [18]. Here, components of μ are counted as classical parameters, and components of ξ are the quantum parameters.

Displaced thermal states, which are a type of infinite-dimensional states frequently encountered in quantum optics, is defined as follows:

$$\rho_{\alpha,\beta} = D_\alpha \, \rho_\beta^{(\mathrm{thm})} D_\alpha^\dagger \qquad \alpha = |\alpha| e^{iT} \quad T \in [0, 2\pi), \quad (1)$$

where $D_\mu = \exp(\mu \hat{a}^\dagger - \bar{\mu}\hat{a})$ is the displacement operator, $\beta \in [0,1)$ is a suitable parameter and

$$\rho_\beta^{(\mathrm{thm})} = (1-\beta) \sum_{i=0}^{\infty} \beta^i |i\rangle\langle i| \quad (2)$$

is a thermal state with $\{|k\rangle\}$ being the photon number basis. For the displaced thermal state family, there is one classical parameter β specifying the probability distribution of the eigenvalues, and two quantum parameters $|\alpha|$ and $T = \arg\alpha$ describing the strength and phase of the displacement.

The main result of our work is the following:

Theorem 1 *Let $\{\rho_\theta^{\otimes n}\}_{\theta\in\mathbf{\Theta}}$ be the state (sub)family of n identical displaced thermal states or non-degenerate qudit states with f_c free classical parameters and f_q free quantum parameters. For any $\delta \in (0, 2/9)$, the state family can be compressed into $[(1/2+\delta)f_c + (1/2)f_q]\log n$ classical bits and $(f_q\delta)\log n$ qubits with an error $\epsilon = O\left(n^{-\delta/2}\right) + O\left(n^{-\kappa(\delta)}\right)$, where $\kappa(\delta) > 0$ is determined by the error of Q-LAN [17]. The compression is optimal, in the sense that any compression protocol requiring a memory of size $[(f_c + f_q)/2 - \delta']\log n$ with $\delta' > 0$ cannot be faithful.*

For the state (sub)family, a free parameter is assumed to be variable in a certain interval, while a fixed parameter can only take one fixed value. By faithful we mean that the worst-case trace distance between the original state and the recovered state $\mathscr{D} \circ \mathscr{E}(\rho_\theta^{\otimes n})$

$$\epsilon := \sup_{\theta\in\mathbf{\Theta}} \frac{1}{2} \|\rho_\theta^{\otimes n} - \mathscr{D} \circ \mathscr{E}(\rho_\theta^{\otimes n})\|_1 \quad (3)$$

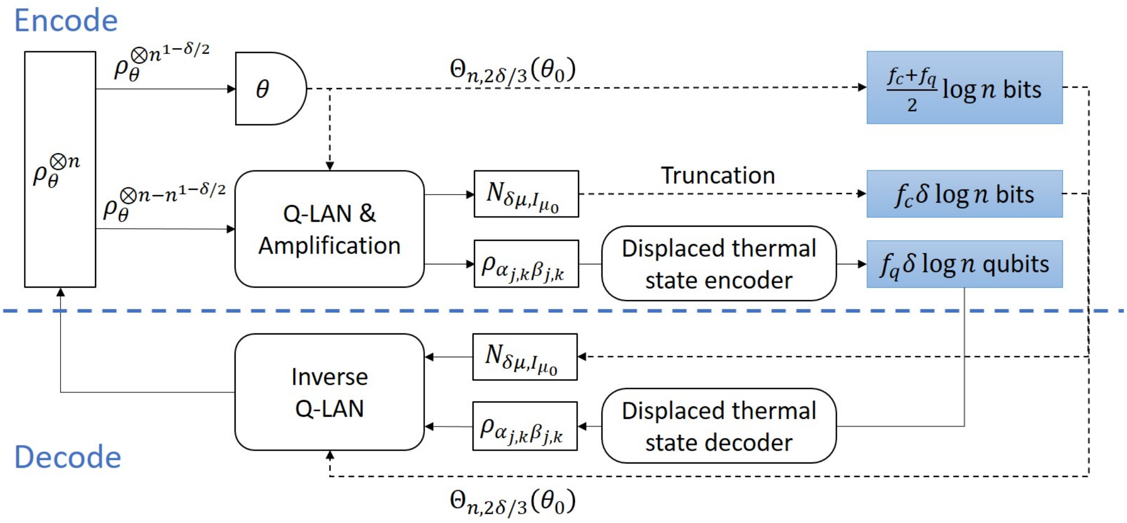
vanishes in the $n \to \infty$ limit.

Theorem 1 states that to encode each free parameter a memory of size $(1/2+\delta)\log n$ is required. When the parameter is classical, the required memory is fully classical; when the parameter is quantum, a quantum memory of $\delta \log n$ qubits is required. Note that Theorem 1 solves the compression of several important (sub)families:

- The full model family of qudits ($f_c = d-1$ and $f_q = d(d-1)$).
- The classical qudit subfamily ($f_c = d-1$ and $f_q = 0$) of d-dimensional classical probability distributions can be compressed into $(d/2)\log n$ classical bits, retrieving the result of [13].
- The phase-covariant qudit subfamily ($f_c = d-1$ and $f_q = d(d-1)/2$).

As we can see from Theorem 1, if we take δ to be small enough, the ratio between the quantum memory cost and the classical memory cost $\delta f_q/((1/2+\delta)f_c + (1/2)f_q)$ can be made close to zero when δ is set close to zero, yet the compression error vanishes more slowly. It is then intuitive to ask whether this ratio can be made equal to zero while keeping the error vanishing, i.e. compressing faithfully using a fully classical memory. The answer to the above question is negative: we proved in the extended version [18] that only commuting quantum states can be faithfully compressed into classical bits. In other words, only classical families of states can be stored using into purely classical bits.

In [18] we provide protocols for compressing a variety of displaced thermal state (sub)families, and here we briefly introduce the protocol for compressing non-degenerate qudit states. Displaced thermal states and non-degenerate qudit states are closely connected by the Q-LAN. Here we use the results derived in [17], which states that n identical copies of a qudit state can be approximated by a classical-quantum Gaussian state in a sufficiently small neighborhood of a point $\theta_0 \in \mathbf{\Theta}$ for large n. The approximation is physically implemented by two quantum channels: $\mathscr{T}_{\theta_0}^{(n)}$ that converts qudit states



Encode
$\rho_\theta^{\otimes n^{1-\delta/2}}$
θ
$\Theta_{n,2\delta/3}(\theta_0)$
$\frac{f_c+f_q}{2}\log n$ bits
$\rho_\theta^{\otimes n}$
$\rho_\theta^{\otimes n-n^{1-\delta/2}}$
Q-LAN & Amplification
$N_{\delta\mu,I_{\mu_0}}$
Truncation
$f_c\delta\log n$ bits
$\rho_{\alpha_{j,k}\beta_{j,k}}$
Displaced thermal state encoder
$f_q\delta\log n$ qubits
Inverse Q-LAN
$N_{\delta\mu,I_{\mu_0}}$
$\rho_{\alpha_{j,k}\beta_{j,k}}$
Displaced thermal state decoder
Decode
$\Theta_{n,2\delta/3}(\theta_0)$

Figure 1: **Compression protocol for finite dimensional states.**

to Gaussian states, and $\mathscr{S}_{\theta_0}^{(n)}$ that do the inverse conversion.

By using Q-LAN, we can reduce the compression for non-degenerate qudit families to the compression of displaced thermal state families as shown in Figure 1.

First, $n^{1-\delta/2}$ copies of ρ_θ in the input are taken out for tomography. In this way, one obtains a neighborhood that contains the input state with high probability. This neighborhood is encoded into classical memory, and is also used for the constructing the Q-LAN channel $\mathscr{T}_{\theta_0}^{(n-n^{1-\delta/2})}$ that converts the remaining $n-n^{1-\delta/2}$ copies to a classical-quantum Gaussian state. Next, the Gaussian state is amplified to compensate the loss of input copies. For each individual quantum mode, a displaced thermal state compressor $\mathscr{P}_{j,k}$ [18] compresses it into quantum memory.

The state can be decompressed from the memory by sending the state of the hybrid memory through the channel $\mathscr{S}_{\theta_0}^{(n)}$, which can be constructed by consulting the outcome of tomography.

The optimality of our compression protocol is justified as follows. To prove that any protocol with an overall memory size of $(f/2-\delta)\log n$ (where $f = f_c + f_q$) for $\delta > 0$ cannot be faithful, it suffices to construct such a family that requires more than this amount of memory to be faithfully compressed.

We define a mesh $\mathbf{M}$ on the parameter space $\mathbf{\Theta}$:

$$\mathbf{M} = \left\{\theta \in \mathbf{\Theta} \mid |(\theta-\theta_0)_i| = z_i \cdot \log n/\sqrt{n}, z_i \in \mathbb{N} \ \forall\ i\right\} \tag{4}$$

where $\theta_0 \in \mathbf{\Theta}$ is a fixed point. The mesh $\mathbf{M}$ is so defined that the states corresponding to this mesh $\{\rho_\theta\}_{\theta\in\mathbf{M}}$ are almost mutually distinguishable. The points can be used to approximately encode $|\mathbf{M}|$ different messages, and the amount of information contained in this ensemble is approximately $\log|\mathbf{M}|$. In [18], we showed that the size of memory required to faithfully encode this ensemble is

$$n_{\text{enc}} \geq \frac{f}{2}\log n - f\log\log n + o(1) \tag{5}$$

which, for any $\delta > 0$, is larger than $(f/2-\delta)\log n$ for large n, showing the optimality of our protocol.

3 Conclusion

In this work we have solved the problem of compressing identically prepared states of finite-dimensional quantum systems and displaced thermal states. We showed that the total size of the required memory is proportional to the number of free parameters of the state. Moreover, we observed the asymptotic ratio between the amount of quantum bits and the amount of classical bits can be set to an arbitrarily small constant. Still, a fully classical memory cannot faithfully encode genuine quantum states: only states that are jointly diagonal in fixed basis (i.e. a classical state family) can be compressed into a purely classical memory.

Acknowledgments

This work is supported by the Canadian Institute for Advanced Research (CIFAR), by the Hong Kong Research Grant Council through Grant No. 17326616, by National Science Foundation of China through Grant No. 11675136, and by the HKU Seed Funding for Basic Research, and the John Templeton Foundation. Y. Y. is supported by a Hong Kong and China Gas Scholarship. MH was supported in part by a MEXT Grant-in-Aid for Scientific Research (C) No. 16K00014 and (B) No. 16KT0017, the Okawa Research Grant and Kayamori Foundation of Informational Science Advancement.

References

[1] Konrad Banaszek, Marcus Cramer, and David Gross. Focus on quantum tomography. *New Journal of Physics*, 15(12):125020, 2013.

[2] Valerio Scarani, Sofyan Iblisdir, Nicolas Gisin, and Antonio Acín. Quantum cloning. *Reviews of Modern Physics*, 77:1225–1256, Nov 2005.

[3] Nicolas J Cerf, A Ipe, and Xavier Rottenberg. Cloning of continuous quantum variables. *Physical Review Letters*, 85(8):1754, 2000.

[4] Stephen M Barnett and Sarah Croke. Quantum state discrimination. *Advances in Optics and Photonics*, 1(2):238–278, 2009.

[5] Benjamin Schumacher. Quantum coding. *Physical Review A*, 51(4):2738, 1995.

[6] Richard Jozsa and Benjamin Schumacher. A new proof of the quantum noiseless coding theorem. *Journal of Modern Optics*, 41(12):2343–2349, 1994.

[7] Hoi-Kwong Lo. Quantum coding theorem for mixed states. *Optics Communications*, 119(5-6):552–556, 1995.

[8] Si Wu, Shun-ichi Amari, and Hiroyuki Nakahara. Population coding and decoding in a neural field: a computational study. *Neural Computation*, 14(5):999–1026, 2002.

[9] Martin Plesch and Vladimír Bužek. Efficient compression of quantum information. *Physical Review A*, 81(3):032317, 2010.

[10] Lee A. Rozema, Dylan H. Mahler, Alex Hayat, Peter S. Turner, and Aephraim M. Steinberg. Quantum data compression of a qubit ensemble. *Physical Review Letters*, 113:160504, Oct 2014.

[11] Yuxiang Yang, Giulio Chiribella, and Daniel Ebler. Efficient quantum compression for ensembles of identically prepared mixed states. *Physical Review Letters*, 116:080501, Feb 2016.

[12] Yuxiang Yang, Giulio Chiribella, and Masahito Hayashi. Optimal compression for identically prepared qubit states. *Physical Review Letters*, 117:090502, Aug 2016.

[13] Masahito Hayashi and Vincent Tan. Minimum rates of approximate sufficient statistics. *arXiv preprint arXiv: 1612.02542*, 2016.

[14] Masahito Hayashi and Keiji Matsumoto. Asymptotic performance of optimal state estimation in qubit system. *Journal of Mathematical Physics*, 49(10):102101, 2008.

[15] Mădălin Guţă and Jonas Kahn. Local asymptotic normality for qubit states. *Physical Review A*, 73(5):052108, 2006.

[16] Mădălin Guţă and Anna Jenčová. Local asymptotic normality in quantum statistics. *Communications in Mathematical Physics*, 276(2):341–379, 2007.

[17] Jonas Kahn and Mădălin Guţă. Local asymptotic normality for finite dimensional quantum systems. *Communications in Mathematical Physics*, 289(2):597–652, 2009.

[18] Yuxiang Yang, Ge Bai, Giulio Chiribella, and Masahito Hayashi. Data compression for quantum population coding. *arXiv preprint arXiv:1701.03372*, 2017.

Detecting metrologically useful asymmetry and entanglement by few local measurements

Chao Zhang[1 2] Benjamin Yadin[3] Zhi-Bo Hou[1 2] Huan Cao[1 2] Bi-Heng Liu[1 2]
Yun-Feng Huang[1 2 *] Reevu Maity[3] Vlatko Vedral[3 4] Chuan-Feng Li[1 2 †]
Guang-Can Guo[1 2] Davide Girolami[3 ‡]

[1] *Key Laboratory of Quantum Information, University of Science and Technology of China, CAS, Hefei, 230026, China*
[2] *Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, 230026, P.R. China*
[3] *Department of Atomic and Laser Physics, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom*
[4] *Centre for Quantum Technologies, National University of Singapore, 117543 Singapore*

Abstract. Important properties of a quantum system are not directly measurable, but they can be disclosed by how fast the system changes under controlled perturbations. In particular, asymmetry and entanglement can be verified by reconstructing the state of a quantum system. Yet, this usually requires experimental and computational resources which increase exponentially with the system size. Here we show how to detect metrologically useful asymmetry and entanglement by a limited number of measurements. This is achieved by studying how they affect the speed of evolution of a system under a unitary transformation. We show that the speed of multi-qubit systems can be evaluated by measuring a set of local observables increasing linearly with the number of qubits. We implement the detection scheme in an all-optical experiment.

Keywords: quantum metrology, speed of evolution, asymmetry and entanglement,quantum fisher information

The ability to engineer quantum coherence and entanglement is one of the main factors determining nonclassical speed-up in information processing. Yet, their experimental verification is a serious challenge. As they are not directly observable, their detection usually implies reconstructing the full state of the system, which requires a number of measurements growing exponentially with the system size. Also, verifying their presence is necessary, but not always sufficient to guarantee a computational advantage.

Studying the rate of change of a system under carefully designed perturbations is a clever way to investigate its key properties. In quantum information and metrology protocols, a system speed determines its computational power. In open system, computing quantum speed limits also provides information about the environment structure, helping develop efficient control strategies, and investigate phase transitions of condensed matter systems. Here we show how to detect metrologically useful coherence and entanglement in systems of arbitrary dimension by measuring the speed of evolution under a generic quantum channel, which for n-qubit systems is a function of a linearly scaling ($O(n)$) number of observables. The system speed is defined by the average rate of change of the state,which is given by mean values of quantum operators $\langle\cdot\rangle_{\rho_t} = \mathrm{Tr}(\cdot\rho_t)$:

$$s_\tau(\rho_t) := \frac{||\rho_\tau - \rho_0||_2}{\tau} = \frac{(\langle\rho_\tau\rangle_{\rho_\tau} + \langle\rho_0\rangle_{\rho_0} - 2\langle\rho_\tau\rangle_{\rho_0})^{1/2}}{\tau},$$

where the Euclidean distance is employed. Measuring the swap operator on two system copies is sufficient to quantify state overlaps. The overlap detection network for n-qubit systems is depicted in Fig. 1. The unitary gates $U_{\tau,i} = e^{-ih_i\tau}$ are applied to the second copy of each pair. The overlap, and therefore the speed function can be extracted by the Bell state measurement on each pair of subsystem copies. We prove a quantitative link between our speed measure, when undertaking a unitary dynamics, and metrological quantum resources. First, we relate speed to asymmetry, i.e. the amount of coherence with respect to an Hamiltonian eigenbasis, which underpins the usefulness of a probe state to phase estimation and reference frame alignment schemes. By extending the analysis to multipartite systems, a superlinear increase

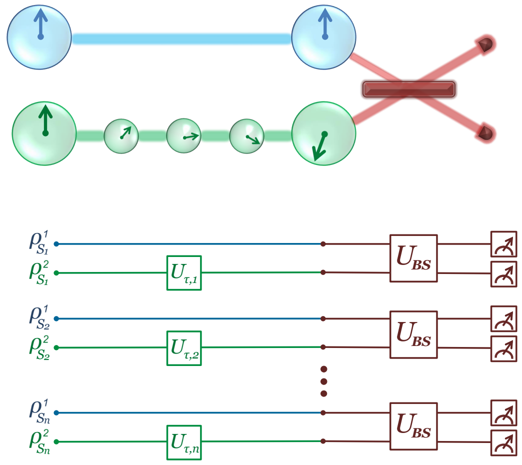


Figure 1: Overlap detection network.

*hyf@ustc.edu.cn
†cfli@ustc.edu.cn
‡davegirolami@gmail.com

of speed with the system size certifies an advantage in phase estimation powered by entanglement.

We demonstrate the scheme in an all-optical experiment. We extract a lower bound to the metrologically useful coherence (i.e. asymmetry) and entanglement of a two-qubit system, by measuring its speed in a controlled unitary evolution. While state tomography would require fifteen measurements, we verify that the proposed protocol needs six. The system is prepared in two-copies of a mixture of Bell states via spontaneous parametric down-conversion (SPDC) sources. We evaluate the speed function from purity and overlap measurements. We implement a six-photon architecture to rule out the case of BSMs measuring two photon pairs emitted by a single SPDC source. We obtain results of excellent quality, being able to experimentally quantify for the first time the asymmetry of a system without state reconstruction. Also, non-classical metrological efficiency due to entanglement is reliably detected.

To summarize, our work answers the crucial question: how can we tell if a genuinely quantum process occurred? We show that speed detection is an extremely powerful strategy to certify quantum processes yielded by coherence and entanglement in large computational registers.

Past of a quantum particle: Common sense prevails

Berthold-Georg Englert,[1, 2, 3, *] Kelvin Horia,[4, †] Jibo Dai,[1, ‡] Yink Loong Len,[1, §] and Hui Khoon Ng[5, 1, 3, ¶]

[1] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543, Singapore*
[2] *Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542, Singapore*
[3] *MajuLab, CNRS-UNS-NUS-NTU International Joint Unit, UMI 3654, Singapore*
[4] *Division of Physics and Applied Physics, School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371, Singapore*
[5] *Yale-NUS College, 16 College Avenue West, Singapore 138527, Singapore*

Abstract. We analyze Vaidman's three-path interferometer with weak path marking [Phys. Rev. A **87**, 052104 (2013)] and find that common sense yields correct statements about the particle's path through the interferometer. This disagrees with the original claim that the particles have discontinuous trajectories at odds with common sense. In our analysis, "the particle's path" has operational meaning as acquired by a path-discriminating measurement. For a quantum-mechanical experimental demonstration of the case, one should perform a single-photon version of the experiment by Danan et al. [Phys. Rev. Lett. **111**, 240402 (2013)] with unambiguous path discrimination. We present a detailed proposal for such an experiment.

Keywords: Foundations of quantum mechanics, Quantum state engineering and measurements, Optical tests of quantum theory

1 Introduction

Vaidman argues that one can meaningfully talk about the past of a quantum particle — specifically: which path it took through an interferometer — by analyzing the faint trace left along the path by weak, almost non-disturbing, measurements within a formalism that uses forward and backward evolving quantum states [1, 2]. This leads to the following criterion [3]:

> The particle was present in paths of the interferometer in which there is an overlap of the forward and backward evolving wave functions. (1)

He then arrives at conclusion that contradicts common sense: The particle can have trajectories that are not continuous, for example, in a three-path interferometer. Later, this assertion is seemingly confirmed by experiments on an optical three-path interferometer [4, 5]. In both experiments, the three-path interferometer is an asymmetric Mach-Zehnder interferometer (MZI) with a symmetric MZI inserted into one arm; see Fig.(1) for illustration.

Various aspects of this matter have been debated; there are at least thirty papers written on the subject in the past few years [6]. The debate is still on-going.

In this work, we first discuss an important aspect that has not yet been recognized and examined in depth: How one extracts path information from faint traces left by an individual particle on its way through the interferometer. It turns out that *common sense prevails if the right question about the path knowledge is asked.* (See points 1 to 3 in the detailed summary below).

Second, we point out that the experimental results in Refs.[4, 5], in fact, do not provide direct and unanimous support to Vaidman's narrative (1) for the past of the photons. (See points 3 to 5 in the detailed summary below.)

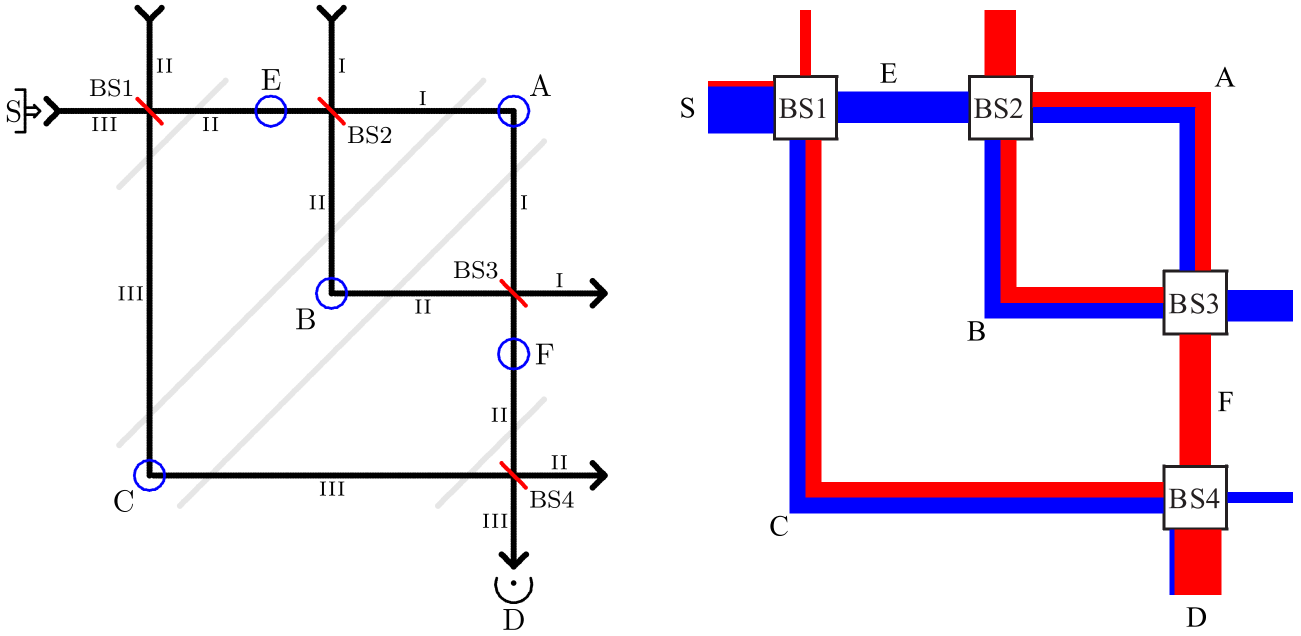
Lastly, we propose *single-photon experiments* for both the two-path interferometer and the three-path interferometer of Ref. [4], with *unambiguous and full-path information extraction.* The proposal for the three-path interferometer has not yet been realized, but once performed, it should demonstrate that common sense does prevail. (See points 6 to 7 in the detailed summary below.)

2 Detailed Summary

1. We review Vaidman's three-path interferometer, and analyze the weak path marking by which a particle leaves faint traces at the various checkpoints on its way from the source to the detector. We conclude that destructive interference suppresses the traces at two checkpoints, which explains why a particle has a discontinuous trajectory in Vaidman's narrative.

2. We further analyze how one acquires such specific knowledge about the path of a *particle just detected* by a suitable measurement of the quantum degrees of freedom that are used to mark the path. In this context, what we learn depends a lot on the question we ask by the chosen measurement, and not all questions are equally relevant. We examine the faint traces left by the particle just detected and show how one extracts path knowledge by a *measurement of unambiguous discrimination* (MUD)

[*] cqtebg@nus.edu.sg
[†] khoria@ntu.edu.sg
[‡] Currently at Data Storage Institute, A*Star; dai_jibo@dsi.a-star.edu.sg
[§] yinkloong@quantumlah.org
[¶] cqtnhk@nus.edu.sg



(a)

(b)

FIG. 1: Fig.(1a) shows Vaidman's three-path interferometer of Refs. [1-5]. The quantum particle is emitted by source S, enters the interferometer at beam splitter BS1, and is detected by detector D after exiting at beam splitter BS4. Both BS1 and BS4 have 2:1 transmission/reflection ratios. On the way from S to D, the particle can take the outer path (path III), identified by path marker at checkpoint C, or the paths through the internal MZI (paths I or II), identified by path markers at checkpoint A or B. The inner MZI is balanced such that for particles entering from checkpoint E, checkpoint F would be the dark port due to destructive interference; BS2 and BS3 are 50:50 nonpolarizing beam splitters. The faint slanted lines connect simultaneous points on the three paths. Fig.(1b) illustrates the description of the past of the particles using two-vector formalism and criterion (1). Since inside the three-path interferometer there are only overlapping of forward wave function (blue/darker lines) and backward (red/lighter lines) wave function at the inner MZI and the outer path, but not at E or F, it is concluded that the trajectory from S to D is discontinuous. Note that the thickness of the blue lines is proportional to the probability of finding the quantum particle there if we look for it.

on the path-marker degree of freedom. Unambiguous which-path information gives operational and quantitative meaning to the otherwise vague concept of "where was the quantum particle." We show that, in fact, for the case of Vaidman's three-path setup with weak path marking, only a small fraction of the particles have path knowledge, while all others have unknowable paths.

3. Upon noting that the probability amplitudes processed by the final beam splitter (BS4) are incoherent, we show how the observed data would in fact *agree with one's common sense*. Specifically, for *post-selected particles* that are detected by D after BS4 (as is done and analyzed in Refs. [4, 5]), we conclude that the particles with unknown path have, in fact, followed the common-sense path, i.e. the outer path of the three-path interferometer. In the limit of ever fainter traces, these are all particles. We provide two analyses, one is through an accounting exercise, and another by examination of the subensembles sorted according to the unambiguous discrimination results, to support our conclusion.

4. In Ref. [4], the experiment is performed with classical light intensities, and no information is available about *individual* photons. One cannot simply invoke a fair-sampling assumption to infer the past of a *single photon*. The formalism and analysis used as in Refs. [1, 2] can only be applied to an *ensemble* of particles [7].

5. In Ref. [5], while the experiment is indeed performed with single photons, their detection method is unable to extract *all the available path information in an unambiguous fashion*, and as such is little better than the original experiment in resolving the narrative.

6. We take a close look at the inner two-path interferometer in Vaidman's three-path setup, and show that every particle detected at the exit for the destructive interference, i.e. at F, has a known path, which is the reason for the incoherent probability amplitudes mentioned in point 3 above. We propose and perform a single-photon experiment for such two-path interferometer with weak path marking. In this experiment, we produce pairs of polarization-entangled photons from down conversion. One of them (signal) enters the MZI, with its

polarization altered gently according to the path taken. Due to the entanglement, the polarization degree of freedom of the partner photon (idler) is then utilized as the path marker. Upon performing MUD for two polarization qubit states, we can then extract the path information of the signal. All the obtained data are in good agreement with our theoretical predictions [8].

7. Furthermore, we propose a single-photon version of the three-path interferometer experiment of Ref. [4], with unambiguous and full-path information extraction. The experiment is similar to the one for the two-path interferometer above, but now *two pairs* of down-converted photons are needed, as markings for three paths are required. In addition, the MUD is performed on the path qutrits of the idler rather than polarization qubits.

Our treatment is entirely within the standard formalism of quantum mechanics and does not rely on the two-state formalism employed by Vaidman. While we do not seek to question the validity of the two-state formalism, we see no particular advantage in using it; the standard formalism offers a transparent way for studying the properties of ensembles that are both pre-selected and post-selected. Our analysis of Vaidman's three-path interferometer with weak path marking has established that common sense does not mislead us, and Vaidman's criterion (1) does not correctly identify the path taken by the particle.

A preprint for this work can be found at arXiv:1704.03722.

[1] L. Vaidman, Phys. Rev. A **87**, 052104 (2013).

[2] L. Vaidman, Phys. Rev. A **89**, 024102 (2014).

[3] L. Vaidman, e-print arXiv:1610.04781 [quant-ph] (2016).

[4] A. Danan, D. Farfurnik, S. Bar-Ad, and L. Vaidman, Phys. Rev. Lett. **111**, 240402 (2013).

[5] Z.-Q. Zhou, X. Liu, Y. Kedem, J.-M. Cui, Z.-F. Li, Y.-L. Hua, C.-F. Li, and G.-C. Guo, Phys. Rev. A **95**, 042121 (2017).

[6] See the Refs.[4-33] in arXiv:1704.03722.

[7] Actually, the power spectrum reported in Ref. [4], the squared Fourier transform of an intensity difference, is quadratic in the light intensity and, therefore, outside the scope of linear optics. Hence, one-to-one correspondence between light intensities and photon probabilities cannot be used here.

[8] Y. L. Len, J. Dai, B.-G. Englert, and L. Krivitsky, in preparation.

Occam's Vorpal Quantum Razor: Memory reduction when simulating continuous-time stochastic processes with quantum devices

Thomas J. Elliott[1] * Mile Gu[1 2 3] †

[1] *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 639673*
[2] *Complexity Institute, Nanyang Technological University, Singapore 639673*
[3] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543*

Abstract. The ubiquity of continuous-time stochastic processes makes the simulation these processes a great utility. Classical simulators of continuous-time processes must typically track unbounded amounts of information about past behaviour, enforcing limits on precision due to finite machine memory. However, quantum machines can require less past information than even their optimal classical counterparts to simulate discrete-time processes, and we demonstrate that this advantage extends to the continuous-time regime. Moreover, we show that this memory reduction can be unboundedly large, allowing for arbitrary precision with a finite quantum memory. We provide a systematic construction for superior quantum machines, and a protocol for analogue quantum simulation of continuous-time renewal processes.

Keywords: Quantum Simulation, Quantum Memory, Computational Mechanics, Quantum Information, Continuous-Time Stochastic Processes

Our experience of the world manifests as a series of observations. These may be characterised by a time series that details what is observed, and when. The goal of scientific theories is to construct models which can provide consistent explanations for past observations, and make predictions about the future. Once equipped with a model, we can build simulators of the processes, to emulate the behaviour of the modelled system.

In general, one can devise many models for a process that make identical predictions, and so it is desirable to have criteria for discerning a preferred model. A guiding philosophy for this is Occam's razor "plurality should not be posited without necessity", which can be interpreted as requiring that a model should be the 'simplest' that accurately describes our observations. The field of computational mechanics [1] provides a quantitive notion of what constitutes the simplest model, defining the optimal model of a process to be that which requires the least information about the past behaviour of the system whilst retaining the same predictive power as though the entire past history was known. Beyond the desire for elegance, a pragmatic reason for investigating such optimal models is that it facilitates the construction of simulators that make efficient use of resources. Here, the resource we optimise is the internal memory of the simulator.

It has recently been shown [2] that quantum mechanics allows for the construction of simulators of discrete-time processes that require less information about the past than their optimal classical counterparts. This is possible because past states that have some, but not complete overlap in future statistics need not be perfectly distinguished, and may be stored as non-orthogonal memory states. This leads to the perhaps surprising conclusion that a quantum device can be more efficient than a classical system even when simulating a purely classical process. This opens a novel avenue for the quantum simulation of classical stochastic processes.

In our work [3] we have demonstrated that this quantum advantage may be extended to the much richer and broader realm of continuous-time stochastic processes. Focussing on the particular case of renewal processes, (which can be used to model a diverse range of systems including queues and neural spike trains), we provide a systematic construction for quantum models in this regime that are more efficient than their corresponding optimal classical counterparts, that can be applied for arbitrary waiting-time distributions.

Classical simulators of such processes [4] typically require an unboundedly large memory to operate exactly in the continuous-time regime, hence enforcing a trade-off between memory usage and precision, ultimately and fundamentally limiting the accuracy of such simulations. In contrast to this, we show that our construction for quantum simulators of the same processes may achieve arbitrarily fine precision whilst still requiring only finite memory [Fig. 1]. This hence allows quantum devices to sidestep the limitations suffered by their classical complements.

Alongside the details for constructing quantum simulators of continuous-time renewal processes superior to the optimal classical models, we characterise the memory required by such quantum models, which we show to be timescale invariant, depending only on the form of the waiting-time distribution. Further, we show that unlike the optimal classical models that require certain properties of the waiting-time distribution to be known and incorporated into their construction, the superior quantum models in effect self-assemble with a naive construction directly from the waiting-time distribution. We outline a protocol for how quantum models can be implemented as analogue simulators of renewal processes, and illustrate our results with two examples that exhibit the unbounded advantage in the memory requirement of quantum devices. We conclude by arguing that this infinite memory saving may be a general property of quan-

*physics@tjelliott.net
†milegu@quantumcomplexity.org

tum simulators of continuous-time renewal processes, and discuss the prospects for extending our results to more general continuous-time stochastic processes.

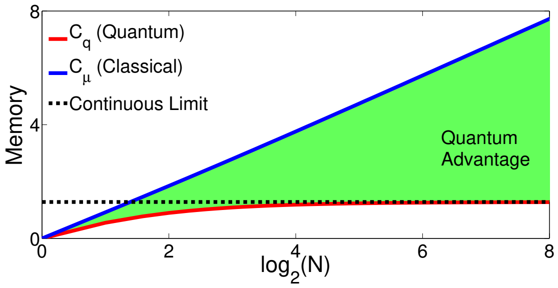


Figure 1: The classical memory requirement to simulate continuous-time stochastic processes typically increases unboundedly with finer coarse-graining, while a quantum simulator may be able to perform the same task to arbitrary precision with only finite memory. The example shown is for a uniform emission density renewal process, where we find a bounded quantum memory requirement in the continuous limit.

References

[1] J. P. Crutchfield and K. Young, Physical Review Letters **63**, 105 (1989).

[2] M. Gu, K. Wiesner, E. Rieper, and V. Vedral, Nature Communications **3**, 762 (2012).

[3] T. J. Elliott and M. Gu, arXiv preprint arXiv:1704.04231 (2017).

[4] S. E. Marzen and J. P. Crutchfield, Entropy **17** 4891 (2015)

Semidefinite programming converse bounds for quantum communication

Xin Wang[1] [*] Kun Fang[1] [†] Runyao Duan[1] [2] [‡] [§]

[1] *Centre for Quantum Software and Information,*
Faculty of Engineering and Information Technology,
University of Technology Sydney, NSW 2007, Australia
[2] *UTS-AMSS Joint Research Laboratory for Quantum Computation and Quantum Information Processing,*
Chinese Academy of Sciences, Beijing 100190, China

Abstract. Recently the power of positive partial transpose preserving (PPT) and no-signalling (NS) codes in quantum communication has been studied in [Leung/Matthews, *IEEE Trans. Inf. Theory* **61**:4486, 2015]. We continue with this line and study the PPT-assisted quantum communication in both non-asymptotic and asymptotic settings. We show improved semidefinite programming (SDP) finite blocklength converse bounds for quantum communication with a given infidelity tolerance and utilize them to study the depolarizing channel and amplitude damping channel in the small blocklength. We then present a general SDP strong converse bound on quantum capacity. All these bounds are efficiently computable and do not rely on any specific structure of the channel. In particular, we prove that our SDP strong converse bound is always smaller than or equal to the *partial transposition bound* introduced by Holevo and Werner, and the inequality could be strict. Furthermore, we show that the SDP strong converse bound can be refined as the *max-Rains information*, which an analog to the Rains information introduced in [Tomamichel/Wilde/Winter, *IEEE Trans. Inf. Theory* **63**:715, 2017]. This also implies that it is always no smaller than the Rains information. Finally, we establish an inequality relationship among the known strong converse bounds on quantum capacity.

Keywords: quantum capacity, strong converse, semidefinite program

1 Introduction

A central topic in quantum information theory is the reliable transmission of quantum information via noisy quantum channels. The quantum capacity of a noisy quantum channel is the highest rate at which it can convey quantum information reliably over asymptotically many uses of the channel. The theorem by Lloyd, Shor, and Devetak (LSD) [1, 2, 3] and the work in [4, 5, 6] show that the quantum capacity is equal to the regularized coherent information. The quantum capacity is notoriously difficult to evaluate since it is characterized by a multi-letter, regularized expression and it is not even known to be computable [7]. Our understanding of the classical and quantum capacities remains limited. Even for qubit channels, the quantum capacity of depolarizing channel is unsolved. Given an arbitrary quantum channel, a previously known efficiently computable strong converse bound is the partial transposition bound introduced in Ref. [8]. Recently, the Rains information [9] was established to be a strong converse bound for quantum communication.

To better understand the channel capacities, one can study the performance of extra free resources in the coding scheme. This scheme, called a *code*, is equivalently a bipartite operation under some physically reasonable restrictions. Recently, the NS and PPT codes are introduced to study quantum capacity [10] while NS-assisted zero-error classical capacity is studied in [11]. These codes are with mathematically tractable structure and could provide insights into the tough problem of determining capacities. Note that NS codes are potentially stronger than entanglement codes and the PPT codes include operations that can be implemented by local operations and classical communication.

Another fundamental problem, of both theoretical and practical interest, is the trade-off between the channel uses, transition rate and error tolerance in the non-asymptotic (or finite blocklength) regime. In a realistic setting, the number of channel uses is necessarily limited in quantum information processing. Therefore one has to make a trade-off between the transmission rate and error tolerance. The study of finite blocklength regime has recently garnered great interest in classical information theory (e.g., [12, 13]) as well as in quantum information

[*] `xin.wang-8@student.uts.edu.au`

[†] `kun.fang-1@student.uts.edu.au`

[‡] `runyao.duan@uts.edu.au`

[§] This work is an extended version of the previous work on arXiv: 1601.06888.

theory (e.g., [14, 15, 16, 17, 18, 10, 19, 20, 21, 22]).

2 Overview of results

Our work follows the quantum communication via quantum channels assisted by NS and PPT codes [10] and focuses on both non-asymptotic and asymptotic regimes. To be specific, we show

- SDP finite blocklength converse bounds for quantum communication with a given error tolerance;
- SDP strong converse bound for quantum capacity: for any code with a rate exceeding our bound, the error probability goes to one exponentially fast in the limit of many channel uses.

Our bounds do not rely on the structure of the channel and can be efficiently computed for any channel since SDP [23] can be solved by polynomial-time algorithms [24].

We first show that the one-shot ε-error capacity assisted by PPT ∩ NS codes, is given by

$$\begin{aligned} Q^{(1)}_{PPT\cap NS}(\mathcal{N},\varepsilon) \\ &= -\log \min\ m \\ &\qquad \text{s.t.}\ \operatorname{Tr} J_{\mathcal{N}} W_{AB} \geq 1-\varepsilon, \\ &\qquad\quad 0 \leq W_{AB} \leq \rho_A \otimes \mathbb{1}_B, \\ &\qquad\quad \operatorname{Tr}\rho_A = 1, \\ &\qquad \textbf{PPT:}\ -m\rho_A \otimes \mathbb{1}_B \leq W_{AB}^{T_B} \leq m\rho_A \otimes \mathbb{1}_B \\ &\qquad \textbf{NS:}\ \operatorname{Tr}_A W_{AB} = m^2 \mathbb{1}_B. \end{aligned} \tag{1}$$

It is worth noting that Eq. (1) is not a convex optimization problem. We do some relaxation on the constraints and obtain SDP converse bounds for quantum communication. These bounds improve the SDP converse bound ($-\log f(\mathcal{N},\varepsilon)$) in Ref. [22]. To be specific, we show that for any quantum channel $\mathcal{N}$ and error tolerance ε, it holds that

$$\begin{aligned} Q^{(1)}(\mathcal{N},\varepsilon) \leq Q^{(1)}_{PPT\cap NS}(\mathcal{N},\varepsilon) \leq -\log \widehat{g}(\mathcal{N},\varepsilon) \\ \leq -\log \widetilde{g}(\mathcal{N},\varepsilon) \leq -\log g(\mathcal{N},\varepsilon) \leq -\log f(\mathcal{N},\varepsilon)\,. \end{aligned} \tag{2}$$

where $\widehat{g}(\mathcal{N},\varepsilon)$, $\widetilde{g}(\mathcal{N},\varepsilon)$, $g(\mathcal{N},\varepsilon)$ are the SDPs we derived, whose explicit formula can be found in the technical version of this work. Note that to get a better result, we may require a few times of successive refinement for $\widehat{g}$ and denote the result after i times of refinement as $\widehat{g}_i$.

Examples of the amplitude damping channel $\mathcal{N}^{(r)}_{AD}$ and the qubit depolarizing channel $\mathcal{N}_D$ (with depolarizing parameter $p = 0.3$) have been given to illustrate that our SDP converse bounds in Eq. (2) can be strictly tighter. The numerical results are shown in Figs. 1 and 2, respectively.

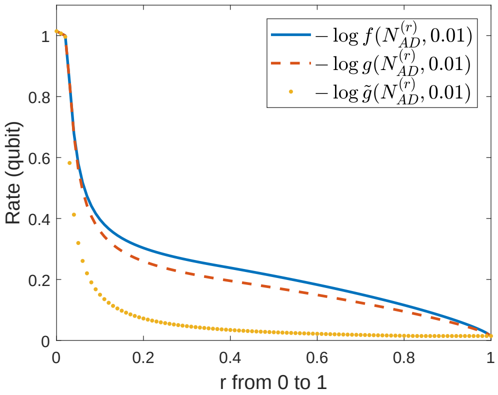


Figure 1: Comparsion of SDP converse bounds $-\log f$, $-\log g$, $-\log \widetilde{g}$ in the case of amplitude damping channels with error tolerance $\varepsilon = 0.01$.

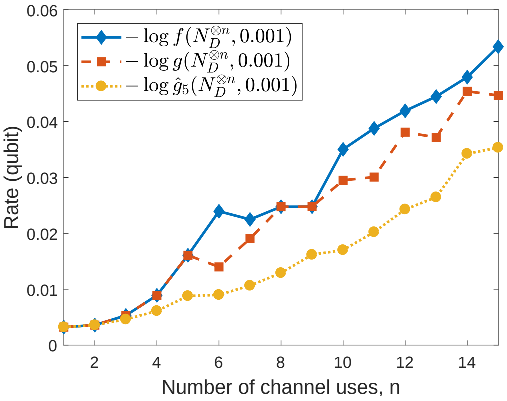


Figure 2: Comparsion of SDP converse bounds $-\log f$, $-\log g$, $-\log \widehat{g}_5$ in the case of the qubit depolarizing channel with $p = 0.3$ and $\varepsilon = 0.001$.

We further study the asymptotic setting and show efficiently computable strong converse bound on quantum capacity. A previously known efficiently computable strong converse bound for general channels is the partial transposition bound [8] introduced by Holevo and Werner: $Q(\mathcal{N}) \leq Q_{\Theta}(\mathcal{N}) := \log \|J_{\mathcal{N}}^{T_B}\|_{cb}$, where $\|\cdot\|_{cb}$ is the completely bounded trace norm. Moreover, Tomamichel, Wilde and Winter [9] introduced the Rains information $R(\mathcal{N})$ and show that it is a strong converse bound on quantum capacity, i.e., $Q(\mathcal{N}) \leq R(\mathcal{N}) :=$

$\max_{\rho_A \in \mathcal{S}(A)} \min_{\sigma \in \mathrm{PPT}'} D\left(\mathcal{N}_{A' \to B}\left(\phi_{AA'}\right) \| \sigma\right)$, where $\phi_{AA'}$ is a purification of ρ_A and the set PPT' = $\{\sigma \geq 0 : \|\sigma^{T_B}\|_1 \leq 1\}$. There are other converse bounds on quantum capacity [25, 26, 27, 28, 29, 30, 31, 32], which require specific settings to be computable and relatively tight. However, our work provides another efficiently computable strong converse bound (Q_Γ) for the quantum capacity of an arbitrary quantum channel, i.e.,

$$Q\left(\mathcal{N}\right) \leq Q_{\to}\left(\mathcal{N}\right) \leq Q_{\mathrm{PPT}}\left(\mathcal{N}\right) \leq Q_\Gamma\left(\mathcal{N}\right) := \log \Gamma\left(\mathcal{N}\right),$$

where

$$\begin{aligned} \Gamma\left(\mathcal{N}\right) := \max\ & \operatorname{Tr} J_{\mathcal{N}} R_{AB} \\ \text{s.t. } & R_{AB} \geq 0, \operatorname{Tr} \rho_A = 1, \\ & -\rho_A \otimes \mathbb{1}_B \leq R_{AB}^{T_B} \leq \rho_A \otimes \mathbb{1}_B \end{aligned} \qquad (3)$$

Moreover, if the rate exceeds Q_Γ, the error probability will go to one exponentially fast.

We further explore the properties of Q_Γ and find that it can be refined as the max-Rains information, which is an analog to the Rains information [9] in the sense of replacing the relative entropy D with the max-relative entropy $D_{\max}$ [33], i.e.,

$$Q_\Gamma\left(\mathcal{N}\right) = \max_{\rho \in S(A)} \min_{\sigma \in \mathrm{PPT}'} D_{\max}\left(\mathcal{N}_{A' \to B}\left(\phi_{AA'}\right) \| \sigma\right).$$

Finally, we establish an inequality relationship among the known strong converse bounds on quantum capacity,

$$Q\left(\mathcal{N}\right) \leq Q_{\to}\left(\mathcal{N}\right) \leq R\left(\mathcal{N}\right) \leq Q_\Gamma\left(\mathcal{N}\right) \leq Q_\Theta\left(\mathcal{N}\right). \qquad (4)$$

We construct an explicit example $\mathcal{N}_r$ showing that the last inequality in Eq. (4) can be strict, where $\mathcal{N}_r = \sum_{i=0}^{1} E_i \cdot E_i^\dagger$ with $E_0 = |0\rangle\langle 0| + \sqrt{r}|1\rangle\langle 1|$ and $E_1 = \sqrt{1-r}|0\rangle\langle 1| + |1\rangle\langle 2|$. The numerical result is given in Fig. 3.

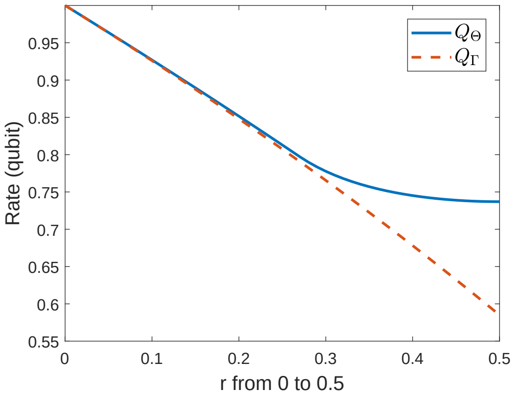


Figure 3: Comparison of strong converse bounds $Q_\Gamma\left(\mathcal{N}_r\right)$ and $Q_\Theta\left(\mathcal{N}_r\right)$. The channel parameter r ranges from 0 to 0.5.

References

[1] S. Lloyd, "Capacity of the noisy quantum channel," *Physical Review A*, vol. 55, no. 3, p. 1613, 1997.

[2] P. W. Shor, "The quantum channel capacity and coherent information," in *lecture notes, MSRI Workshop on Quantum Computation*, 2002.

[3] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 44–55, 2005.

[4] B. Schumacher and M. A. Nielsen, "Quantum data processing and error correction," *Physical Review A*, vol. 54, no. 4, p. 2629, 1996.

[5] H. Barnum, E. Knill, and M. A. Nielsen, "On quantum fidelities and channel capacities," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1317–1329, 2000.

[6] H. Barnum, M. A. Nielsen, and B. Schumacher, "Information transmission through a noisy quantum channel," *Physical Review A*, vol. 57, no. 6, p. 4153, 1998.

[7] T. Cubitt, D. Elkouss, W. Matthews, M. Ozols, D. Pérez-García, and S. Strelchuk, "Unbounded number of channel uses may be required to detect quantum capacity," *Nature Communications*, vol. 6, 2015.

[8] A. S. Holevo and R. F. Werner, "Evaluating capacities of bosonic Gaussian channels," *Physical Review A*, vol. 63, no. 3, p. 32312, 2001.

[9] M. Tomamichel, M. M. Wilde, and A. Winter, "Strong Converse Rates for Quantum Communication," *IEEE Transactions on Information Theory*, vol. 63, no. 1, pp. 715–727, jan 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7586115/

[10] D. Leung and W. Matthews, "On the power of PPT-preserving and non-signalling codes," *IEEE Transactions on Information Theory*, vol. 61, no. 8, pp. 4486–4499, 2015.

[11] R. Duan and A. Winter, "No-signalling-assisted zero-error capacity of quantum channels and an information theoretic interpretation of the Lovász number," *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 891–914, 2016.

[12] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[13] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 4947–4966, 2009.

[14] W. Matthews and S. Wehner, "Finite blocklength converse bounds for quantum channels," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7317–7329, 2014.

[15] L. Wang and R. Renner, "One-shot classical-quantum capacity and hypothesis testing," *Physical Review Letters*, vol. 108, no. 20, p. 200501, 2012.

[16] J. M. Renes and R. Renner, "Noisy channel coding via privacy amplification and information reconciliation," *IEEE Transactions on Information Theory*, vol. 57, no. 11, pp. 7377–7385, 2011.

[17] M. Tomamichel and M. Hayashi, "A hierarchy of information quantities for finite block length analysis of quantum tasks," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7693–7710, 2013.

[18] M. Berta, M. Christandl, and R. Renner, "The quantum reverse Shannon theorem based on one-shot information theory," *Communications in Mathematical Physics*, vol. 306, no. 3, pp. 579–615, 2011.

[19] M. Tomamichel and V. Y. F. Tan, "Second-order asymptotics for the classical capacity of image-additive quantum channels," *Communications in Mathematical Physics*, vol. 338, no. 1, pp. 103–137, 2015.

[20] S. Beigi, N. Datta, and F. Leditzky, "Decoding quantum information via the Petz recovery map," *Journal of Mathematical Physics*, vol. 57, no. 8, p. 082203, aug 2016.

[21] M. Tomamichel, *Quantum Information Processing with Finite Resources: Mathematical Foundations.* Springer, 2015, vol. 5.

[22] M. Tomamichel, M. Berta, and J. M. Renes, "Quantum coding with finite resources," *Nature Communications*, vol. 7, p. 11419, 2016.

[23] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, 1996.

[24] L. G. Khachiyan, "Polynomial algorithms in linear programming," *USSR Computational Mathematics and Mathematical Physics*, vol. 20, no. 1, pp. 53–72, 1980.

[25] G. Smith, J. Smolin, and A. Winter, "The quantum capacity with symmetric side channels," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4208–4217, 2008.

[26] D. Sutter, V. B. Scholz, A. Winter, and R. Renner, "Approximate Degradable Quantum Channels," *arXiv:1412.0980*, dec 2014.

[27] L. Gao, M. Junge, and N. LaRacuente, "Capacity Bounds via Operator Space Methods," *arXiv:1509.07294*, 2015.

[28] D. Bruß, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin, "Optimal universal and state-dependent quantum cloning," *Physical Review A*, vol. 57, no. 4, p. 2368, 1998.

[29] N. J. Cerf, "Pauli cloning of a quantum bit," *Physical Review Letters*, vol. 84, no. 19, p. 4497, 2000.

[30] M. M. Wolf and D. Perez-Garcia, "Quantum capacities of channels with small environment," *Physical Review A*, vol. 75, no. 1, p. 12303, 2007.

[31] G. Smith and J. A. Smolin, "Additive extensions of a quantum channel," in *Proceedings of IEEE Information Theory Workshop (ITW)*. IEEE, 2008, pp. 368–372.

[32] F. Leditzky, N. Datta, and G. Smith, "Useful states and entanglement distillation," *arXiv:1701.0308*, jan 2017.

[33] N. Datta, "Min-and max-relative entropies and a new entanglement monotone," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2816–2826, 2009.

Locality Preserving Logical Operators in Topological Stabiliser Codes

Paul Webster[1] * Stephen D. Bartlett[1]

[1] *Centre for Engineered Quantum Systems, School of Physics, The University of Sydney, Sydney, NSW 2006, Australia*

Abstract. Locality preserving logical operators are naturally fault tolerant in topological codes, since they preserve correctability of errors. We provide a framework for finding all locality preserving logical operators admitted by a large and important class of topological stabiliser codes; those equivalent to a finite number of copies of a toric code. We use this approach to explicitly classify the operators for these codes with two types of boundary conditions, and illustrate how it may similarly be applied to codes with different boundaries from these cases. This framework and classification provides a more complete understanding of the potential and limitations of these codes for allowing fault tolerant quantum computing, which we believe will be valuable in guiding future quantum computing architectures.

Keywords: fault tolerant quantum computing, topological stabilizer code, locality preserving, gapped domain wall, toric code, color code

Topological stabiliser codes are a class of quantum error correcting codes that have attracted widespread attention due to their simplicity and natural realisation as physical systems [1, 2]. The codestates of such codes are topologically protected, so that all local errors are correctable [3]. As such, logical operators implemented on information encoded in such codes are fault tolerant provided they are locality preserving [4]. However, Bravyi and König [4], and subsequently Pastawski and Yoshida [5], have shown that strong constraints apply to the types of locality preserving operations that may be implemented fault tolerantly in a topological stabiliser code. Specifically, these constraints are upper bounds on the level of the Clifford hierarchy from which gates may be locality preserving in a given topological stabiliser code. Recently, Yoshida has also observed correspondences between gapped domain walls and locality preserving logical operators, and domain walls and symmetry protected excitations, that suggest an approach to building a framework for finding all locality preserving logical operators which may be implemented in a given topological stabiliser code [6]. Our work builds such a framework and applies it to classify the locality preserving logical operators which are implementable in any topological stabiliser code which is locally equivalent to a finite number of toric codes. This is a large class of codes, containing all non-chiral, translationally invariant two dimensional topological stabiliser codes [3], and a wide range of higher dimensional topological stabiliser codes, including all colour codes [7].

1 Main Ideas

The key ideas behind this work build on Yoshida's observations in [6]. Specifically, we argue that in a d dimensional topological stabiliser code there are one-to-one correspondences between $k+1$ dimensional locality preserving logical operators, k dimensional transparent, gapped domain walls and k dimensional excitations for $1 \leq k \leq d-1$. We also note that transparent, gapped domain walls correspond to permutations of excitations which, in addition to preserving exchange, braiding and fusion statistics [8], are constrained by the dimensions of the code, domain wall and excitation.

We use these correspondences to build a recursive approach to finding all locality preserving logical operators which may be implemented in a given topological stabiliser code. Specifically, all domain walls which permute eigenstate excitations are first found. Such domain walls correspond to Clifford locality preserving logical operators. They also may be viewed as symmetry protected excitations of the code, which we label as $\mathcal{C}_2$ excitations. We then consider domain walls which map eigenstate excitations to excitations which may include $\mathcal{C}_2$ excitations. These domain walls correspond to locality preserving logical operators from the third level of the Clifford hierarchy, and may in turn be viewed as new symmetry protected excitations of the code. Considering domain walls which involve these new excitations we can then find locality preserving logical gates in the fourth level of the Clifford hierarchy. This process is then continued on for higher and higher levels of the hierarchy. The bound of Bravyi and König ensures that the process terminates, since above a certain level of the hierarchy there are no new locality preserving logical gates, and so no new domain walls will be found. Once this termination occurs, all locality preserving logical gates admitted by the code will have been found.

2 Results

We apply this approach to the class of topological stabiliser codes that are locally equivalent to a finite number of identical d dimensional toric codes, considering explicitly two types of boundary conditions. Firstly, we consider codes locally equivalent to a disjoint union of a finite number of generalised surface codes in d spatial dimensions. For such codes we demonstrate that all non-Clifford locality preserving logical operators are products of Pauli X operators and Pauli Z operators controlled by k qubits, which we denote C^kZ. We also determine a nec-

*pweb6304@uni.sydney.edu.au

essary and sufficient condition for such C^kZ operators to be admitted as a function of k, d and the dimensionality of electric charges of the surface codes involved. Secondly, we consider the d-dimensional colour code, which may be viewed as d identical copies of a toric code folded together to encode a single logical qubit [7]. We show how the above results may be adapted for these different boundary conditions. From this, we show that all non-Clifford locality preserving logical gates in such codes are products of Pauli X operators and rotations of the Bloch sphere about the Z axis of $2^{1-k}\pi$, for $k \in \mathbb{N}$, which we denote by R_k. We show that an analogous necessary and sufficient condition to that above determines which operators of this type are allowed. More generally, we indicate how these results could be further generalised to a more diverse range of boundary conditions for such codes. This can allow our framework to have broad applicability to a wide range of codes. We also suggest a generalisation of our results to the larger class of quantum codes referred to as abelian quantum double models [9].

3 Impact and Importance

The approach we develop could be valuable in assisting with finding low overhead schemes for implementing quantum algorithms. Specifically, magic state distillation, which is necessary where locality preserving implementations of gates are not possible, requires a very large number of physical qubits [10]. By providing an approach to determining which gates are locality preserving in a topological stabiliser code we offer the potential for algorithms which minimise the need for magic state distillation to be chosen for such a code. This could make implementation of these algorithms feasible even on relatively small quantum computers which may be expected to precede larger, more powerful quantum computers.

More specifically, our results regarding surface and colour codes are particularly relevant, since these codes are believed to be promising approaches for achieving both small-scale protected quantum memories in the short term [2], and large-scale memories in the long term [10]. Understanding the potential of these codes to allow for fault tolerant quantum computation, therefore, is an important aspect of working towards achieving quantum computing with such memories. While previous work had identified particular locality preserving logical operators admitted by these codes [7, 9], and general constraints [4, 5], we provide the first complete classification of all such operators. This allows for more certain assessment of these codes, and for comparisons with other proposed schemes for quantum computation. We hope this may help to guide realisations of these schemes, or motivate searches for alternatives.

Finally, our work provides deeper understanding of the codes we study, which may illuminate approaches that may be taken to explore important open questions. Specifically, our work illustrates how the bounds of Bravyi and König, and of Pastawski and Yoshida, can be viewed to arise for the codes we consider from dimensional constraints on the actions of domain walls on excitations. We hope that this may provide insight that may allow for similar bounds to be derived for a broader class of codes, such as non-abelian quantum double models. Another area our work may illuminate is explorations of alternative approaches to achieving fault tolerant operations based on braiding defects [11]. Such defects are known to emerge as boundaries of domain walls in two dimensional codes [12]. We hope that our exploration of domain walls and excitations in higher dimensional codes may provide insight for future explorations of analogous higher dimensional defects, which may prove to allow for a rich braiding structure, and a range of fault tolerant operators.

References

[1] H. Bombin, “Structure of 2D Topological Stabilizer Codes”, Commun. Math. Phys., **327**, pp 387-432, 2014.

[2] D. Nigg, M. Müller, E. A. Martinez, P. Schindler, M. Hennrich, T. Monz, M. A. Martin-Delgado and R. Blatt. “Quantum computations on a topologically encoded qubit”, Science, **345**, 6194, 2014.

[3] H. Bombin, G. Duclos-Cianci and D. Poulin, “Universal topological phase of two-dimensional stabilzer codes”, New J. Phys, **14**, 073048, 2012.

[4] S. Bravyi and R. König, “Classification of topologically protected gates for local stabilizer codes”, Phys. Rev. Lett., **110**, 170503, 2013.

[5] F. Pastawski and B. Yoshida, ”Fault-tolerant logical gates in quantum error-correcting codes”, Phys. Rev. A, **91**, 012305, 2015.

[6] B. Yoshida, “Topological color code and symmetry-protected topological phases”, Phys. Rev. B, **91**, 245131, 2015.

[7] A. Kubica, B. Yoshida and F. Pastawski, “Unfolding the color code”, New J. Phys., **17**, 083026, 2015.

[8] T. Lan, J. Wang and X.-G. Wen, “Gapped Domain Walls, Gapped Boundaries and Topological Degeneracy”, Phys. Rev. Lett., **114**, 076402, 2015.

[9] B. Yoshida, “Gapped boundaries, group cohomology and fault-tolerant logical gates”, arXiv:1509.03626v1 [cond-mat. str-el], 2015.

[10] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, “Surface codes: Towards practical large-scale quantum computation”, Phys. Rev. A, **86**, 032324, 2012.

[11] H. Bombin, Topological Order with a Twist: Ising Anyons from an Abelian Model, Phys. Rev. Lett., **105**, 030403, 2010.

[12] M. Barkeshli, P. Bonderson, M. Cheng and Z. Wang, Symmetry, Defects and Gauging of Topological Phases, arXiv:1410.4540v2 [cond-mat.str-el], 2014.

Self-testing of binary observables based on commutation

Jędrzej Kaniewski[1] *

[1]*QMATH, Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark*

Abstract. We consider the problem of certifying binary observables based on a Bell inequality violation alone, a task known as *self-testing of measurements.* We introduce a family of commutation-based measures, which encode all the distinct arrangements of two projective observables on a qubit. These quantities by construction take into account the usual limitations of self-testing and since they are "weighted" by the (reduced) state, they automatically deal with rank-deficient reduced density matrices. We show that these measures can be estimated from the observed Bell violation in several scenarios and the proofs rely only on standard linear algebra. The trade-offs turn out to be tight and, in particular, they give non-trivial statements for arbitrarily small violations. On the other extreme, observing the maximal violation allows us to deduce precisely the form of the observables, which immediately leads to a complete rigidity statement. In particular, we show that for all $n \geq 3$ the n-partite Mermin-Ardehali-Belinskii-Klyshko inequality self-tests the n-partite Greenberger-Horne-Zeilinger state and maximally incompatible qubit measurements on every party. Our results imply that any pair of projective observables on a qubit can be certified in a truly robust manner. Finally, we show that commutation-based measures give a convenient way of expressing relations among more than two observables.

Keywords: self-testing, Bell nonlocality, device-independence

1 Introduction

The fact that quantum mechanics is incompatible with the concept of local realism [9] is arguably one of the most surprising features of the quantum world. It should therefore come as no surprise that Bell nonlocality is an attractive field of research for both theoreticians (see Ref. [10] for a review) and experimentalists (see e.g. the recent loophole-free Bell tests [23, 22, 39, 24]). An important practical application of Bell nonlocality is *device-independent quantum cryptography*, whose goal is to prove the security of protocols executed using potentially untrusted devices (see Refs. [7, 2, 18, 1, 19] for the early contributions and Ref. [21] for a relatively up-to-date review). What makes this task possible is the fact that observing nonlocal correlations allows us to draw conclusions about the inner workings of the untrusted devices. In fact, certain extremal quantum correlations identify exactly the quantum system under consideration (up to well-understood equivalences). For example the only manner to achieve the maximal violation of the Clauser-Horne-Shimony-Holt (CHSH) [15] inequality is to perform anticommuting measurements on the maximally entangled state of two qubits [41, 43, 38]. Mayers and Yao realised that this allows us to certify quantum devices under minimal assumptions and they also coined the term *self-testing* [28, 29]. The general question is simple: "We have conducted a Bell test and observed certain nonlocal correlations. What can we rigorously deduce about the state shared between the devices and the measurements performed?".

The first self-testing results only applied in the case of observing the ideal statistics. While interesting from the foundational point of view, it is not sufficient for practical applications. In order to make statements relevant for experiments we must make them *robust*, i.e. we have to show that if the observed statistics are close to the ideal ones, then the quantum device should be close (in some well-defined sense) to the perfect realisation. To simplify the problem, instead of looking at the entire probability distribution, we often only look at the violation of some fixed Bell inequality. A Bell inequality is given by a vector of real coefficients $c_{abxy} \in \mathbb{R}$ and the corresponding Bell value β is defined as

$$\beta := \sum_{abxy} c_{abxy} \Pr[a, b|x, y],$$

where $\Pr[a, b|x, y]$ is the probability of observing outputs a, b given inputs x, y. Let β^L and β^Q be the largest values achievable by local-realistic theories and quantum mechanics, respectively, and suppose that $\beta^L < \beta^Q$. A necessary condition to make a self-testing statement is to observe some violation ($\beta > \beta_L$) and a self-testing result is called robust if we can make conclusions even if the violation is not maximal ($\beta < \beta^Q$). It is important to distinguish self-testing results which only apply if the violation is close to maximal from the ones that cover a sizeable portion of the interval $[\beta^L, \beta^Q]$. The latter apply to real-world experiments and, therefore, might actually be useful in designing robust and efficient testing procedures for real devices. Deriving such experimentally-relevant self-testing statements is precisely the focus of this work.

The main challenge in deriving robust self-testing statements lies in finding a natural mathematical formulation of the problem. Since our goal is to make statements even for statistics significantly differing from the ideal setup, we cannot aim for a complete description. We should instead pin down the relevant property and certify precisely that property. This is how our approach differs from the standard formulation, which attempts to certify closeness (in trace distance) to the perfect realisation.

*jkaniewski@math.ku.dk

Our primary goal is to certify two-outcome (binary) projective measurements. We propose a novel formulation based on commutation, which recovers several previous results as extreme cases. Commutation-based measures are easily computable, have a simple physical interpretation and demonstrate that all pairs of projective qubit observables can be certified in a robust fashion.

2 Previous self-testing results

While self-testing of quantum states has received significant attention in the regime of small [40, 33, 46, 30, 4, 31, 42, 36, 16, 20, 32, 12, 35, 17, 13] and experimentally-relevant [6, 47, 45, 37, 5, 44, 25] robustness, self-testing of measurements is a significantly less studied topic. Although most results in the small robustness regime come as complete rigidity statements (i.e. they also characterise the optimal measurements), there are only two approaches that yield experimentally-relevant robustness [5, 11, 14].

3 Self-testing of observables based on commutation

In our framework we certify observables of one party at a time, i.e. we have a separate statement for each party, which only depends on the local observables and the reduced state. This is in line with the idea of focusing on a single property, instead of certifying the whole setup.

The two inherent limitations of self-testing (i.e. properties that cannot be deduced from the outcome statistics) are: the presence of auxiliary degrees of freedom and the application of local unitaries. It is clear that these two equivalences do not affect commutation relations between observables. We might therefore conclude that what we should be certifying is precisely the commutation structure between the observables. This is, however, not quite correct as we can only make statements about the observables *on the support of the (reduced) state.* Therefore, instead of making statements about the observables, we will consider scalar quantities of the form $t := \operatorname{tr}(T\rho_A)$, where ρ_A is the reduced state on the subsystem to be measured and T is a Hermitian operator constructed from the observables (whose exact definition depends on the commutation structure we wish to certify). An appealing feature of these measures is the fact that the maximal value of t is achieved by essentially just one arrangement of observables. Let us stress that whenever we make a statement directly about the operators, we implicitly assume that the reduced state is full-rank.

4 Methods

A binary observable is a Hermitian operator A satisfying $-\mathbb{I} \leq A \leq \mathbb{I}$ (we do not a priori assume projectivity). It is well known that in the case of binary observables commutators and anticommutators of observables appear in the square of the Bell operator. Let W be the Bell operator, let $\beta := \operatorname{tr}(W\rho_{AB})$ be the Bell value and from the Cauchy-Schwarz inequality we deduce that $\beta^2 = [\operatorname{tr}(W\rho_{AB})]^2 \leq \operatorname{tr}(W^2\rho_{AB}) \cdot \operatorname{tr}\rho_{AB} = \operatorname{tr}(W^2\rho_{AB})$. Therefore, proving an operator inequality

$$W^2 \leq g(A_0, A_1) \otimes \mathbb{I}, \tag{1}$$

where A_0 and A_1 are the observables of Alice and g is a function which outputs a Hermitian operator, immediately implies $\beta \leq \sqrt{\operatorname{tr}\left(g(A_0, A_1)\rho_A\right)}$. If the right-hand side provides a useful characterisation of the observables of Alice, this constitutes a self-testing statement. Let us stress that for projective observables W^2 can often be written explicitly as a function of their commutators and anticommutators, which provides helpful intuition on the possible form of the function $g(A_0, A_1)$.

5 Certifying anticommuting observables

In the CHSH scenario Alice and Bob measure one of two binary observables denoted by A_j and B_k for $j, k \in \{0, 1\}$. The CHSH operator is defined as

$$W := (A_0 + A_1) \otimes B_0 + (A_0 - A_1) \otimes B_1$$

for which $\beta^L = 2$ and $\beta^Q = 2\sqrt{2}$. We prove that

$$W^2 \leq 4 \cdot \mathbb{I} \otimes \mathbb{I} - [A_0, A_1] \otimes [B_0, B_1]$$

and by noticing that $|[B_0, B_1]| \leq 2 \cdot \mathbb{I}$ we obtain

$$W^2 \leq 4 \cdot \mathbb{I} \otimes \mathbb{I} + 2\,|[A_0, A_1]| \otimes \mathbb{I}.$$

From the argument outlined above we deduce that

$$\beta \leq 2\sqrt{1 + t}, \tag{2}$$

where $t := \frac{1}{2} \operatorname{tr}\left(|[A_0, A_1]|\rho_A\right) \in [0, 1]$ is the *effective commutator.* This scalar quantity is invariant under local unitaries and adding extra degrees of freedom, it avoids making any statement about the observables outside the support of ρ_A and is easily computable. The physical interpretation is clear: t measures the incompatibility of Alice's observables "weighted" by the reduced state ρ_A. The matrix modulus, which arises in the derivation, avoids cancellations, e.g. $t = 0$ implies that the observables commute on the support of ρ_A, which prevents us from observing any violation. On the other extreme, the maximal value $t = 1$ implies the existence of a unitary U_A such that

$$A_0 = U_A(\sigma_x \otimes \mathbb{I})U_A^\dagger \quad \text{and} \quad A_1 = U_A(\sigma_y \otimes \mathbb{I})U_A^\dagger \tag{3}$$

(recall the assumption that ρ_A is full-rank). This shows that t is a useful measure of how close Alice's observables are to a pair of anticommuting observables on a qubit. The inequality (2) is interesting for several reasons: it gives a non-trivial statement as soon as $\beta > 2$, it is tight and observing the maximal violation $\beta = 2\sqrt{2}$ implies $t = 1$, which allows us to deduce the exact form of the observables. Although our primary goal is certifying observables, in the case of perfect statistics this argument immediately gives a complete rigidity statement. An extension of this method allows us to certify maximally incompatible observables in the multipartite setting using the Mermin-Ardehali-Belinskii-Klyshko

(MABK) [34, 3, 8] family. The resulting trade-offs are, again, tight and we obtain complete rigidity statements for the perfect statistics.

6 Certifying non-maximally incompatible observables

In the previous cases the optimal observables on every party correspond to anticommuting observables on a qubit. Here, we show that an arbitrary pair of qubit observables, not necessarily maximally incompatible, is exactly characterised through their commutation relation. For $\alpha \geq 1$ we consider the generalisation of the CHSH inequality:

$$W_\alpha := \alpha(A_0 + A_1) \otimes B_0 + (A_0 - A_1) \otimes B_1$$

introduced by Lawson, Linden and Popescu [27]. The local-realistic and quantum bounds for this inequality equal $\beta^L_\alpha = 2\alpha$ and $\beta^Q_\alpha = 2\sqrt{\alpha^2+1}$ (hence $\beta^L_\alpha < \beta^Q_\alpha$ for all $\alpha \geq 1$) and the maximal violation is achieved by measuring a maximally entangled two-qubit state, but the optimal observables of Alice are no longer maximally incompatible. We show that

$$W_\alpha^2 \leq 2(\alpha^2+1) \cdot \mathbb{I} \otimes \mathbb{I} + T_\alpha \otimes \mathbb{I}$$

for $T_\alpha := (\alpha^2-1)\{A_0, A_1\} + 2\alpha|[A_0, A_1]|$. Defining $t_\alpha := \frac{1}{4}\operatorname{tr}(T_\alpha \rho_A) - \frac{1}{2}(\alpha^2-1)$, which recovers the effective commutator for $\alpha = 1$, allows us to write $\beta_\alpha \leq 2\sqrt{\alpha^2 + t_\alpha}$. If the observables of Alice commute, we have $t_\alpha \leq 0$, which immediately recovers the classical bound. On the other hand, observing the maximal violation $\beta_\alpha = \beta^Q_\alpha$ implies that $t_\alpha = 1$ and that there exists of a unitary U_A such that

$$A_0 = U_A(\sigma_x \otimes \mathbb{I})U_A^\dagger,$$
$$A_1 = U_A\big([\cos\theta_\alpha\, \sigma_x + \sin\theta_\alpha\, \sigma_y] \otimes \mathbb{I}\big)U_A^\dagger$$

for $\theta_\alpha := \arccos\big(\frac{\alpha^2-1}{\alpha^2+1}\big) \in (0, \pi/2]$. This characterises the exact commutation structure between the observables and by considering $\alpha \in [1, \infty)$ we can certify any angle between two projective observables on a qubit. The maximal violation is only possible if the observables of Bob anticommute, which leads directly to a rigidity statement for the generalised CHSH inequality.

All the details can be found in the full version of this work [26]. In addition we show that this method is useful for treating the case of more than two observables and discuss potential extensions to the case of measurements with more than two outcomes.

References

[1] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani. Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.*, 98(230501), 2007. DOI: 10.1103/PhysRevLett.98.230501.

[2] A. Acín, N. Gisin, and L. Masanes. From Bell's theorem to secure quantum key distribution. *Phys. Rev. Lett.*, 97(120405), 2006. DOI: 10.1103/PhysRevLett.97.120405.

[3] M. Ardehali. Bell inequalities with a magnitude of violation that grows exponentially with the number of particles. *Phys. Rev. A*, 46(5375), 1992. DOI: 10.1103/PhysRevA.46.5375.

[4] C. Bamps and S. Pironio. Sum-of-squares decompositions for a family of Clauser-Horne-Shimony-Holt-like inequalities and their application to self-testing. *Phys. Rev. A*, 91(052111), 2015. DOI: 10.1103/PhysRevA.91.052111.

[5] J.-D. Bancal, M. Navascués, V. Scarani, T. Vértesi, and T. H. Yang. Physical characterization of quantum devices from nonlocal correlations. *Phys. Rev. A*, 91(022115), 2015. DOI: 10.1103/PhysRevA.91.022115.

[6] C.-E. Bardyn, T. C. H. Liew, S. Massar, M. McKague, and V. Scarani. Device independent state estimation based on Bell's inequalities. *Phys. Rev. A*, 80(062327), 2009. DOI: 10.1103/PhysRevA.80.062327.

[7] J. Barrett, L. Hardy, and A. Kent. No signaling and quantum key distribution. *Phys. Rev. Lett.*, 95(010503), 2005. DOI: 10.1103/PhysRevLett.95.010503.

[8] A. V. Belinskii and D. N. Klyshko. Interference of light and Bell's theorem. *Phys. Usp.*, 36(653), 1993. DOI: 10.1070/PU1993v036n08ABEH002299.

[9] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1(195), 1964.

[10] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86(419), 2014. DOI: 10.1103/RevModPhys.86.419.

[11] D. Cavalcanti and P. Skrzypczyk. Quantitative relations between measurement incompatibility, quantum steering, and nonlocality. *Phys. Rev. A*, 93(052112), 2016. DOI: 10.1103/PhysRevA.93.052112.

[12] R. Chao, B. W. Reichardt, C. Sutherland, and T. Vidick. Test for a large amount of entanglement, using few measurements. 2016. arXiv: 1610.00771.

[13] R. Chao, B. W. Reichardt, C. Sutherland, and T. Vidick. Overlapping qubits. 2017. arXiv: 1701.01062.

[14] S.-L. Chen, C. Budroni, Y.-C. Liang, and Y.-N. Chen. Natural framework for device-independent quantification of quantum steerability, measurement incompatibility, and self-testing. *Phys. Rev. Lett.*, 116(240401), 2016. DOI: 10.1103/PhysRevLett.116.240401.

[15] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23(880), 1969. DOI: 10.1103/PhysRevLett.23.880.

[16] A. W. Coladangelo. Parallel self-testing of (tilted) EPR pairs via copies of (tilted) CHSH. *Quant. Inf. Comp.*, 17(831), 2017. arXiv: 1609.03687.

[17] A. W. Coladangelo, K. T. Goh, and V. Scarani. All pure bipartite entangled states can be self-tested. *Nat. Commun.*, 8(15485), 2017. DOI: 10.1038/ncomms15485.

[18] R. Colbeck. *Quantum and relativistic protocols for secure multi-party computation*. PhD thesis, University of Cambridge, 2006. arXiv: 0911.3814.

[19] R. Colbeck and A. Kent. Private randomness expansion with untrusted devices. *J. Phys. A: Math. Theor.*, 44(095305), 2011. DOI: 10.1088/1751-8113/44/9/095305.

[20] M. Coudron and A. Natarajan. The parallel-repeated magic square game is rigid. 2016. arXiv: 1609.06306.

[21] A. Ekert and R. Renner. The ultimate physical limits of privacy. *Nature*, 507(443), 2014. DOI: 10.1038/nature13132.

[22] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger. Significant-loophole-free test of Bell's theorem with entangled photons. *Phys. Rev. Lett.*, 115(250401), 2015. DOI: 10.1103/PhysRevLett.115.250401.

[23] B. Hensen, H. Bernien, A. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(682), 2015. DOI: 10.1038/nature15759.

[24] B. Hensen, N. Kalb, M. S. Blok, A. Dréau, A. Reiserer, R. F. L. Vermeulen, R. N. Schouten, M. Markham, D. J. Twitchen, K. Goodenough, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. Loophole-free Bell test using electron spins in diamond: second experiment and additional analysis. *Sci. Rep.*, 6(30289), 2016. DOI: 10.1038/srep30289.

[25] J. Kaniewski. Analytic and nearly optimal self-testing bounds for the Clauser-Horne-Shimony-Holt and Mermin inequalities. *Phys. Rev. Lett.*, 117(070402), 2016. DOI: 10.1103/PhysRevLett.117.070402.

[26] J. Kaniewski. Self-testing of binary observables based on commutation. *Phys. Rev. A*, 95(062323), 2017. DOI: 10.1103/PhysRevA.95.062323.

[27] T. Lawson, N. Linden, and S. Popescu. Biased nonlocal quantum games. 2010. arXiv: 1011.6245.

[28] D. Mayers and A. Yao. Quantum cryptography with imperfect apparatus. *Proceedings 39th Annual Symposium on Foundations of Computer Science*, (503), 1998. DOI: 10.1109/SFCS.1998.743501.

[29] D. Mayers and A. Yao. Self testing quantum apparatus. *Quant. Inf. Comp.*, 4(273), 2004. arXiv: quant-ph/0307205.

[30] M. McKague. Self-testing graph states. *Theory of Quantum Computation, Communication, and Cryptography. TQC 2011. Lecture Notes in Computer Science*, 6745(104), 2014. DOI: 10.1007/978-3-642-54429-3_7.

[31] M. McKague. Self-testing in parallel. *New J. Phys.*, 18(045013), 2016. DOI: 10.1088/1367-2630/18/4/045013.

[32] M. McKague. Self-testing in parallel with CHSH. 2016. arXiv: 1609.09584.

[33] M. McKague, T. H. Yang, and V. Scarani. Robust self-testing of the singlet. *J. Phys. A: Math. Theor.*, 45(455304), 2012. DOI: 10.1088/1751-8113/45/45/455304.

[34] N. D. Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65(1838), 1990. DOI: 10.1103/PhysRevLett.65.1838.

[35] A. Natarajan and T. Vidick. Robust self-testing of many-qubit states. 2016. arXiv: 1610.03574.

[36] D. Ostrev and T. Vidick. Entanglement of approximate quantum strategies in XOR games. 2016. arXiv: 1609.01652.

[37] K. F. Pál, T. Vértesi, and M. Navascués. Device-independent tomography of multipartite quantum states. *Phys. Rev. A*, 90(042340), 2014. DOI: 10.1103/PhysRevA.90.042340.

[38] S. Popescu and D. Rohrlich. Which states violate Bell's inequality maximally? *Phys. Lett. A*, 169, 1992. DOI: 10.1016/0375-9601(92)90819-8.

[39] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr, F. Marsili,

M. D. Shaw, J. A. Stern, C. Abellán, W. Amaya, V. Pruneri, T. Jennewein, M. W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill, and S. W. Nam. Strong loophole-free test of local realism. *Phys. Rev. Lett.*, 115(250402), 2015. DOI: 10.1103/PhysRevLett.115.250402.

[40] W. Slofstra. Lower bounds on the entanglement needed to play XOR non-local games. *J. Math. Phys.*, 52(102202), 2011. DOI: 10.1063/1.3652924.

[41] S. J. Summers and R. F. Werner. Maximal violation of Bell's inequalities is generic in quantum field theory. *Commun. Math. Phys.*, 110(247), 1987. DOI: 10.1007/BF01207366.

[42] I. Šupić, R. Augusiak, A. Salavrakos, and A. Acín. Self-testing protocols based on the chained Bell inequalities. *New J. Phys.*, 18(035013), 2016. DOI: 10.1088/1367-2630/18/3/035013.

[43] B. S. Tsirelson. Some results and problems on quantum Bell-type inequalities. *Hadronic J. Suppl.*, 8(329), 1993.

[44] X. Wu, J.-D. Bancal, M. McKague, and V. Scarani. Device-independent parallel self-testing of two singlets. *Phys. Rev. A*, 93(062121), 2016. DOI: 10.1103/PhysRevA.93.062121.

[45] X. Wu, Y. Cai, T. H. Yang, H. N. Le, J.-D. Bancal, and V. Scarani. Robust self-testing of the three-qubit W state. *Phys. Rev. A*, 90(042339), 2014. DOI: 10.1103/PhysRevA.90.042339.

[46] T. H. Yang and M. Navascués. Robust self-testing of unknown quantum systems into any entangled two-qubit states. *Phys. Rev. A*, 87(050102(R)), 2013. DOI: 10.1103/PhysRevA.87.050102.

[47] T. H. Yang, T. Vértesi, J.-D. Bancal, V. Scarani, and M. Navascués. Robust and versatile black-box certification of quantum devices. *Phys. Rev. Lett.*, 113(040401), 2014. DOI: 10.1103/PhysRevLett.113.040401.

Forging the culture of quantum information science

Charles Bennett

IBM, USA

Abstract: Physicists, mathematicians and engineers, guided by what has worked well in their respective disciplines, have historically developed different scientific tastes, different notions of what constitutes an interesting, well-posed problem or an adequate solution. While this has led to some frustrating misunderstandings, it has invigorated the theory of communication and computation, enabling it to outgrow its brash beginnings with Turing, Shannon and von Neumann, and develop a coherent scientific taste of its own, adopting and domesticating ideas from thermodynamics and quantum mechanics that physicists had mistakenly thought belonged solely to their field, to better formalize the core concepts of communication and computation.

Role of Hypothesis Testing in Quantum Information

Masahito Hayashi[1] [2] *

[1] *Graduate School of Mathematics, Nagoya University, Nagoya, 464-8602, Japan*
[2] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117542, Singapore*

Abstract. Recently, it is well recognized that hypothesis testing has deep relations with other topics in quantum information theory as well as in classical information theory. These relations enable us to derive precise evaluation in the finite-length setting. However, such usefulness of hypothesis testing is not limited to information theoretical topics. For example, it can be used for verification of entangled state and quantum computer as well as guaranteeing the security of keys generated via quantum key distribution. In this talk, we overview these kinds of applications of hypothesis testing.

Keywords: Quantum hypothesis testing, classical-quantum channel coding, quantum key distribution, entangled state, verification of quantum computer

1 Quantum information theory and binary hypothesis testing

In information theory community it is well known that many information theoretical tasks can be analyzed by using the terminology of the binary asymmetric hypothesis testing. While there are many studies to focus on this relation, the first series study with this direction is the method of information spectrum, which was initiated by Han and Verdú [8][9][39], in which we convert the optimization problems in various information tasks into the binary asymmetric hypothesis testing, and the asymptotic behavior of the likelihood ratio plays a key role. This correspondence is valid without any assumption for the information source and/or the information channel, i.e., we do not need the independent and identical distributed condition nor Markovian condition. Due to the generality of the method of information spectrum, Nagaoka [31] considered to employ this method for quantum information theory. As a result, he found a remarkable relation between the classical-quantum channel coding and the quantum binary hypothesis testing, in which the correctly decoding probability is upper bounded by the performance of the corresponding quantum binary hypothesis testing in a canonical way. Later, Polyanskiy, Poor and Verdú [33] showed the same inequality only with the classical channel coding, which is called meta-converse theorem, nowadays. Nagaoka [31] also pointed out the notable relation between the quantum binary hypothesis testing and the Rényi relative entropy. These his results were presented in the first conference of ERATO Workshop on Quantum Information Science, which is the forerunner of AQIS conference series [31].

Based on this study, the author jointly with Nagaoka proved another remarkable relation between the classical-quantum channel coding and the quantum binary hypothesis testing [24, 11]. That is, they showed that the decoding error probability is upper bounded by the error probability of the corresponding quantum binary hypothesis testing, which is chosen slightly differently from the meta converse. Based on this method, the author derived the lower bound of the error exponent in the classical-quantum channel [15]. Wang and Renner [40] reformulated this result by introducing the hypothesis testing entropy. Later, Polyanskiy, Poor and Verdú showed the same inequality only with the classical channel coding, which is called the dependence test (DT) bound [33]. These two remarkable relations lead the breakthrough of the second order analysis of channel coding [16, 33].

Also, quantum data compression can be treated via the quantum binary hypothesis testing [10, 32]. Since reduction to a quantum analogue of likelihood ratio test, i.e., the quantum binary hypothesis testing is a very powerful method [20, 11], the following topics can be treated in this direction; quantum wiretap channel [19], universal (compound) channel coding [18, 4], entanglement concentration [12], entanglement dilution [20, Section 8.6], classical data compression with quantum side information [37], quantum Slepian-Wolf problem [3], classical random number generation with quantum side information [37], quantum state redistribution [2], and entanglement assisted communication over (quantum-quantum) point to point quantum channel, Gel'fand- Pinsker quantum channel, and quantum broadcast channel [1].

2 Verification of bipartite entangled state

Application of quantum hypothesis testing is not limited to the above type of theoretical aspects. Quantum hypothesis testing can be applied to more practical topics. One is verification of a bipartite entangled state. When an entangled state is generated experimentally, to use the generated bipartite entangled state, we need to verify whether the generated state is truly the intent bipartite entangled state. In the conventional qubit system, our verification is written as a binary POVM on the bipartite system. In this case, the direction of the error cannot be expected, it is suitable that the performance of this testing does not depend on the direction. That is, the POVM of the testing is preferred to be invariant for the group action preserving the entangled state. Such a testing method is formulated by using the irreducible decomposition of the group representation theory [22, 17].

However, in a usual optical device, like, spontaneous

*masahito@math.nagoya-u.ac.jp

parametric down-conversion (SPDC), a binary POVM is often constructed by a filter and a detector. That is, when the filter is passed, we have detection. Otherwise, we have no detection. In this situation, it is impossible to distinguish the following two cases, both of which correspond to no detection. One is the event that the photon pair is not generated so that it is not detected. The other is the event that the photon pair is generated, but the filter is not passed so that it is not detected. Then, the performance of the following two cases are not the same. We have the detection when the generated state is close to the intent entangled state. We have the detection when the generated state is far from the intent entangled state. Surprisingly, when the photon generation rate is known, the latter has better performance for this testing [27], whose experimental demonstration was also done [26].

3 Quantum key distribution

Another important application of hypothesis testing is its application to quantum key distribution, which is a method to share secure keys via quantum communications and classical communications [5]. Its security is trivial when the quantum communication channel has no noise. However, a real quantum communication channel has a certain amount of noise. When the amount is less than a threshold, we can generate secure keys by combining the error correction and the privacy amplification [36]. Quantum key distribution focuses on the bit basis and the phase basis. The error of the bit basis expresses the sacrifice rate in the error correction and the error of the phase basis expresses the sacrifice rate in the privacy amplification [13]. Later, a similar observation was also done via smooth entropy [38]. While we randomly choose check bits in quantum key distribution, its purpose is verification of the error rates of both bases.

However, in a realistic quantum key distribution, we often employ weak coherent pulses, which generates multiphoton state with some probability. In this situation, only a part of generated photons arrive at the receiver side. When a multi-photon state is generated, the eavesdropper, in principle, can obtain the transmitted information. Therefore, the required sacrifice rate in the privacy amplification is determined by the rate of pulses generated as multi-photon among the pluses arrived at the receiver side and the error rate of phase basis among pulses generated as single-photon and arrived at the receiver side [14]. That is, when we employ weak coherent pulses, we need to know these two ratios as well as the bit error ratio of the received pulses. For this purpose, we need to guarantee that these two rates are not greater than certain values. In this verification process, we randomly chooses several values of intensity of pulses [28]. Then, we obtain the detection rate and the error rates of the phase basis depending on the intensity. Using this data, we apply the method of hypothesis testing or the interval estimation, which employs the percent point [25]. Hence, we can verify these two ratios with certain intervals.

4 Verification of quantum computer

Hypothesis testing can be applied to the verification of quantum computer. In the conventional circuit model, it is quite difficult to predict the outcome of the circuit because the aim of the computation is to know the outcome. As an alternative model of quantum computer, measurement-based quantum computer (MBQC) is known [34], and is composed of a limited number of local measurements and a graph state, which is an entangled state of large size. Since the components of MBQC have known forms, its verification can be done by verifying these components. Therefore, when we can trust these local measurement devices, we can verify our computation outcome under the MBQC model by verifying the graph state. However, since available measurements are restricted to a limited number of local measurements, we need to realize the verification only with this limited class of measurements. Fortunately, in the graph state, the outcome of the Z basis predicts the behavior of the connected site. In the case of two-colorable graph, we can deterministically predict the outcome of the X basis on sites of one color from the Z basis outcome on sites of the other color. Using this property, we can verify whether the generated state is the intent graph state [23]. Since the above prediction is deterministic, this verification is can be done very efficiently. That is, the required number of sampling does not depend on the size of the graph state. Since this test checks whether the state belongs to the stabilizer defined by the pair of the X basis measurement and the Z basis measurement, it is called the stabilizer test.

Further, this method can be extended to the case when the measurement devices has noises and the generated graph state has noise [6]. In this case, we need to attach the fault-tolerant MBQC, which is often based on a topological surface code. When the noises of measurement devices are independent, they can be theoretically converted to the noises in the generated graph state. Once we fix the scheme of the fault-tolerant MBQC, we can define the set of correctable errors. When the noise belongs to the set of correctable error, the fault-tolerant MBQC works properly, i.e., the computation outcome is the correct value. Hence, it is sufficient to verify whether the error belongs to the the set. Since this test is also deterministic, the required number of sampling still does not depend on the size of the graph state.

Furthermore, we can make this kind of test even when the measurement device cannot be trusted. This kind of test is called self-testing, and the currently proposed method works with the noiseless case. In this setting, the testing of a graph state can be reduced to the testing of the Bell state in a canonical way. McKague et al [30] proposed the self-testing of the Bell state only with the CHSH test. However, the recently proposed method [21] combines the CHSH test and the stabilizer test so that the performance is much improved. This self-testing of the Bell state yields a self-testing of the graph state. When it applied to the verification of MBQC, the obtained scaling is much better than previously obtained verification methods [35, 29], and is the same as that of the paper [7], which employs a different method.

Acknowledgements

The author is grateful to all the collaborators of these presented results. Also, he is thankful for all the financial supports of these presented results.

References

[1] A. Anshu, R. Jain, and N. A. Warsi, arXiv:1702.01940 2017.

[2] A. Anshu, R. Jain, and N. A. Warsi, arXiv:1702.02396 2017.

[3] A. Anshu, R. Jain, and N. A. Warsi, arXiv:1703.09961 2017.

[4] A. Anshu, R. Jain, and N. A. Warsi, arXiv:1706.08286 2017.

[5] C. H. Bennett and G. Brassard, Quantum cryptography: public key distribution and coin tossing, *Proc. IEEE International Conference on Computers, Systems and Signal Processing*, (Bangalore, India), pp. 175–179, 1984.

[6] K. Fujii and M. Hayashi, Verifiable fault tolerance in measurement-based quantum computation, *Phys. Rev. A, Rapid Communication*, Accepted; arXiv:1610.05216.

[7] M. Hajdušek, C. A. Pérez-Delgado, and J. F. Fitzsimons, arXiv:1502.02563, 2015.

[8] T.-S. Han, *Information-Spectrum Methods in Information Theory*, Springer, Berlin, 2003. (Original Japanese version: Baifukan, 1998)

[9] T. S. Han and S. Verdú, *IEEE Trans. Inf. Theory*, **39** 752–772, 1993.

[10] M. Hayashi, *Phys. Rev. A*, **66** 032321, 2002.

[11] M. Hayashi, Hypothesis testing approach to quantum information theory, *Proceeding of COE Symposium on Quantum Information Theory*, Kyoto, Japan, September 2–3, 2003.

[12] M. Hayashi, *IEEE Trans. Inf. Theory*, **52** 1904–1921, 2006.

[13] M. Hayashi, *Phys. Rev. A*, **74**, 022307, 2006.

[14] M. Hayashi, *Phys. Rev. A*, **76** 012329, 2007.

[15] M. Hayashi, *Phys. Rev. A*, **76**, 062301, 2007.

[16] M. Hayashi, *IEEE Trans. Inf. Theory*, **55** 4947–4966, 2009.

[17] M. Hayashi, *New J. Phys.*, **11** 043028, 2009.

[18] M. Hayashi, *Comm. Math. Phys.*, **289** 1087-1098, 2009.

[19] M. Hayashi, *IEEE Trans. Inf. Theory*, **61** 5595–5622, 2015.

[20] M. Hayashi, *Quantum Information Theory: Mathematical Foundation*, *Graduate Texts in Physics*, Springer, 2017. (First edition: Springer, 2006).

[21] M. Hayashi and M. Hajdušek, arXiv:1603.02195, 2016.

[22] M. Hayashi, K. Matsumoto, and Y. Tsuda, *J. Phys. A: Math. Gen.* **39** 14427 – 14446, 2006.

[23] M. Hayashi and T. Morimae, *Phys. Rev. Lett.,* **115** 220502, 2015.

[24] M. Hayashi and H. Nagaoka, *IEEE Trans. Inf. Theory*, **49** 1753 – 1768, 2003.

[25] M. Hayashi and R. Nakayama, *New. J. Phys.*, **16** 063009, 2014.

[26] M. Hayashi, B.-S. Shi, A. Tomita, K. Matsumoto, Y. Tsuda, and Y.-K. Jiang, *Phys. Rev. A*, **74** 062321, 2006.

[27] M. Hayashi, A. Tomita, and K. Matsumoto, *New J. Phys.*, **10** 043029, 2008.

[28] W.-Y. Hwang, *Phys. Rev. Lett.* **91** 057901, 2003.

[29] M. McKague, *Theory of Computing* **12** 1, 2016.

[30] M. McKague, T. H. Yang, and V. Scarani, *J. Phys. A: Math. Theor.*, **45** 455304, 2012.

[31] H. Nagaoka, Strong converse theorems in quantum information theory, In *Proceedings of ERATO Workshop on Quantum Information Science 2001*, Univ. Tokyo, Tokyo, Japan, September 6-8, 2001, pp. 33.

[32] H. Nagaoka and M. Hayashi, *IEEE Trans. Inf. Theory*, **53** 534–549, 2007.

[33] Y. Polyanskiy, H.V. Poor, and S. Verdú, *IEEE Trans. Inf. Theory*, **56** 2307 – 2359, 2010.

[34] R. Raussendorf and H. J. Briegel, *Phys. Rev. Lett.* **86** 5188 (2001).

[35] B. W. Reichardt, F. Unger, and U. Vazirani, *Nature* **496** 456, 2013.

[36] P. W. Shor and J. Preskill, *Phys. Rev. Lett.*, **85** 441–444, 2000.

[37] M. Tomamichel and M. Hayashi, *IEEE Trans. Inf. Theory*, **59** 7693–7710, 2013.

[38] M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner, *Nat. Com.* **3** 634, 2012.

[39] S. Verdú and T. S. Han, *IEEE Trans. Inf. Theory*, **40** 1147–1157, 1994.

[40] L Wang and R Renner, *Phys. Rev. Lett.*, **108** 200501, 2012.

Practical round-robin-differential-phase-shift quantum key distribution

Zhen-Qiang Yin[1 2 3] Shuang Wang[1 2 3 *] Yun-Guang Han[1 2 3] Wei Chen[1 2 3 †]
Guang-Can Guo[1 2 3] Zheng-Fu Han[1 2 3]

[1] *CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei 230026, P. R. China*

[2] *Synergetic Innovation Center of Quantum Information & Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China*

[3] *State Key Laboratory of Cryptology, P. O. Box 5159, Beijing 100878, P. R. China*

Abstract. Round-robin-differential-phase (RRDPS) quantum key distribution (QKD) protocol has attracted intensive studies due to its distinct security characteristic, e.g., information leakage in RRDPS can be bounded without learning error rate of key bits. Nevertheless, its implementation is still far from practical due to the complication of its measurement device. Moreover, on the theoretical side, its security is still not clear in view of error rate. Here, by observing a potential phase randomization in the encoding states, we develop a theory to bound information leakage quite tightly and differently. Besides, the error rate is incorporated to improve the secret key rate, which is significant for the understanding of RRDPS. Based on our novel security proof, the practicality and performance of RRDPS can be both improved dramatically. Furthermore, we realize an experiment up to 140km fiber distance which is the longest achievable distance of RRDPS system until now, while the original security proof predicts no secret key can be generated in our experiment.

Keywords: QKD, RRDPS

1 Introduction

Unlike classical cryptography whose security relies on unproven mathematical assumptions, quantum key distribution (QKD) [1, 2] can information-theoretically distribute secret key bits between distant peers (such as Alice and Bob). According to quantum mechanics, any eavesdropping on quantum channel will inevitably introduce signal disturbance, which implies that Alice and Bob can bound the information leakage for the eavesdropper (Eve) through collecting the error rate of their raw key bits. Thus, monitoring signal disturbance is indispensable for almost all QKD protocols.

Surprisingly, recently proposed round-robin-differential-phase-shift (RRDPS) [3] protocol is an exception. In RRDPS protocol, Alice prepares a series of pulse trains, each consisting of L weak coherent pulses. The pulses are individually modulated to random phases out of 0 and π, and every L-pulse train can be handled as a packet. Upon receiving these packets, Bob measures the phase shift between the i-th pulse and $(i+r)$-th pulse of each packet, where r is randomly chosen from $[1, L-1]$ for each packet and $i+r \leqslant L$. Through a simple and comprehensive security proof [3], it has been pointed out that Eve's information on raw key bits I_{AE} is no larger than $h_2(n/(L-1))$, where n is the photon number of a packet. The main merit of RRDPS protocol is that I_{AE} does not depend on error rate of key bits, and thus can be treated as a constant experimentally. It's obvious that the information leakage will be deeply suppressed when L becomes large, which is the reason why a RRDPS experiment with large L is important.

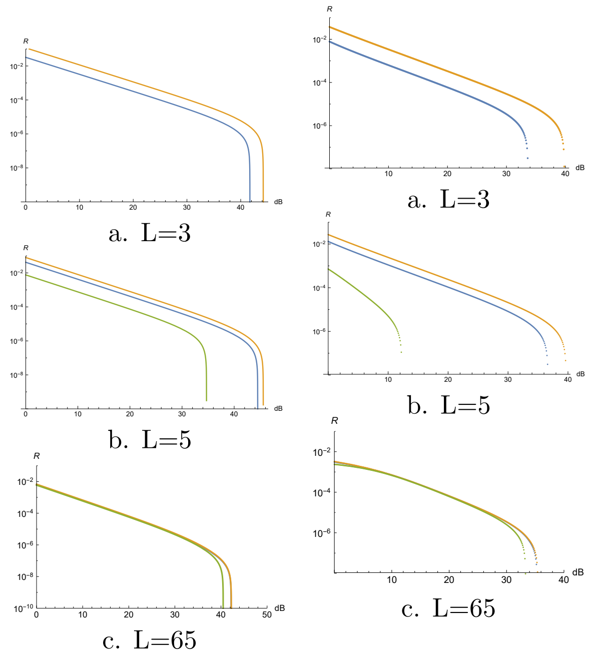
There have been several successful demonstrations of this protocol with passive interferometers [4, 5] and actively-selectable components [6, 7]. The longest achievable distance is around 90km [7]. Albeit great progresses on experiments of RRDPS protocol have been made, it's still a great challenge to realize a practical measurement system with large L value. Besides, large L value will decrease the secret key rate per pulse obviously. Therefore, it is highly desired if I_{AE} can be further lowered while L is maintained small. Additionally, although I_{AE} given in Ref.[3] does not depend on the error rate, theorists are still not clear how does Eve's attack introduce error bits, and if it is possible to use the error rate in RRDPS to improve its performance. To address these issues, we first report a new theory to bound I_{AE} greatly tighter than before especially for small L values. Interestingly, error rate can be also taken into account in our method to estimate I_{AE} further tightly. Through numerical simulation, we show that with our theory, the performance of real-life RRDPS implementation can be improved dramatically. It is remarkable that the most simple real-life RRDPS protocol with $L = 3$, which is not permitted in the original RRDPS protocol, can outperform the ones with very large L. Finally, we verify our theory through an experiment with $L = 3$, which achieves the longest achievable distance (140km) so far.

2 Results

In Tab.1, the maximum tolerant error rate for RRDPS with conventional method and the proposed formula are given. Our formula can increase the tolerance of error rate dramatically, especially when L is small. It's remarkable to note that for the case $L = 3$, our bound can tolerant E up to 8%, while original RRDPS protocol can not generate secure key bits at all. One may also note that the difference between these methods become little

[*] wshuang@ustc.edu.cn

[†] weich@ustc.edu.cn



a. L=3

b. L=5

c. L=65

a. L=3

b. L=5

c. L=65

(a) Fig.1. Secret key rate R (per pulse) versus channel loss under SPS.

(b) Fig.2. Secret key rate R (per pulse) versus channel loss under WCS.

in large L cases. The reason is quite simple, e.g., the original bound $h_2(1/(L-1))$ has been close to 0 for large L, so the potential improvement made by our analyses will be very little. However, the significance of our theory will not be compromised for real-life implementations, since we show that the RRDPS systems with small L values can outperform the ones with larger L values .

Table 1: The maximum value of tolerant error rate of RRDPS with different method.

method / L	original RRDPS	new without E	new with E
3	–	0.0546	0.0811
5	0.0289	0.122	0.144
65	0.302	0.347	0.347

We simulate the secret key rate R per pulse versus total loss for $L=3$, $L=5$ and $L=65$ using single photon source and weak coherent source respectively. The simulation results are shown in Fig.1 and Fig.2. In each figure, the upper line and middle line correspond to the proposed calculations of I_{AE} with and without error rate statistic, while the lower line accounts for the original RRDPS protocol. Note that in Fig.1 (a) and Fig.2 (a), there is no line for the original RRDPS protocol, since original protocol's key rate is 0 for $L=3$ case.

We tested the $L=3$ RRDPS system with standard telecom fiber channels at the distance of 50 km, 100 km, and 140 km. Thus, we have successfully verified the feasibility of RRDPS with the smallest $L=3$, which is impossible based on original theory.

3 Conclusion

In conclusion, we develop a theory to estimate Eve's information on raw key bits I_{AE} in a quite different way. Briefly speaking, the new physics behind our method is that the potential phase randomization can be utilized for the security analysis of RRDPS. The main merit of our method is that I_{AE} could be bounded more tightly than before, especially when L is small. In theory, the relation between the information leakage and error rate in RRDPS is present clearly, which is particularly meaningful for the completeness of security analysis of QKD. Our results pave an avenue towards practical RRDPS[8].

References

[1] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. in Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing (IEEE Press, New York, 1984) pp. 175179.

[2] A. K. Ekert. Quantum cryptography based on Bells theorem. Phys. Rev. Lett. 67, 661 (1991).

[3] T. Sasaki, Y. Yamamoto, and M. Koashi. Practical quantum key distribution protocol without monitoring signal disturbance. Nature 509, 475 (2014).

[4] J.-Y. Guan, et al. Experimental Passive Round-Robin Differential Phase-Shift Quantum Key Distribution. Phys. Rev. Lett. 114, 180502 (2015).

[5] H. Takesue, T. Sasaki, K. Tamaki, and M. Koashi. Experimental quantum key distribution without monitoring signal disturbance. Nat Photon 9, 827 (2015).

[6] S. Wang, et al. Experimental demonstration of a quantum key distribution without signal disturbance monitoring. Nat Photon 9, 832 (2015).

[7] Y.-H. Li, et al. Experimental round-robin differential phase-shift quantum key distribution. Phys. Rev. A 93, 030302 (2016).

[8] Z. Yin, et al. Practical round-robin-differential-phase-shift quantum key distribution with and without monitoring signal disturbance. Arxiv preprint arXiv:1702.01260 (2017).

Flow Ambiguity: A Path Towards Classically Driven Blind Quantum Computation

Atul Mantri[1 2 *] Tommaso F. Demarie[1 2 †] Nicolas C. Menicucci[3 4 ‡]
Joseph F. Fitzsimons[1 2 §]

[1] *Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372*
[2] *Centre for Quantum Technologies, National University of Singapore, Block S15, 3 Science Drive 2, Singapore 117543*
[3] *School of Science, RMIT University, Melbourne, Victoria 3001, Australia*
[4] *School of Physics, The University of Sydney, Sydney, New South Wales 2006, Australia*

Abstract. Blind quantum computation protocols allow a user to delegate a computation to a remote quantum computer in such a way that the privacy of their computation is preserved, even from the device implementing the computation. To date, such protocols are only known for settings involving at least two quantum devices: either a user with some quantum capabilities and a remote quantum server or two or more entangled but noncommunicating servers. In this work, we take the first step towards the construction of a blind quantum computing protocol with a completely classical client and single quantum server. Specifically, we show how a classical client can exploit the ambiguity in the flow of information in measurement-based quantum computing to construct a protocol for hiding critical aspects of a computation delegated to a remote quantum computer.

Keywords: blind quantum computation, secure delegated quantum computation, measurement-based quantum computation, generalized information flow

It is very likely that when a universal quantum computer will finally become available, it will be hosted by large institutions and accessed remotely by clients. For example, companies like D-Wave [1] and IBM [2], as well as academic institutions including the University of Bristol [3], have begun making their quantum devices available for remote access. This situation will inevitably lead to questions as to the integrity and privacy of the client's computation. In the past, quantum protocols have been proposed to address similar problems. Protocols which provide security of client's quantum computation, as well as input and output, are known as blind quantum computing protocols [4, 5, 6, 7]. Similarly, protocols which capture the idea of verification of quantum computing, i.e., the ability to detect, with very high probability, any attempt by a malicious server to deviate from the computation are known as verifiable quantum computing protocols [8, 9, 10]. A common feature among known protocols for these tasks is that either the client require a small quantum device on their side or there must exist at least two non-communicating quantum servers [11, 12]. In other words, there is a requirement that two or more parties involved in the protocol possess quantum processors. Ideally, we would like to have a secure delegated quantum computing protocol between a completely classical client and a quantum server. In this work, we take the first steps towards this problem. We construct a blind quantum computing protocol which maintains security of the client's computation even against the quantum server. In the next section, we briefly describe our main ideas and results, but the full details can be found in [13].

Main Ideas and Results

Our aim is to explore the possibility of blind quantum computation with a purely classical client. We demonstrate this fact by constructing a protocol for a task we call as classically driven blind quantum computing (CDBQC) and analyse its security in the stand-alone setting. Our protocol uses measurement-based quantum computing (MBQC) [14] as the underlying principle. We show that the protocol allows a client to hide non-trivial information about their computation from the powerful quantum server by making use of a novel technique that we call flow ambiguity. In particular, we analyse the case of a single instance of the protocol and show that the amount of information obtained by the server is bounded below what is necessary to unambiguously distinguish the computation.

In the MBQC framework we denote by Δ the computation of the client such that $\Delta = \{\mathcal{G}, \boldsymbol{\alpha}, \boldsymbol{f}\}$. Here $\mathcal{G}$ denotes the graph state, $\boldsymbol{\alpha}$ is the set of measurement angles on the graph state, and $\boldsymbol{f}$ represents the information flow [15] which captures how angles are to be adapted based on results of previous measurements. Formally, generalised information flow or g-flow [16] is defined as follows: For an open graph $\mathcal{G}(I, O)$, there exists a *g-flow* $(g, \succ)$ if one can define a function $g : O^c \to P(I^c)$ and a partial order $\succ$ on $\mathcal{V}$ such that $\forall i \in O^c$, all of the following conditions hold:

(G1) if $j \in g(i)$ and $j \neq i$, then $j \succ i$;
(G2) if $j \not\succ i$ and $i \neq j$, then $j \notin \mathrm{Odd}(g(i))$; and
(G3) $i \notin g(i)$ and $i \in \mathrm{Odd}(g(i))$.

Intuitively, g-flow is used to assign a set of local corrections to a subset of unmeasured qubits to ensure de-

[*] atul_mantri@mymail.sutd.edu.sg
[†] tommaso_demarie@sutd.edu.sg
[‡] ncmenicucci@gmail.com
[§] joseph_fitzsimons@sutd.edu.sg

terministic computation, despite the random nature of the measurement outcomes obtained during the computation. It is important to note that for a fixed graph there exist multiple choices of the input and output vertex sets that result in deterministic measurement patterns consistent with the same fixed total ordering of vertices. Specifically, we show that the transcript of any run of the protocol is consistent with multiple non-equivalent computations. This is due to the fact that the information about the g-flow for the underlying resource state is hidden from the server. This particular ambiguity in the flow enables the classical user to hide the essential aspects of the computation.

The CDBQC protocol is interactive and proceeds as follows. Firstly, the client sends the dimension of the graph to the server to prepare the graph state $|G\rangle$. At each step i: Client chooses a bit r_i uniformly random. Using r_i and the previous measurement outcome $b_{<i'}$, client updates the angle α_i to construct α_i' in the following way:

$$\alpha' = (-1)^{s^x}\alpha + (r \oplus s^z)\pi$$

where s^x and s^z denote the corrections dictated by flow based on previous measurement results. The server performs a projective measurement of i-th vertex in the XY-plane of the Bloch sphere, denoted $M_i^{\alpha_i'} = \{|\pm_{\alpha_i'}\rangle\langle\pm_{\alpha_i'}|\}$, where $\{|\pm_{\alpha'}\rangle\} = \frac{1}{\sqrt{2}}(|0\rangle \pm e^{i\alpha'}|1\rangle)$ and sends the measurement outcome b_i' to the client. The client records $b_i = b_i' \oplus r_i$ in $\boldsymbol{b}$ and then updates the set $(\boldsymbol{s}^x, \boldsymbol{s}^z)$. If the i-th vertex is output qubit then the bit b_i is registered in the set $\boldsymbol{p}_B^C$. The client and the server repeat this procedure for all the vertex of the graph in the given total order. The client implements the final round of corrections on the string $\boldsymbol{p}^C$ (equivalent to $\boldsymbol{p}_B^C$ in the case of honest server) to obtain the output string $\boldsymbol{p}$. At the end of the protocol, the server possesses information about the angles $\boldsymbol{\alpha'}$ and the measurement outcome $\boldsymbol{b'}$ and whereas the client's secret consists of the actual measurement angle $\boldsymbol{\alpha}$ and the flow bits $\boldsymbol{f}$. Hereafter we will denote the variables with the upper-case letters and particular instances of such variable with the lower-case letters. For example, $\boldsymbol{A}$ will be used to denote the angle variable and $\boldsymbol{F}$ is used to represent the flow variable. For simplicity, let's take the case when $\boldsymbol{A}$ and $\boldsymbol{F}$ are uniformly variables. We quantify the amount of information that on average remains hidden from the server about the client's computation at the end of the protocol. This is given by the conditional entropy $H(\boldsymbol{A}, \boldsymbol{F}|\boldsymbol{B'}, \boldsymbol{A'})$.

$$H(\boldsymbol{A}, \boldsymbol{F}|\boldsymbol{B'}, \boldsymbol{A'}) = H(\boldsymbol{A}, \boldsymbol{F}) - I(\boldsymbol{B'}, \boldsymbol{A'}; \boldsymbol{A}, \boldsymbol{F}). \quad (1)$$

where $H(\boldsymbol{A}, \boldsymbol{F}) = H(\boldsymbol{A}) + H(\boldsymbol{F}) := \log_2 N_{\boldsymbol{A}} + \log_2 N_{\boldsymbol{F}}$. Here $N_{\boldsymbol{A}}$ and $N_{\boldsymbol{F}}$ denote the number of possible choices for the angle and flow variable respectively. Using tools from information theory we explicitly calculate that, in a single run of CDBQC protocol, the mutual information between the client's secret input ($\boldsymbol{\alpha}$, $\boldsymbol{f}$) and the information received by the server ($\boldsymbol{\alpha'}$, $\boldsymbol{b'}$) is bounded by

$$I(\boldsymbol{B'}, \boldsymbol{A'}; \boldsymbol{A}, \boldsymbol{F}) \leq H(\boldsymbol{A'})$$

This in turn gives a lower bound on the conditional entropy

$$H(\boldsymbol{A}, \boldsymbol{F}|\boldsymbol{B'}, \boldsymbol{A'}) \geq \log_2 N_{\boldsymbol{F}} \quad (2)$$

We derive a non-trivial lower bound on the conditional entropy by calculating the value of $N_{\boldsymbol{F}}$. Note that flow is a property of the underlying graph and therefore depends on the chosen graph $\mathcal{G}$. We will consider the case of cluster states as they are known to be universal for quantum computation with (X, Y)-plane measurements [17]. To calculate $N_{\boldsymbol{F}}$ for the cluster state, we put a lower bound on the number of different input and output choices (open graphs) $\#\mathcal{G}(I, O)_{n,m}$ satisfying flow conditions for a cluster state and a certain fixed total order. Mathematically, this corresponds to calculating the flows that satisfy the conditions (G1)-(G3) as mentioned above. To simplify the counting argument we put an additional constraint:

(G4) If $k \in \mathcal{N}(i) \cup \mathcal{N}(j)$, and if $k \in g(i)$, then $k \notin g(j)$.

This is not required strictly by the definition of g-flow, but it simplifies the flow counting problem and so we obtain a lower bound on the number of flows rather than the exact number. For a generic cluster state $\mathcal{G}_{(n,m)}$ with the fixed total ordering of measurements, the number of different open graphs $G(I, O)$ satisfying the conditions (G1)- (G4) is given by:

$$\#\mathcal{G}(I, O)_{n,m} = F_{2\min(n,m)+1}^{|n-m|} \prod_{\mu=2}^{\min(n,m)} F_{2\mu}^2. \quad (3)$$

where F_i is the ith Fibonacci number. Further simplifying the above equation gives us $\#\mathcal{G}(I, O)_{n,m} = 2^{2N\log_2\phi + O(N^\epsilon)}$ for $\epsilon < 1$, $N = nm$ and assuming $m = \text{poly}(n)$. This shows that there exists at least an exponential number (in the dimension of the graph) of information flows corresponding to a cluster state for a given total order of measurements. To demonstrate this we take a simple example of 2×2 cluster state $G(I, O)_{(2\times2)}$ in Fig. 1. The figure shows 9 possible open graphs compatible with the flow conditions (G1)-(G4). In general different flows correspond to different computations.

Using the following relation $N_{\boldsymbol{F}} \geq \#\mathcal{G}(I, O)_{n,m}$ with the the above result, we get $\log_2 N_{\boldsymbol{F}} \geq \log_2 \#\mathcal{G}(I, O)_{n,m} \approx 1.388N$. Therefore, the conditional entropy is given by

$$H(\boldsymbol{A}, \boldsymbol{F}|\boldsymbol{B'}, \boldsymbol{A'}) \geq 1.388N. \quad (4)$$

This shows that it is indeed possible for a client to hide their chosen computation, by using the ambiguity in the flow of information, from a quantum server. Importantly, we show that it is not possible for the quantum server to guess the client's computation perfectly, since a large number of other computations are still compatible with the information server receives.

For more details, we refer to the published version of this work [13] and references therein.

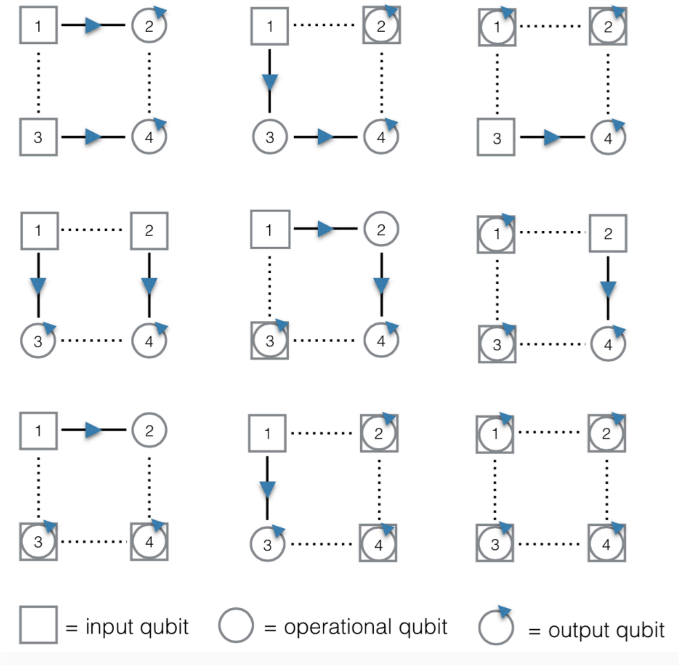


Figure 1: All the possible $\mathcal{G}(I, O)_{2,2}$ combinations that satisfy g-flow conditions for the 2×2 cluster state are shown. The arrows indicate the direction of the quantum information flow. Note that overlapping input and output sets are allowed. All patterns implement unitary embeddings on the input state.

References

[1] D-Wave. `https://www.dwavesys.com/`

[2] IBM. The quantum experience. `http://www.research.ibm.com/quantum/`

[3] Quantum in the Cloud. `http://www.bristol.ac.uk/physics/research/quantum/engagement/qcloud/`

[4] Childs, A. M. (2005). Secure assisted quantum computation. Quantum Information & Computation, 5(6), 456-466.

[5] Arrighi, P., & Salvail, L. (2006). Blind quantum computation. International Journal of Quantum Information, 4(05), 883-898.

[6] Broadbent, A., Fitzsimons, J., & Kashefi, E. (2009, October). Universal blind quantum computation. In Foundations of Computer Science, 2009. FOCS'09. 50th Annual IEEE Symposium on (pp. 517-526). IEEE.

[7] Morimae, T., & Fujii, K. (2013). Blind quantum computation protocol in which Alice only makes measurements. Physical Review A, 87(5), 050301.

[8] Fitzsimons, J. F., & Kashefi, E. (2017). Unconditionally verifiable blind quantum computation. Physical Review A, 96(1), 012303.

[9] Broadbent, A. (2015). How to verify a quantum computation. arXiv preprint arXiv:1509.09180.

[10] Hayashi, M., & Morimae, T. (2015). Verifiable measurement-only blind quantum computing with stabilizer testing. Physical review letters, 115(22), 220502.

[11] Reichardt, B. W., Unger, F., & Vazirani, U. (2013). Classical command of quantum systems. Nature, 496(7446), 456-460.

[12] Fitzsimons, J. F. (2017). Private quantum computation: an introduction to blind quantum computing and related protocols. npj Quantum Information, 3(1), 23.

[13] Mantri, A., Demarie, T. F., Menicucci, N. C., & Fitzsimons, J. F. (2017). Flow ambiguity: A path towards classically driven blind quantum computation. Physical Review X, 7(3), 031004.

[14] Raussendorf, R., & Briegel, H. J. (2001). A one-way quantum computer. Physical Review Letters, 86(22), 5188.

[15] Danos, V., & Kashefi, E. (2006). Determinism in the one-way model. Physical Review A, 74(5), 052310.

[16] Browne, D. E., Kashefi, E., Mhalla, M., & Perdrix, S. (2007). Generalized flow and determinism in measurement-based quantum computation. New Journal of Physics, 9(8), 250.

[17] Mantri, A., Demarie, T. F., & Fitzsimons, J. F. (2017). Universality of quantum computation with cluster states and (X, Y)-plane measurements. Scientific Reports, 7.

A Cost-Effective Approach for Satellite Based Quantum Key Distribution

Alexander Lohrmann[1] * Aitor Villar[1] Debashis Demunshi[1] Zhongkan Tang[1]
Rakhitha Chandrasekara[1] Alexander Ling[1]

[1] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, S117543*

Abstract. Space-based quantum key distribution (QKD) is a widely discussed and greatly anticipated key technology for emerging quantum communication applications. However, high launch costs of conventional satellites prevent rapid progress in the development of space-based QKD. Here, we present a robust high-brightness source of entangled photons that can be employed on a small-scale nanosatellite to enable space based QKD.

Keywords: Space-based quantum key distribution, entangled-photon sources

A key technology for quantum experiments is quantum cryptography, and in particular QKD, which bridges the gap between fundamental tests of the concepts of quantum mechanics and potential applications in communication technology. QKD enables secure communication with forward secrecy based on the laws of quantum mechanics. While ground-based QKD systems are now in place in several institutions across the globe, their range on earth is fundamentally limited (absorption losses, line-of-sight).

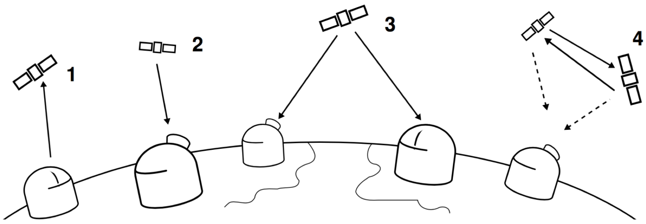


Figure 1: Different schemes for space-based QKD. (1) Ground-to-satellite, (2) satellite-to-ground, (3) entanglement based, (4) inter satellite. Illustration reproduced from [1].

To overcome this limitation, space-based QKD satellites have been proposed and first proof-of-concept satellites have been developed and deployed in the recent years [1, 2]. Fig. 1 shows several different concepts for space-based QKD. Our approach envisiones the generation of entangled photon pairs in orbit (3 in Fig. 1) to enable QKD over long distances via an entanglement-based protocol.

Despite its advantages, space-based QKD suffers from high costs and system complexities. In order to minimize the costs for the demonstration of space-based QKD, we are working on an iterative approach in which entangled photon sources are operated in a low-earth orbit aboard low-cost nanosatellites. This provides a cost-effective alternative to a full-scale QKD satellite.

At the heart of our technology lies the Small Photon Entangling Quantum System (SPEQS), a specifically designed, small entangled photon source based on spontaneous parametric down conversion (SPDC). SPDC is routinely exploited to create entangled photon pairs from a material with a strong second order non-linearity, such as β-Barium Borate (BBO) or Lithium Niobate. In short, in the presence of a non-linear crystal, a pump photon can split into two lower energy photons that are correlated regarding their polarization state (for example $|H_1\rangle\,|H_2\rangle$ or $|V_1\rangle\,|V_2\rangle$ in type-I phasematching depending on the crystal orientation). If SPDC photon pairs are created in two coherent processes, e.g. in crossed crystal or double pass configuration [3, 4], one can produce the entangled state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|H_1\rangle\,|H_2\rangle + e^{i\Phi}\,|V_1\rangle\,|V_2\rangle\right) \qquad (1)$$

The first generation of SPEQS that produced correlated photon pairs was successfully launched and its in-orbit operation was demonstrated [5]. The source design will be further improved to produce entangled photon pairs. We give a brief overview on the progress made on this source (SPEQS-1) which will be launched within a year.

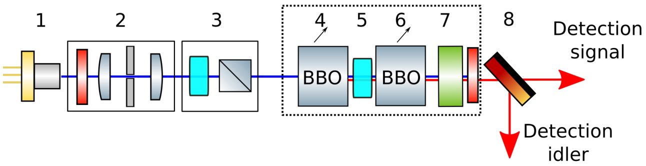


Figure 2: Schematic of the current design of our SPEQS-2 entangled-photon source. The components in the aligned crystal configuration are (1) 405 nm pump, (2) spatial and spectral filtering, (3) preparation of pump polarization, (4) BBO crystal, (5) quarter waveplate,(6) BBO crystal, (7) temporal compensator YVO4 and (8) dichroic mirror. The dotted rectangle highlights the SPDC generation and temporal compensation.

Our final goal is the development of a source (SPEQS-2) that provides entangled photon pairs with higher pair rates to overcome atmospheric turbulence and absorption losses [6] and enable space based QKD on a small and cost-effective platform. Here, we present the progress on our final source design (see Fig. 2). We use two BBO

*cqtal@nus.edu.sg

crystals with aligned optical axes in type-I phase matching condition to create non-degenerate entangled photon pairs via SPDC. Collinear emission of the SPDC photons enables a compact and robust source design. By compensating the temporal walkoff effects, we can prepare the entangled two-photon state in Eq. 1. First experiments performed with this source design show that an in-orbit rate of 1 million entangled photon pairs per second is within reach. In ground-based experiments, we achieve a pair rate of more than 70 k/s/mW at high visibility which is unprecedented for BBO based entangled pair sources.

With the employment of a compact, high brightness QKD source in space, we lay the foundation for pratical entanglement based QKD on a multitude of platforms, such as aircrafts, drones and fiber-coupled, ground-based systems.

References

[1] Bedington, R., Bai, X., Truong-Cao, E., Tan, Y.C., Durak, K., Zafra, A.V., Grieve, J.A., Oi, D.K. and Ling, A. Nanosatellite experiments to enable future space-based QKD missions. In EPJ Quantum Technology, 3(1), p.12, 2016.

[2] Vallone, G., Bacco, D., Dequal, D., Gaiarin, S., Luceri, V., Bianco, G. and Villoresi, P. Experimental satellite quantum communications. In Physical review letters, 115(4), p.040502, 2015.

[3] Trojek, P. and Weinfurter, H. Collinear source of polarization-entangled photon pairs at nondegenerate wavelengths. In Applied Physics Letters, 92(21), p.211103, 2008.

[4] Steinlechner, F., Ramelow, S., Jofre, M., Gilaberte, M., Jennewein, T., Torres, J.P., Mitchell, M.W. and Pruneri, V. Phase-stable source of polarization-entangled photons in a linear double-pass configuration. Optics express, 21(10), pp.11943-11951, 2013.

[5] Tang, Z., Chandrasekara, R., Tan, Y.C., Cheng, C., Sha, L., Hiang, G.C., Oi, D.K. and Ling, A. Generation and analysis of correlated pairs of photons aboard a nanosatellite. In Physical Review Applied, 5(5), p.054022, 2016.

[6] Bonato, C., Tomaello, A., Da Deppo, V., Naletto, G. and Villoresi, P. Feasibility of satellite quantum key distribution. In New Journal of Physics, 11(4), p.045017, 2009.

Lorentz invariant entanglement distribution for the space-based quantum network

Tim Byrnes[1 2 *] Batyr Ilyas[2 †] Louis Tessler[1 ‡] Masahiro Takeoka[3 §]
Segar Jambulingam[1 ¶] Jonathan Dowling[3 ‖]

[1] *New York University Shanghai, 1555 Century Ave, Pudong, Shanghai 200122, China*
[2] *State Key Laboratory of Precision Spectroscopy, School of Physical and Material Sciences, East China Normal University, Shanghai 200062, China*
[3] *NYU-ECNU Institute of Physics at NYU Shanghai, 3663 Zhongshan Road North, Shanghai 200062, China*
[4] *National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan*
[5] *Department of Physics, New York University, New York, NY 10003, USA*
[6] *Department of Physics, Nazarbayev University, 53 Kabanbay Batyr Ave., Astana 010000 Kazakhstan*
[7] *Department of Physics, Ramakrishna Mission Vivekananda College, Mylapore, Chennai 600004, India*
[8] *Hearne Institute for Theoretical Physics, Department of Physics & Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803-4001, USA*

Abstract. In recent years there has been a great deal of focus on a globe spanning quantum network, including linked satellites for applications ranging from quantum key distribution to distributed sensors and clocks. In many of these schemes, relativistic transformations may have deleterious effects on the purity of the distributed entangled pairs. This becomes particularly important for the application of distributed clocks. In this paper, we have developed a Lorentz invariant entanglement distribution protocol that completely removes the effects due to the relative motions of the satellites.

Keywords: Entanglement distribution, relativity, space, purification

One of the main roadblocks to the widespread utilization of quantum communication such as quantum cryptography is the difficulty of producing long-distance entanglement. Photons are a natural way of generating such entanglement due to their excellent coherence properties and the fact that they are "flying qubits". However optical fiber quantum communication is limited to distances of approximately $\sim$ 200 km due to photon loss, which make them practical for only for a limited region, not a global scale. Broadly two approaches have been considered to overcome this challenge – the use of quantum repeaters to cascade entanglement generation for longer distances [1, 2], and space-based schemes [3, 4, 5, 6]. Quantum communication in space is attractive due to the negligible effects of the atmosphere which is the origin of decoherence effects such as photon loss. This allows for the possibility of globe-scale quantum network where the photons can be transmitted at distances of the order of the diameter of the Earth without the need of additional infrastructure such as quantum repeaters.

In this paper, we investigate various strategies for space-based entanglement distribution using photons. We examine three popular alternatives for entanglement generation: (I) a polarization entangled photons; (II) single photon entangled state; and (III) dual rail entangled photons. The advantages and disadvantages of each will be investigated in the context of low Earth orbit (LEO) satellites producing and detecting the photons (see Fig. 1). The photonic states (II) and (III) are particularly interesting as they are based on RI quantities. It is known that photonic Fock states are RI quantities [10]. It is therefore natural to choose entangled states involving these degrees of freedom to develop a truly Lorentz invariant (LI) entanglement distribution. Choosing such manifestly LI states bypasses the need for any correction that would need to be made for states such as (I).

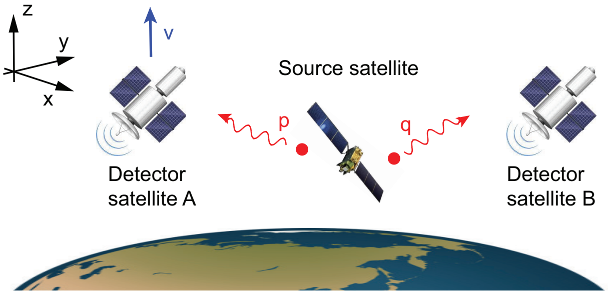


Figure 1: Entanglement distribution between three satellites in low Earth orbit. The photons heading to the two satellites may have different momenta $\boldsymbol{p}, \boldsymbol{q}$, due to their different directions. We choose Alice's satellite to be moving in the z-direction without loss of generality.

Let us first introduce the three types of entangled photon states that will be analyzed in this paper for creating long-distance entanglement using photons. The first is simply a polarization entangled photon pair, produced for example by parametric down conversion. The state

[*] tim.byrnes@nyu.edu
[†] batyr.ilyas@nu.edu.kz
[‡] IIAOPSW@gmail.com
[§] takeoka@nict.go.jp
[¶] segar@rkmvc.ac.in
[‖] jdowling@lsu.edu

is written

$$|\Psi_{\mathrm{I}}^{(S)}\rangle = \frac{1}{\sqrt{2}}\left(|\boldsymbol{p},h\rangle|\boldsymbol{q},h\rangle - |\boldsymbol{p},v\rangle|\boldsymbol{q},v\rangle\right), \tag{1}$$

where $|\boldsymbol{p},\sigma\rangle$ is a single photon state of four momentum $\boldsymbol{p}$ and polarization $\sigma = h, v$, and the S refers to the fact that the photons are in the reference frame of the source satellite. We label the modes for Alice and Bob's satellites with A and B respectively. The second type of entangled state is the single photon entangled state, which can be produced by a single photon source mounted on the source satellite entering a 50:50 beamsplitter. The state is

$$|\Psi_{\mathrm{II}}^{(S)}\rangle = \frac{1}{\sqrt{2}}(|\boldsymbol{p},\lambda\rangle_A|0\rangle_B - |0\rangle_B|\boldsymbol{q},\lambda\rangle_B). \tag{2}$$

where $\lambda = \pm 1$ labels the helicity, and $|0\rangle$ is the electromagnetic vacuum. Finally, the third type of entangled state is using a dual rail encoding, where Alice and Bob each posses two distinct modes $A1, A2$ and $B1, B2$ respectively, and the same helicity is used for both photons and modes:

$$\begin{aligned}|\Psi_{\mathrm{III}}^{(S)}\rangle = \frac{1}{\sqrt{2}}(&|0\rangle_{A1}|\boldsymbol{p},\lambda\rangle_{A2}|0\rangle_{B1}|\boldsymbol{q},\lambda\rangle_{B2}\\ &- |\boldsymbol{p},\lambda\rangle_{A1}|0\rangle_{A2}|\boldsymbol{q},\lambda\rangle_{B1}|0\rangle_{B2}).\end{aligned} \tag{3}$$

Each of these states will have a different behavior under a Lorentz transformation, and our task will be to identify which is the best for entanglement generation.

First, let us examine how single photon states transform. For a photon of helicity λ and momentum $\boldsymbol{p}$ in the Source frame, the state in Alice's frame is

$$U(\Lambda)|\boldsymbol{p},\lambda\rangle = e^{-i\lambda\Theta(\Lambda,\boldsymbol{p})}|\Lambda\boldsymbol{p},\lambda\rangle \tag{4}$$

where Θ is the Wigner phase, and Λ is the Lorentz transformation to the frame of A. Since we assume that the photon momentum is in an arbitrary direction, without loss of generality we may take the Lorentz transformation to be a pure boost in the z direction $\Lambda = L_z(\beta)$. In this case $L_z(\beta)$ is the standard Lorentz transformation matrix with dimensionless velocity $\beta = v/c$ (c is the speed of light). Polarized vectors in the original frame are defined as

$$\begin{aligned}|\boldsymbol{p},h\rangle &= R(\hat{\boldsymbol{p}})(0,\cos\phi,-\sin\phi,0)^T\\ |\boldsymbol{p},v\rangle &= R(\hat{\boldsymbol{p}})(0,\sin\phi,\cos\phi,0)^T\\ |\boldsymbol{p},\lambda\rangle &= R(\hat{\boldsymbol{p}})(0,1,i\lambda,0)^T/\sqrt{2}\end{aligned} \tag{5}$$

where the rotation matrix is $R(\hat{\boldsymbol{p}}) = R_z(\phi)R_y(\theta)$, with $R_{y,z}$ being the standard SO(3) rotation matrices, and $\hat{\boldsymbol{p}} = (\sin\theta\cos\phi, \sin\theta, \sin\phi, \cos\theta)$ is the normalized 3-momentum. For a pure boost in the z direction, the effect is to transform the coordinates as

$$\begin{aligned}\sin\theta \to \sin\theta' &= \frac{\sin\theta}{\sqrt{\sin^2\theta + \gamma^2(\cos\theta-\beta)^2}}\\ \phi \to \phi' &= \phi.\end{aligned} \tag{6}$$

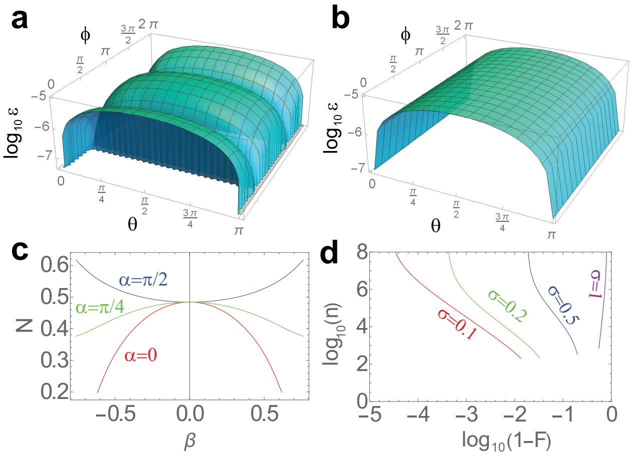


Figure 2: Performance of the entanglement distribution for various protocols. Trace distance ε between the original state and that observed in a moving frame for (a) a single horizontally (or vertically) polarized photon (b) a polarization entangled photon pair moving in opposite directions $\theta = \theta_A = -\theta_B$. Parameters are $\beta = 10^{-5}$. (c) Negativity of (10) under Lorentz boosts with different orientations. Photons are taken to move in opposite directions $\theta_A = -\theta_B$, $\phi_A = \phi_B$ and the spread due to the diffraction is $\sigma = 1$. (d) Number of entangled photon states (10) with $\sigma = 1$ required to reach purities as marked. We assume a photon attenuation factor of $\mathcal{A} = 100$, and the number of photons required for k purification steps to be 2^k.

To a good approximation, for $\beta \ll 1$ the variation in angle has the effect of

$$\theta' \approx \pi\left(\frac{\theta}{\pi}\right)^{1-\frac{2}{\pi\ln 2}\beta}. \tag{7}$$

This effectively broadens or contracts the angular variation around the z-axis. The angular variation is the origin of the variation in entanglement that was observed in works such as Ref.

To quantify the change we measure the trace distance of the polarization vector

$$\varepsilon = \mathrm{Tr}(\sqrt{(\rho^{(S)}-\rho^{(A)})^2})/2 \tag{8}$$

where $\rho^{(S)} = \mathrm{Tr}_{\boldsymbol{p}}(|\boldsymbol{p},\sigma\rangle\langle\boldsymbol{p},\sigma|)$ and $\rho^{(A)} = \mathrm{Tr}_{\boldsymbol{p}}(|\Lambda\boldsymbol{p},\sigma\rangle\langle\Lambda\boldsymbol{p},\sigma|)$ for this case. Here we trace over the momentum degrees of freedom in order to obtain a 4×4 matrix that is with respect to the polarization degrees of freedom. Fig. 2(a) shows the trace distance between a horizontally polarized photon with momentum $\boldsymbol{p}$ as observed by the source and Alice's satellite. For small velocities $\beta \ll 1$ as will be true for all satellites orbiting the Earth, expansion of the density matrices reveals that

$$\varepsilon_h \approx \beta\sin\theta\cos\phi, \tag{9}$$

which very accurately summarizes the numerical results in Fig. 2(a). We see that the basic effect of the relativistic correction on the polarization is at the level of $\varepsilon_h \sim O(\beta)$. We note that the trace distance is the most appropriate quantity (than the fidelity for instance which scales as $F \sim 1 - O(\beta^2)$), as it is most closely related to distances on the Bloch sphere.

Let us now examine the effect on the entangled states. For the type I entangled state, in Alice's frame we have

$$|\Psi_{\mathrm{I}}^{(A)}\rangle = \frac{1}{\sqrt{2}} \left(|\Lambda \boldsymbol{p}, h\rangle |\Lambda \boldsymbol{q}, h\rangle - |\Lambda \boldsymbol{p}, v\rangle |\Lambda \boldsymbol{q}, v\rangle \right). \quad (10)$$

The Wigner phase does not affect the state in this case as the state is transformed only by a pure Lorentz boost. The sole effect in terms of the trace distance is the rotation of the polarization vectors, as given in (6). The trace distance between the states in the Source and Alice's frames $\rho^{(S,A)} = \mathrm{Tr}_{\boldsymbol{p},\boldsymbol{q}}(|\Psi_{\mathrm{I}}^{(S,A)}\rangle\langle\Psi_{\mathrm{I}}^{(S,A)}|)$ is shown in Fig. 2(b).

$$\varepsilon_{\mathrm{I}} \approx \beta \sin\theta. \quad (11)$$

We again see that the relativistic correction again occurs at the level of $\sim O(\beta)$.

In this regard, the type II and III entangled states are a better choice. Fock states, including the vacuum, are known to be invariant states under Lorentz transforms, and remain orthogonal in all reference frames. For the single photon entangled states, transforming to the reference frame of satellite A, we find

$$\begin{aligned} |\Psi_{\mathrm{II}}^{(A)}\rangle =& \frac{1}{\sqrt{2}} (e^{-i\lambda\Theta(\Lambda,\boldsymbol{p})} |-\Lambda \boldsymbol{p}, \lambda\rangle_A |0\rangle_B \\ & - e^{-i\lambda\Theta(\Lambda,\boldsymbol{q})} |0\rangle_A |\Lambda \boldsymbol{q}, \lambda\rangle_B). \quad (12) \\ |\Psi_{\mathrm{III}}^{(A)}\rangle =& e^{-i\lambda(\Theta(\Lambda,\boldsymbol{p})+\Theta(\Lambda,\boldsymbol{q}))} (|0\rangle_{A1} |\mathbf{p}, \lambda\rangle_{A2} |0\rangle_{B1} |\mathbf{q}, \lambda\rangle_{B2} \\ & - |\mathbf{p}, \lambda\rangle_{A1} |0\rangle_{A2} |\mathbf{q}, \lambda\rangle_{B1} |0\rangle_{B2}), \quad (13) \end{aligned}$$

Similarly to type I, the photons are Lorentz transformed and there are separate Wigner phase terms due to the photon traveling with different momenta. Helicity is a Lorentz invariant quantity. The Wigner phase, which depends on the Lorentz transformation and the momentum does not show up in the measure we calculate. In both these cases, the entanglement is present in the photon number, rather than polarization. We thus define the density matrices for these states according to

$$\rho = \mathrm{Tr}_{\boldsymbol{p},\boldsymbol{q},\lambda}(|\Psi\rangle\langle\Psi|). \quad (14)$$

The trace distance between $\rho^{(S)}$ and $\rho^{(A)}$ is always zero, hence it is a manifestly LI state.

To take into account of diffraction, we integrate with a momentum distribution [11]

$$|\tilde{\Psi}\rangle = \int \tilde{d\boldsymbol{p}} \tilde{d\boldsymbol{q}} f_A(\boldsymbol{p}) f_B(\boldsymbol{q}) |\Psi(\boldsymbol{p}, \boldsymbol{q})\rangle \quad (15)$$

where the $|\Psi(\boldsymbol{p}, \boldsymbol{q})\rangle$ are the states (1), (2), (3) in the source satellite's frame. Here $\tilde{d\boldsymbol{p}} \equiv \frac{d^3\boldsymbol{p}}{2|\boldsymbol{p}|}$ is a Lorentz-invariant momentum integration measure and the $f(\mathbf{p})$ is a normalized diffraction function.

$$f(\boldsymbol{p}) = \frac{1}{\sqrt{M}} e^{-\frac{\theta^2}{2\sigma^2}} \delta(|\boldsymbol{p}| - p_0). \quad (16)$$

This gives a Gaussian spread for a photon traveling in primarily the z-direction. σ is a parameter controlling the angular spread of the beam and M is a suitable normalization factor. To have photons traveling in directions other than the z-direction, we make rotation of the coordinates around the y-axis by changing variables in the integrand

$$\begin{aligned} \theta \to \theta'' &= \cos^{-1}(\cos\alpha\cos\theta + \sin\alpha\sin\theta\cos\phi) \\ \phi \to \phi'' &= \tan^{-1}\left(\frac{\sin\theta\sin\phi}{\cos\alpha\sin\theta\cos\phi - \sin\alpha\cos\theta} \right) \end{aligned} \quad (17)$$

which gives photons traveling in primarily the direction $(\theta, \phi) = (\alpha, 0)$. To transform to Alice's frame, one then applies a boost in the z-direction to the states, which amounts to making the transformation (6).

We now estimate the order to which the relativistic corrections affect the entanglement. To gauge this we calculate the effect of the boost on the purity of the states $P = \mathrm{Tr}\rho^2$. The purity is directly related to the entanglement in this case as for the case with no diffraction, the entanglement is invariant under all boosts. The degradation in the entanglement observed in Fig. 2(c) arises from an effective decoherence entering the system due to tracing out the momentum degrees of freedom. Performing an expansion for $\beta \ll 1$ we find that the purity behaves as

$$P \approx 1 - 2\sigma^2(1 + |\beta|)^2. \quad (18)$$

Turning to type II and III states, we find that the effects of diffraction that afflicted type I states are not present in terms of entanglement degradation. The reason is that the type of entanglement is encoded in the orthogonality of the Fock states, which are preserved as they are relativistically invariant. For the type II state the main effect that one must account for is simply photon loss, which is captured by the photon attenuation $\mathcal{A}$ which is the same as the above. For the dual rail type III states, there is however the issue that the diffraction cone for the two rails will start to overlap unless they are separated by a sufficiently large distance, which is impractical for satellite based sources and detectors.

In summary, we have analyzed several photon based entanglement distribution protocols for the space-based quantum network. We find that standard polarization based photon entanglement (type I) can experience significant errors for satellites that are in low Earth orbit. While in principle these are correctable if the velocities of the satellites are known to high precision, this can still introduce errors at the $\delta\beta$, which is the error on the estimate of the satellite velocity.We note that other types of encodings, such as in energy or time, would also undergo Lorentz transformations. Combined with the fact that diffraction effects degrade the entanglement for type I states, our results point to the fact that single photon entangled states (type II) and dual rail photon entanglement (type III) are a superior choice in terms of robustness to relativistic transformations.

This work is supported by the Shanghai Research Challenge Fund, New York University Global Seed Grants for Collaborative Research, NYU-ECNU Institute of Physics at NYU Shanghai, National Natural Science Foundation of China (Grant No. 61571301), the Thousand Talents Program for Distinguished Young Scholars (Grant No. D1210036A), and the NSFC Research Fund for International Young Scientists (Grant No. 11650110425).

References

[1] Gisin, Nicolas and Ribordy, Grégoire and Tittel, Wolfgang and Zbinden, Hugo Quantum cryptography Reviews of Modern Physics vol 74 number 1 page 145, 2002

[2] Briegel, H.-J. and Dür, W. and Cirac, J. I. and Zoller, P. Quantum Repeaters: The Role of Imperfect Local Operations in Quantum Communication Phys. Rev. Lett. vol. 81 issue 26 pages 5932-5935, 1998

[3] Aspelmeyer, M. and Jennewein, Thomas. and Pfennigbauer, M. and Leeb, W. and Zeilinger, A. Long-Distance Quantum Communication with Entangled Photons using Satellites IEEE Journal of Selected Topics in Quantum Electronics, 2003

[4] Kaltenbaek, Rainer and Aspelmeyer, Markus and Jennewein, Thomas and Brukner, Caslav and Zeilinger, Anton and Pfennigbauer, Martin and Leeb, Walter R Proof-of-concept experiments for quantum physics in space Optical Science and Technology, SPIE's 48th Annual Meeting pages 252-268, 2004

[5] Ursin, Rupert and Tiefenbacher, F and Schmitt-Manderbach, T and Weier, H and Scheidl, Thomas and Lindenthal, M and Blauensteiner, B and Jennewein, T and Perdigues, J and Trojek, P and others Entanglement-based quantum communication over 144 km Nature physics vol 3 number 7 pages 481-486, 2007

[6] Yin, Juan and Cao, Yuan and Liu, Shu-Bin and Pan, Ge-Sheng and Wang, Jin-Hong and Yang, Tao and Zhang, Zhong-Ping and Yang, Fu-Min and Chen, Yu-Ao and Peng, Cheng-Zhi and others Experimental quasi-single-photon transmission from satellite to earth optics express volume 21 number 17 pages 20032-20040, 2013

[7] Ekert, Artur K. Quantum cryptography based on Bell's theorem Phys. Rev. Lett.vol 67 pages 661-663 August 1991

[8] Jae-Weon Lee and Eok Kyun Lee and Yong Wook Chung and Hai-Woong Lee and Jaewan Kim Quantum cryptography using sinlge-particle entanglement Phys Rev A vol 68 2003

[9] Paul M Alsing and Gerard G Milburn Lorentz invariance of entanglement arXiv:quant-ph/0203051 2002

[10] J.E. Avron and E. Berg and D. Goldsmith and A. Gordon Is the number of photons is a classical invariant Eur. J. Phys 1999 vol 20 pages 153-159

[11] Gingrich, Robert M and Bergou, Attila J and Adami, Christoph Entangled light in moving frames Physical Review A vol 68 number 4, 2003

[12] Bednorz, Adam Relativistic invariance of the vacuum The European Physical Journal C vol 73 number 12 pages 1-14, 2013

[13] Richard Jozsa and Daniel S. Abrams and Jonathan P. Dowling and Colin P. Williams Quantum Clock Synchronization Based on Shared Prior Entanglement Physical Review Letters vol 85, 2000

[14] Artur K. Ekert Quantum cryptography based on Bells theorem Physical Review Letters vol 67 number 661, 1991

[15] Louis Tessler and Jonathan Dowling and Tim Byrnes (in preparation)

Efficient classical verification of quantum computations*

Richard Jozsa[1] Sergii Strelchuk[1]

[1] *DAMTP, Centre for Mathematical Sciences, University of Cambridge, Cambridge CB3 0WA, U.K.*

Abstract. We propose an efficient scheme for verifying quantum computations in the 'high complexity' regime i.e. beyond the remit of classical computers. Previously proposed schemes remarkably provide confidence against arbitrarily malicious adversarial behaviour in the misfunctioning of the computer. Our scheme is not secure against arbitrarily adversarial behaviour, but may nevertheless be sufficiently acceptable in many practical situations. In contrast to previous schemes, our verifier is entirely classical. It is based on the fact that adaptive Clifford circuits on general product state inputs provide universal quantum computation, while the same processes without adaptation are always classically efficiently simulatable.

Keywords: verification, quantum computation

Establishing confidence in the output of a quantum computing device operating in the 'high complexity' regime i.e. beyond the remit of classical computers, will be an important issue as our first quantum computers become available. Some previously proposed schemes for such verification [1, 2, 3, 4] have been based on adaptations of the formalism of interactive proof systems (IP) from classical complexity theory [5], while others [12, 13, 14] have been based on the formalism of measurement based computing (MQC). Verification has also played an important role in the cognate subjects of blind quantum computing [10, 9], and self testing and device independent protocols [15, 12, 11].

Most of the previously proposed schemes have been designed to provide confidence in correctness of the output against the most general prospective malicious or adversarial behaviour in the misfunctioning of the computer. But although a stance of such extreme guardedness may be appropriate and relevant in fields such as cryptography, information security and financial transaction, it is not normally adopted in standard scientific method, and it comes at a considerable cost. All previous schemes (except some [12, 11] having multiple provers, so not directly relevant for our setting) require the verifier to have a quantum communication channel to the prover and some quantum processing capability (to the extent of the verifier having the ability to perform individual computational steps that would suffice for universal quantum computing). As such, these schemes, while being significant developments in the theory of verification, are perhaps less relevant for issues of realistically efficient verification capability *per se.*

In contrast to the above, in our scheme the verifier is entirely classical, using only polynomially-bounded classical computing resources and only classical communication with the prover. However there are associated limitations too: the confidence in correctness of the output will not be secure against arbitrarily malicious malfunctioning of the computer, but nevertheless we would expect that it could be acceptable in many realistic situations, to an extent that's not dissimilar from commonly accepted scientific practice. Indeed with this concession we gain in manifest simplicity and transparency.

Our scheme will be based on the fact that adaptive Clifford computations (on product state inputs that can include 1-qubit magic states) provide universal quantum computation [7], whereas non-adaptive Clifford computations with the same inputs are always classically efficiently simulatable [6]. Intuitively, this will enable us to reduce the the running of a universal quantum computer to a classically simulatable process, *after* the machine has completed its run, and hence, with further testing runs of the same sequence of quantum operations that occurred in the initial run, we can efficiently verify that the machine was able to correctly run that sequence.

The verification scheme. Suppose we have available a quantum computing device that can allegedly perform Clifford gates and computational basis measurements, and we can also reliably prepare the computational basis states $|0\rangle$ and $|1\rangle$ as well as the magic state $|A\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\pi/4}|1\rangle)$. We also assume that the classical actions of hardware choice involved in adaptation in response to the classical outcomes of intermediate measurements, can be reliably performed.

It is well known [7] that the above resources provide universal quantum computation. Indeed in addition to the Clifford gates it suffices to be able to implement the gate $T = \mathrm{diag}(1\,,\, \mathrm{e}^{\mathrm{i}\pi/4})$, which can be achieved by an adaptive Clifford process called "the T-gadget", consuming one copy of $|A\rangle$ as follows: to apply T to line k using $|A\rangle$ on line a, first apply Clifford CX_{ka} to those two lines, measure line a to obtain measurement outcome m, and then apply the Clifford gate $S^m = \mathrm{diag}(1\,,\,\mathrm{i})^{\mathrm{m}}$ to line k. This results in T being applied on line k (up to an irrelevant overall phase).

In view of all the above, suppose now that we have solved an instance of a BQP decision task by implementing an adaptive Clifford circuit on our quantum computing device, to obtain output 0 or 1. By further efficient use of the machine and efficient classical computation, we wish to develop confidence in the correctness of the quantum operation of the machine and its classical output for the adaptive choices in the computational run that occurred.

*Full arXiv version available at https://arxiv.org/abs/1705.02817

Having run our adaptive Clifford circuit we record its output as well as the sequence of (adaptively chosen) gates that were actually applied. We will refer to this particular sequence as the 'computational run'. To develop confidence in the correctness of its output we next run, polynomially many times, the same sequence of gates non-adaptively (including here also the intermediate measurement operations but ignoring their outcomes). We refer to these runs as 'gate test runs'. From the frequencies of their outputs, we obtain an estimate of the output probabilities of this non-adaptive process to within $1/\text{poly}(n)$ additive error (with probability exponentially close to 1, by the Chernoff bound cf Appendix of [8]). Now this non-adaptive process of Clifford gates and measurements, the same sequence of operations that actually occurred in the computational run, is classically efficiently simulatable [6]. We classically compute its output probabilities and compare them to the experimentally obtained values, verifying that the single actually implemented sequence of operations used in the computational run, provided a true sample of its output distribution to within $1/\text{poly}(n)$ additive error in the probabilities (assuming that the computing device behaves in the same way for repeated trials of a given process).

The non-adaptive process in the gate test runs differs from the computational run only in that the intermediate measurement results are uniformly randomly varying and unlikely to reproduce those that occurred in the computational run itself. Nevertheless we still develop confidence in the computing device's ability to correctly implement the same sequence of operations that occurred in the computational run, albeit in a slightly different scenario. Indeed since the measured qubit is always disjoint from rest of the computer and never used again, the measurement operation on it can have no effect on the reduced state of the rest of the computer (by the no-signalling principle). As such, the unitary gates of the (adaptive) computational run acted on the same (reduced) input states that occur in the (now classically simulatable, non-adaptive) gate test runs. Also the computer cannot physically function differently for different intermediate measurement results (assuming suitable non-communication in implementation of local operations, and having the measured qubits reasonably isolated from the other qubits). So although the gate test runs produce generally different intermediate measurement outputs, they may nevertheless still be viewed as providing evidence for validity of the implementation of the quantum operations that actually occurred in the computational run.

In the adaptive process, it is important that the different gate sequences (assumed now to be sufficiently faithfully implemented themselves) are chosen with their correct respective probabilities, for a single run of the adaptive process to represent a valid sampling of the desired BQP problem's solution. We can develop further confidence in the device's correctness of its operation in this respect as follows. We consider the initial part of the circuit up to the first T-gadget and run it polynomially many times to estimate the measurement probability to within a $1/\text{poly}(n)$ additive error, and verify that it is within (say) $O(1/t^2)$ of the value half, where $t = O(\text{poly}(n))$ is the number of T-gadgets in the circuit. Being adequately satisfied with the first T-gadget's measurement operation, we apply the same process to the second T-gadget, while now treating the first T-gadget's measurement output non-adaptively and using the gate sequence from the computational run up to the second T-gadget. Similarly we work through all the T-gadgets in order. We refer to such test runs as 'measurement test runs'.

The theoretical probability of any adaptive gate sequence is $1/2^t$ and we have developed confidence that the computing device has selected the gate sequence used in the computational run with probability $\pi = \left(\frac{1}{2} + O(\frac{1}{t^2})\right)^t = \frac{1}{2^t}\left(1 + O(\frac{1}{t^2})\right)^t$. Since $(1 + \frac{1}{m^2})^m \to 1$ as $m \to \infty$ we see that we can thus (with polynomially bounded computing effort) confirm that $|\pi - \frac{1}{2^t}| < \frac{\epsilon}{2^t}$ for any chosen constant $\epsilon > 0$ and all sufficiently large t.

Finally let p_{out} be the true theoretical probability of output 0 in the adaptive Clifford process, and let $p_{out}^{(j)}$, for $j = 1, \ldots, 2^t$, be the corresponding output probability for the j^{th} adapted gate sequence. Also let $\pi^{(j)}$ be the true theoretical probability that the j^{th} gate sequence occurs in an adaptive run. (In fact $\pi^{(j)} = 1/2^t$ here). Then $p_{out} = \sum_{j=1}^{2^t} p_{out}^{(j)} \pi^{(j)}$. Let j_0 be the label of the adaptive sequence that was actually used in the computational run. With polynomially bounded quantum and classical computational resources we have developed confidence that:
(i) (from gate test runs) the output probability $\tilde{p}_{out}^{(j_0)}$ of the implemented computational run is within additive error $\eta = 1/\text{poly}(n)$ of its theoretical value $p_{out}^{(j_0)}$, and
(ii) (from measurement test runs) the gate sequence labelled j_0 has been chosen by the device with probability $\tilde{\pi}^{(j_0)}$ that is within $\epsilon/2^t$ of its theoretical value $\pi^{(j_0)} = 1/2^t$.
(i) and (ii) then imply that (with suitably chosen η and ϵ) our quantum computing device has provided a sample of a probability distribution that is within any desired $\epsilon' > 0$ of the theoretical distribution $\{p_{out}, 1 - p_{out}\}$, so its output is then (within the requirements of the bounded error condition) the solution to our BQP decision problem e.g. we could assume without loss of generality that the BQP algorithm used has bounded error margin $|p_{out} - 1/2| > 0.49$, and choose ϵ and η to provide $\epsilon' < 0.01$, to then establish confidence that the machine's output is the correct answer to the decision problem with probability at least 0.98.

References

[1] D. Aharonov, M. Ben-Or and E. Eban. Interactive Proofs For Quantum Computations. Arxiv preprint arXiv:0810.5375, 2008.

[2] D. Aharonov and U. Vazirani, Is Quantum Mechanics Falsifiable? A computational perspective on the foundations of Quantum Mechanics. *Philosophy of*

Science anthology "Computability: Godel, Turing, Church, and beyond", Editors: Copeland, Posy, Shagrir, MIT press, 2012.

[3] A. Broadbent, How to Verify a Quantum Computation. Arxiv preprint arXiv:1509.09180, 2015.

[4] D. Aharonov, M. Ben-Or, E. Eban and U. Mahadev. Interactive Proofs For Quantum Computations. Arxiv preprint arXiv:1704.04487, 2017.

[5] S. Arora and B. Barak. Computational Complexity: A Modern Approach. Cambridge Univ. Press, 2009.

[6] R. Jozsa and M. Van den Nest. Classical simulation complexity of extended Clifford circuits. *Quant. Inf. Comp.* **14**, pp. 633-648, 2014.

[7] S. Bravyi and A. Kitaev. Universal Quantum computation with ideal Clifford gates and noisy ancillas. *Phys. Rev.* **A 71**, 022316, 2005.

[8] M. Van den Nest. Simulating quantum computers with probabilistic methods. *Quant. Inf. Comp.* **11**, pp. 784-812, 2011.

[9] A. Broadbent, J. Fitzsimons and E. Kashefi. Universal blind quantum computation. In Proceedings of the 50th Annual Symposium on Foundations of Computer Science, FOCS '09, pp. 517-526. IEEE Computer Society, 2009.

[10] A. Childs. Secure assisted quantum computation. *Quant. Inf. Comp.*, **5**, pp. 456-466, 2005.

[11] B. Reichardt, F. Unger, and U. Vazirani. A classical leash for a quantum system: Command of quantum systems via rigidity of CHSH games. In *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science*, ITCS ?13, pages 321?322, ACM, New York, NY, USA, 2013.

[12] M. McKague. Interactive proofs for BQP via self-tested graph states. *Theory of Computing*, **12**, pp. 1-42, 2016.

[13] T. Morimae and J. Fitzsimons. Post hoc verification with a single prover, 2016. Arxiv preprint arXiv:1603.06046, 2016.

[14] M. Hayashi and M. Hajdusek. Self-guaranteed measurement-based quantum computation. Arxiv preprint arXiv:1603.02195, 2016.

[15] F. Magniez, D. Mayers, M. Mosca and H. Ollivier. Self-testing of Quantum Circuits. In: Bugliesi M., Preneel B., Sassone V., Wegener I. (eds) Automata, Languages and Programming. ICALP 2006. Lecture Notes in Computer Science, vol 4051. Springer, Berlin, Heidelberg, 2006.

Quantum Sphere-Packing Bounds with Polynomial Prefactors

Hao-Chung Cheng[1 2 *] Min-Hsiu Hsieh[2 †] Marco Tomamichel[2 3 ‡]

[1] *Graduate Institute Communication Engineering, National Taiwan University, Taiwan (R.O.C.)*
[2] *Centre for Quantum Computation & Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology Sydney, NSW 2007, Australia*
[3] *School of Physics, The University of Sydney, NSW 2006, Australia*

Abstract. We study lower bounds on the optimal error probability in channel coding at rates below capacity, commonly termed sphere-packing bounds. In this work, we establish a sphere-packing bound for classical-quantum channels, which significantly improves previous prefactor from the order of subexponential to polynomial. Furthermore, the gap between the obtained error exponent for constant composition codes and the best known classical random coding exponent vanishes in the order of $o(\log n/n)$, indicating our sphere-packing bound is almost exact in the high rate regime. The main technical contributions are two converse Hoeffding bounds for quantum hypothesis testing and the saddle-point properties of error exponent functions. Our complete paper can be found in arXiv:1704.05703 [quant-ph].

Shannon's noisy coding theorem [1] states that a message in an appropriately coded form can be reliably transmitted through a discrete memoryless channel $\mathscr{W}$, provided the coding rate R is below the channel capacity $C_{\mathscr{W}}$. More precisely, the probability of decoding errors can be made arbitrarily small as the coding blocklength grows. Later, Shannon himself pioneered the study of the exponential dependency of the optimal error probability $\epsilon^*(n,R)$ for a blocklength n and transmission rate R [2]. He defined the *reliability function* to be, for any fixed coding rate $R < C_{\mathscr{W}}$, $E(R) := \limsup_{n\to+\infty} -\frac{1}{n}\log\epsilon^*(n,R)$. The quantity $E(R)$ then provides a measure of how rapidly the error probability approaches zero with an increase in blocklength. This characterization of the reliability function is hence called the the *error exponent analysis*. For a classical channel, the upper bounds of the optimal error can be established using a random coding argument [3]. On the other hand, the lower bound was first developed by Shannon, Gallager, and Berlekamp [4] and was called the *sphere-packing bound.*

Error exponent analysis in classical-quantum (c-q) channels is much more difficult because of the noncommutative nature of quantum mechanics. Dalai [8] employed Shannon-Gallager-Berlekamp's approach to establish a sphere-packing bound with Gallager's expression [4]. It was later pointed out that these two sphere-packing exponents are not equal for general c-q channels . The sphere-packing bound obtained by Dalai [8] had a pre-factor $\mathrm{e}^{-O(\sqrt{n})}$, which is loose in the situation where the transmission rate is close to channel capacity. The main contribution of this paper is to establish a sphere-packing bound with a better pre-factor $O(n^{-t})$ for some $t > 1/2$, which notably improves Dalai's bound [8] from the order of subexponential to polynomial (Corollary 2). When restricting to constant composition codes, we can be more explicit about the obtained pre-factor, namely, $n^{-\frac{1}{2}\left(1+\left|E'_{\mathrm{sp}}(R)\right|+o(1)\right)}$ (Theorem 1). Furthermore, this sphere-packing bound and the best known random coding upper bound [9] in the classical case coincide up to the third-order term. Hence, our result yields an almost exact asymptotics of the sphere-packing bound for constant composition codes.

Our main ingredients are a tight concentration inequality in strong large deviation theory [6] and Blahut's approach of hypothesis testing reduction [5]. The strategy of the proof consists of three steps: (i) formulate the error probability of a certain codebook to a hypothesis testing problem; (ii) give a lower (or called the converse) bound to the type-I error in quantum hypothesis testing; and (iii) relate the error with the strong sphere-packing exponent. In Section 2, we provide two converse bounds for quantum hypothesis testing. The first bound generalizes Blahut's one-shot converse Hoeffding bound [5, Theorem 10] to the quantum case (Proposition 4). Unlike Blahut's result derived in the weak form, we establish a strong sphere-packing bound for c-q channels. For the second bound (Proposition 5), we employ Bahadur-Ranga Rao's inequality [6] to prove a sharp converse bound in step (ii). Finally, we combine these two results to obtain a refined strong sphere-packing bound with a polynomial pre-factor.

1 Notation and Main Result

1.1 Notation

Throughout this paper, we consider a finite-dimensional Hilbert space $\mathcal{H}$. The set of density operators (i.e. positive semi-definite operators with unit trace) on $\mathcal{H}$ are defined as $\mathcal{S}(\mathcal{H})$. We write $\rho \ll \sigma$ if $\mathtt{supp}(\rho) \subset \mathtt{supp}(\sigma)$, where $\mathtt{supp}(\rho)$ denotes the support of ρ. The identity operator on $\mathcal{H}$ is denoted by $\mathbb{1}_{\mathcal{H}}$. When there is no possibility of confusion, we skip the subscript $\mathcal{H}$. Let $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{R}_{>0}$ denote the set of integers, real numbers, and positive real numbers,, respectively. Define $[n] := \{1,2,\ldots,n\}$ for $n \in \mathbb{N}$. Given a pair of positive semi-definite operators $\rho,\sigma \in \mathcal{S}(\mathcal{H})$,

[*] F99942118@ntu.edu.tw
[†] Min-Hsiu.Hsieh@uts.edu.au
[‡] marco.tomamichel@uts.edu.au

we define the (quantum) relative entropy as $D(\rho\|\sigma) := \operatorname{Tr}\left[\rho\left(\log\rho - \log\sigma\right)\right]$, when $\rho \ll \sigma$, and $+\infty$ otherwise. For every $\alpha \in [0,1)$, we define the (Petz) quantum Rényi divergences $D_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\operatorname{Tr}\left[\rho^\alpha\sigma^{1-\alpha}\right]$. For $\alpha = 1$, $D_1(\rho\|\sigma) := \lim_{\alpha\to 1} D_\alpha(\rho\|\sigma) = D(\rho\|\sigma)$. Let $\mathcal{X} = \{1, 2, \ldots, |\mathcal{X}|\}$ be a finite alphabet, and let $\mathscr{P}(\mathcal{X})$ be the set of probability distributions on $\mathcal{X}$. A classical-quantum (c-q) channel W maps elements of the finite set $\mathcal{X}$ to the density operators in $\mathcal{S}(\mathcal{H})$, i.e., $W : \mathcal{X} \to \mathcal{S}(\mathcal{H})$. Let $\mathcal{M}$ be a finite alphabetical set with size $M = |\mathcal{M}|$. An (n-block) *encoder* is a map $f_n : \mathcal{M} \to \mathcal{X}^n$ that encodes each message $m \in \mathcal{M}$ to a codeword $\mathbf{x}^n(m) := x_1(m)\ldots x_n(m) \in \mathcal{X}^n$. The codeword $\mathbf{x}^n(m)$ is then mapped to a state $W^{\otimes n}_{\mathbf{x}^n(m)} = W_{x_1(m)} \otimes \cdots \otimes W_{x_n(m)} \in \mathcal{S}(\mathcal{H}^{\otimes n})$. The *decoder* is described by a positive operator-valued measurement (POVM) $\Pi_n = \{\Pi_{n,1}, \ldots, \Pi_{n,M}\}$ on $\mathcal{H}^{\otimes n}$, where $\Pi_{n,i} \geq 0$ and $\sum_{i=1}^M \Pi_{n,i} = \mathbb{1}$. The pair $(f_n, \Pi_n) =: \mathcal{C}_n$ is called a *code* with *rate* $R = \frac{1}{n}\log|\mathcal{M}|$. The error probability of sending a message m with the code $\mathcal{C}_n$ is $\epsilon_m(W, \mathcal{C}_n) := 1 - \operatorname{Tr}\left(\Pi_{n,m} W_{\mathbf{x}^n(m)}\right)$. We use $\epsilon_{\max}(W, \mathcal{C}_n) = \max_{m\in\mathcal{M}} \epsilon_m(W, \mathcal{C}_n)$ and $\bar{\epsilon}(W, \mathcal{C}_n) = \frac{1}{M}\sum_{m\in\mathcal{M}} \epsilon_m(W, \mathcal{C}_n)$ to denote the *maximal* error probability and the *average* error probability, respectively. Given a sequence $\mathbf{x}^n \in \mathcal{X}^n$, we denote by $P_{\mathbf{x}^n}(x) := \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{x = x_i\}$ the empirical distribution of $\mathbf{x}^n$. A constant composition code with a composition $P_{\mathbf{x}^n}$ refers to a codebook whose codewords all have the same distribution $P_{\mathbf{x}^n}$.

We define the following conditional entropic quantities for the channel W with $P \in \mathscr{P}(\mathcal{X})$: $D_\alpha\left(W\|\sigma|P\right) := \sum_{x\in\mathcal{X}} P(x) D_\alpha\left(W_x\|\sigma\right)$. The *mutual information* of the c-q channel $W : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ with prior distribution $P \in \mathscr{P}(\mathcal{X})$ is defined as $I(P, \mathscr{W}) :== D\left(\mathscr{W}\|P\mathscr{W}|P\right)$, where $P\mathscr{W} := \sum_{x\in\mathcal{X}} P(x) W_x$. The (classical) *capacity* of the channel $\mathscr{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ is denoted by: $C_{\mathscr{W}} := \max_{P\in\mathscr{P}(\mathcal{X})} I(P, \mathscr{W})$. We define two related information quantities: for every $\alpha \in [0,1]$,

$$I^{(1)}_\alpha(P, \mathscr{W}) := \min_{\sigma\in\mathcal{S}(\mathcal{H})} D_\alpha\left(P \circ \mathscr{W}\|P \otimes \sigma\right); \qquad (1)$$

$$I^{(2)}_\alpha(P, \mathscr{W}) := \min_{\sigma\in\mathcal{S}(\mathcal{H})} D_\alpha\left(\mathscr{W}\|\sigma|P\right). \qquad (2)$$

The term $I^{(1)}_\alpha(P, \mathscr{W})$ is called the *α-Rényi mutual information*. The second term $I^{(2)}_\alpha(P, \mathscr{W})$ can be viewed as a variant of the α-Rényi mutual information. For the case of $\alpha = 1$, they both equal conventional mutual information, i.e. $I^{(1)}_1(P, \mathscr{W}) = I^{(2)}_1(P, \mathscr{W}) = I(P, \mathscr{W})$. Mosonyi and Ogawa [10, Proposition IV.2] showed that for all $\alpha \in [0,1]$, $C_{\alpha,\mathscr{W}} := \max_{P\in\mathscr{P}(\mathcal{X})} I^{(1)}_\alpha(P, \mathscr{W}) = \max_{P\in\mathscr{P}(\mathcal{X})} I^{(2)}_\alpha(P, \mathscr{W})$, and it is termed the *Rényi radius* of order α. Let

$$E^{(1)}_{\text{sp}}(R, P) := \sup_{0<\alpha\leq 1} \frac{1-\alpha}{\alpha}\left(I^{(1)}_\alpha(P, \mathscr{W}) - R\right); \qquad (3)$$

$$E^{(2)}_{\text{sp}}(R, P) := \sup_{0<\alpha\leq 1} \frac{1-\alpha}{\alpha}\left(I^{(2)}_\alpha(P, \mathscr{W}) - R\right). \qquad (4)$$

The *sphere-packing exponent* is defined by

$$E_{\text{sp}}(R) := \max_{P\in\mathscr{P}(\mathcal{X})} E^{(1)}_{\text{sp}}(R, P) = \max_{P\in\mathscr{P}(\mathcal{X})} E^{(2)}_{\text{sp}}(R, P),$$

where the last equality follows from [10, Proposition IV.2]. Further, we define a rate, [8]: $R_\infty := C_{0,\mathscr{W}}$. It follows that $E_{\text{sp}}(R) = +\infty$ for any $R \leq R_\infty$ (see also [4, p. 69] and [3, Eq. (5.8.5)]). In this paper, we assume the channel satisfies $R_\infty < C_{\mathscr{W}}$.

Given any $R \in (R_\infty, C_{\mathscr{W}})$ and $P \in \mathscr{P}_R(\mathcal{X})$, we denote a *maximum absolute value subgradient* of the sphere-packing exponent at R by

$$\left|E'_{\text{sp}}(R)\right| := \max_{P: E^{(2)}_{\text{sp}}(R,P) = E_{\text{sp}}(R)} \frac{1-\alpha^\star_{R,P}}{\alpha^\star_{R,P}},$$

where $\alpha^\star_{R,P}$ is the optimizer in Eq. (4).

Consider a binary hypothesis whose null and alternative hypotheses are $\rho \in \mathcal{S}(\mathcal{H})$ and $\sigma \in \mathcal{S}(\mathcal{H})$, respectively. The *type-I error* and *type-II error* of the hypothesis testing, for an operator $0 \leq Q \leq \mathbb{1}$, are defined as $\alpha\left(Q;\rho\right) := \operatorname{Tr}\left[(\mathbb{1} - Q)\rho\right]$, and $\beta\left(Q;\sigma\right) := \operatorname{Tr}\left[Q\sigma\right]$. There is a trade-off between these two errors. Thus, we can define the minimum type-I error, when the type-II error is below $\mu \in (0,1)$, as

$$\widehat{\alpha}_\mu\left(\rho\|\sigma\right) := \min_{0\leq Q\leq\mathbb{1}} \left\{\alpha\left(Q;\rho\right) : \beta\left(Q;\sigma\right) \leq \mu\right\}. \qquad (5)$$

1.2 Main Result

Theorem 1 (Refined Strong Sphere-Packing Bound of Constant Composition Codes)**.** *Consider a classical-quantum channel $\mathscr{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ and $R \in (R_\infty, C_{\mathscr{W}})$. For every $\gamma > 0$, there exist an $N_0 \in \mathbb{N}$ and a constant $A > 0$ such that for all constant composition codes $\mathcal{C}_n$ of length $n \geq N_0$ with message size $|\mathcal{C}_n| \geq \exp\{nR\}$, we have*

$$\bar{\epsilon}\left(\mathcal{C}_n\right) \geq \frac{A}{n^{\frac{1}{2}\left(1+|E'_{\text{sp}}(R)|+\gamma\right)}} \exp\left\{-nE_{\text{sp}}(R)\right\}. \qquad (6)$$

The following corollary generalizes the refined sphere-packing bound for constant composition codes to arbitrary codes via a standard argument [4, p. 95].

Corollary 2 (Refined Strong Sphere-Packing Bound for General Codes)**.** *Consider a classical-quantum channel $\mathscr{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ and $R \in (R_\infty, C_{\mathscr{W}})$. There exist some $t > 1/2$ and $N_0 \in \mathbb{N}$ such that for all codes of length $n \geq N_0$, we have*

$$\epsilon^*\left(n, R\right) \geq n^{-t}\exp\left\{-nE_{\text{sp}}(R)\right\}. \qquad (7)$$

We provide the proof in the full version [12, Section 4.2].

Theorem 1 yields

$$\log\frac{1}{\bar{\epsilon}(\mathcal{C}_n)} \leq nE_{\text{sp}}(R) + \frac{1}{2}\left(1 + \left|E'_{\text{sp}}(R)\right|\right)\log n + o(\log n),$$

On the other hand, for the case of classical *non-singular* channels, it was shown that [9, Theorem 3.6], for all constant composition codes $\mathcal{C}_n$ and rate $R \in (R_{\text{crit}}, C_{\mathscr{W}})$,

$$\log\frac{1}{\bar{\epsilon}(\mathcal{C}_n)} \geq nE_{\text{r}}(R) + \frac{1}{2}\left(1 + |E'_{\text{r}}(R)|\right)\log n + \Omega(1), \qquad (8)$$

where $E_{\text{r}}(R)$ is the *random coding exponent*, and R_{crit} is the critical rate such that $E_{\text{r}}(R) = E_{\text{sp}}(R)$ for all $R \geq R_{\text{crit}}$ [3, p. 160]. Hence our result, Theorem 1, matches the achievability up to the logarithmic order.

2 Proof Ideas

To establish our main result, we combine Blahut's insight of relating a channel coding problem to binary hypothesis testing [5] with a sharp concentration inequality employed in Ref. [7]. Our proof consists of three major steps: (i) reduce the channel coding problem to binary hypothesis testing (Lemma 3); (ii) bound its type-I error from below (Propositions 4 and 5); (iii) relate the derived bound to the sphere-packing exponent.

Lemma 3. *For any classical-quantum channel $W : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ and any code $\mathcal{C}_n$ with message size M, it follows that*

$$\epsilon_{\max}\left(\mathcal{C}_n\right) \geq \max_{\sigma\in\mathcal{S}(H)} \min_{\mathbf{x}^n\in\mathcal{C}_n} \widehat{\alpha}_{\frac{1}{M}}\left(W_{\mathbf{x}^n}^{\otimes n}\|\sigma^{\otimes n}\right). \tag{9}$$

We provide the proof in the full version [12, Section 4].

Proposition 4 (Chebyshev-Type Converse Hoeffding Bound)**.** *Let $R \in (R_\infty, C_{\mathscr{W}})$. Consider the hypotheses*

$$\mathsf{H}_0 : \rho^n = W_{\mathbf{x}^n}^{\otimes n}; \tag{10}$$

$$\mathsf{H}_1 : \sigma^n = \left(\sigma^\star\right)^{\otimes n}, \tag{11}$$

where $\mathbf{x}^n \in \mathcal{X}^n$ and $\sigma^\star \in \arg\min_{\sigma\in\mathcal{S}(\mathcal{H})}\sup_{0<\alpha\leq 1}\frac{1-\alpha}{\alpha}\left(D_\alpha\left(\mathscr{W}\|\sigma|P_{\mathbf{x}^n}\right)-R\right)$. Then, for every $c > 0$, there exist $N_0 \in \mathbb{N}$ and $\kappa_1, \kappa_2 \in \mathbb{R}_{>0}$ such that for all $n \geq N_0$ we have

$$\widehat{\alpha}_{c\exp\{-nR\}}\left(\rho^n\|\sigma^n\right) \geq \kappa_1 \exp\left\{-\kappa_2\sqrt{n} - nE_{\rm sp}^{(2)}\left(R, P_{\mathbf{x}^n}\right)\right\},$$

We provide the proof in the full version [12, Section 4].

Proposition 5 (Sharp Converse Hoeffding Bound)**.** *Let $R \in (R_\infty, C_{\mathscr{W}})$. Consider the hypothesese:*

$$\mathsf{H}_0 : \rho^n = W_{\mathbf{x}^n}^{\otimes n}; \tag{12}$$

$$\mathsf{H}_1 : \sigma^n = \left(\sigma^\star\right)^{\otimes n}, \tag{13}$$

where $\mathbf{x}^n \in \mathcal{X}^n$, and $\sigma^\star := \arg\min_{\sigma\in\mathcal{S}(\mathcal{H})}\sup_{0<\alpha\leq 1}\frac{1-\alpha}{\alpha}\left(D_\alpha\left(\mathscr{W}\|\sigma|P_{\mathbf{x}^n}\right)-R\right)$ satisfying $E_{\rm sp}^{(2)}\left(R, P_{\mathbf{x}^n}\right) \in [\nu, +\infty)$ for some positive $\nu > 0$. For every $c > 0$, there exists a constant $N_0 \in \mathbb{N}$, independent of the sequences ρ^n and σ^n, such that for all $n \geq N_0$ we have

$$\widehat{\alpha}_{c\exp\{-nR\}}\left(\rho^n\|\sigma^n\right) \geq \frac{A}{n^{\frac{1}{2}(1+s^\star)}}\exp\left\{-nE_{\rm sp}^{(2)}\left(R, P_{\mathbf{x}^n}\right)\right\},$$

where $s^\star := -\left.\frac{\partial E_{\rm sp}^{(2)}(r,P)}{\partial r}\right|_{r=R}$, and $A \in \mathbb{R}_{>0}$ is a finite constant depending on R, ν and $\mathscr{W}$.

We provide the proof in the full version [12, Section 4].

3 Discussion

In this paper, we obtained a refined strong sphere-packing bound for c-q channels and constant composition codes with a polynomial pre-factor $n^{-\frac{1}{2}\left(1+|E'_{\rm sp}(R)|+o(1)\right)}$. Moreover, the established result matches the best known random coding bound (i.e. achievability) up to the logarithmic order [7, 9]. For the case of general codes, the derived pre-factor is of the polynomial order, i.e. $O(n^{-t})$ for some $t > 1/2$. We are able to obtain the exact pre-factor without the assumption of constant composition codes for a class of symmetric c-q channels. We note that the exact pre-factor for general codes is still open even in the classical case. Finally, our refinement enables a moderate deviation analysis in c-q channels [13].

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.

[2] ——, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, no. 3, pp. 611–656, 1959.

[3] R. Gallager, *Information Theory and Reliable Communication*. Wiley, 1968.

[4] C. Shannon, R. Gallager, and E. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. I," *Inform. Control*, vol. 10, no. 1, pp. 65–103, Jan 1967.

[5] R. E. Blahut, "Hypothesis testing and information theory," *IEEE Trans. Inform. Theory*, vol. 20, no. 4, pp. 405–417, Jul 1974.

[6] R. R. Bahadur and R. R. Rao, "On deviations of the sample mean," *Ann. Math. Stat.*, vol. 31, no. 4, pp. 1015–1027, 1960.

[7] Y. Altuğ and A. B. Wagner, "Refinement of the sphere-packing bound: Asymmetric channels," *IEEE Trans. Inform. Theory*, vol. 60, no. 3, pp. 1592–1614, Mar 2014.

[8] M. Dalai, "Lower bounds on the probability of error for classical and classical-quantum channels," *IEEE Trans. Inform. Theory*, vol. 59, no. 12, pp. 8027–8056, Dec 2013.

[9] J. Scarlett, "Reliable communication under mismatched decoding," *PhD Thesis (University of Cambridge)*, 2014.

[10] M. Mosonyi and T. Ogawa, "Strong converse exponent for classical-quantum channel coding," `arXiv:1409.3562`.

[11] M. Nussbaum and A. Szkoła, "The Chernoff lower bound for symmetric quantum hypothesis testing," *Ann. Stat.*, vol. 37, no. 2, pp. 1040–1057, apr 2009.

[12] H.-C. Cheng, M.-H. Hsieh, and M. Tomamichel, "Quantum Sphere-Packing Bounds with Polynomial Prefactors," `arXiv:1704.05703 [quant-ph]`.

[13] H.-C. Cheng and M.-H. Hsieh, "Moderate deviation analysis for classical-quantum channels and quantum hypothesis testing," `arXiv:1701.03195 [quant-ph]`.

Quantum Sphere-Packing Bounds with Polynomial Prefactors

Hao-Chung Cheng[1,2], Min-Hsiu Hsieh[2], and Marco Tomamichel[2,3]

[1] *Graduate Institute Communication Engineering, National Taiwan University, Taiwan (R.O.C.)*
[2] *Centre for Quantum Software and Information (UTS:Q$|$SI$\rangle$),*
Faculty of Engineering and Information Technology, University of Technology Sydney, Australia
[3] *School of Physics, The University of Sydney, Australia*

Abstract. We study lower bounds on the optimal error probability in classical coding over classical-quantum channels at rates below the capacity, commonly termed quantum sphere-packing bounds. Winter and Dalai have derived such bounds for classical-quantum channels; however, the exponents in their bounds only coincide when the channel is classical. In this paper, we show that these two exponents admit a variational representation and are related by the Golden-Thompson inequality, reaffirming that Dalai's expression is stronger in general classical-quantum channels. Second, we establish a sphere-packing bound for classical-quantum channels, which significantly improves Dalai's prefactor from the order of subexponential to polynomial. Furthermore, the gap between the obtained error exponent for constant composition codes and the best known classical random coding exponent vanishes in the order of $o(\log n/n)$, indicating our sphere-packing bound is almost exact in the high rate regime. Finally, for a special class of symmetric classical-quantum channels, we can completely characterize its optimal error probability without the constant composition code assumption. The main technical contributions are two converse Hoeffding bounds for quantum hypothesis testing and the saddle-point properties of error exponent functions.

1. Introduction

Shannon's noisy coding theorem [1] states that a message in an appropriately coded form can be reliably transmitted through a discrete memoryless channel $\mathcal{W}$, provided the coding rate R is below the channel capacity $C_{\mathcal{W}}$. More precisely, the probability of decoding errors can be made arbitrarily small as the coding blocklength grows. Later, Shannon himself pioneered the study of the exponential dependency of the optimal error probability $\epsilon^*(n, R)$ for a blocklength n and transmission rate R [2]. He defined the *reliability function* to be, for any fixed coding rate $R < C_{\mathcal{W}}$,

$$E(R) := \limsup_{n\to+\infty} -\frac{1}{n}\log \epsilon^*(n,R). \tag{1}$$

The quantity $E(R)$ then provides a measure of how rapidly the error probability approaches zero with an increase in blocklength. This characterization of the reliability function is hence called the *reliability function analysis* or the *error exponent analysis*.

For a classical channel, lower bounds for the reliability function can be established by random coding arguments [3, 4, 5, 6]. However, upper bounds require different techniques since the code-dependent bounds on the error probability need to be optimized over all codebooks. The first result—the *sphere-packing bound* $E(R) \le E_{\text{sp}}(R)$—was developed by Shannon, Gallager, and Berlekamp [7]. The *sphere-packing exponent* $E_{\text{sp}}(R)$ is defined as

$$E_{\text{sp}}(R) := \sup_{s\ge 0}\left\{ \max_P E_0(s,P) - sR \right\}, \tag{2}$$

where P is maximized over all probability distributions on the input alphabet, and $E_0(s,P)$ is the *auxiliary function* or *Gallager's exponent* [5]. Unlike Shannon-Gallager-Berlekamp's technique which relates

E-mail address: F99942118@ntu.edu.tw, Min-Hsiu.Hsieh@uts.edu.au, marco.tomamichel@uts.edu.au.
Part of this paper was accepted in 2017 IEEE International Symposium on Information Theory (ISIT).

channel coding to binary hypothesis testing, Haroutunian [8, 9] employed a combinatorial method and obtained an upper bound for the reliability function in terms of the following expression

$$\widetilde{E}_{\text{sp}}(R) := \max_{P} \min_{\mathcal{V}} \left\{ D\left(\mathcal{V}\|\mathcal{W}|P\right) : I(P,\mathcal{V}) \leq R \right\}, \tag{3}$$

where $\mathcal{V}$ is minimized over all channels with the same output alphabet as $\mathcal{W}$, $D(\mathcal{V}\|\mathcal{W}|P)$ is the conditional relative entropy between the dummy channel $\mathcal{V}$ and the true channel $\mathcal{W}$, and $I(P,\mathcal{V})$ is the mutual information of the channel $\mathcal{V}$ (the detailed definitions are given in Section 2). It was later realized that the two quantities in Eqs. (2) and (3) are equivalent: they are related by convex program duality [10, 11, 12]. Therefore, these two expressions, Eqs. (2) or (3), are both called sphere-packing exponents.

Error exponent analysis in classical-quantum (c-q) channels is more challenging because of the noncommutative nature of quantum mechanics. Burnashev and Holevo [13] introduced a quantum version of the auxiliary function [14, 15] and initialized the study of reliability functions in c-q channels. Winter [16] derived a sphere-packing bound for c-q channels in the form of $\widetilde{E}_{\text{sp}}(R)$ in Eq. (3), generalizing Haroutunian's idea [8]. Dalai [17] employed Shannon-Gallager-Berlekamp's approach [7] to establish a sphere-packing bound with Gallager's exponent in Eq. (2). In the follow-up work [18], Dalai and Winter pointed out that these two exponents are not equal in c-q channels. In this work, we explicitly demonstrate a relationship between the two quantities. Precisely, we show that they individually admit a variational representation (Theorem 6 in Section 3):

$$E_{\text{sp}}(R) = \max_{P} \sup_{0<\alpha\leq 1} \min_{\sigma} \left\{ \frac{1-\alpha}{\alpha} \left(\sum_{x} P(x) D_{\alpha}\left(W_x \| \sigma\right) - R \right) \right\}; \tag{4}$$

$$\widetilde{E}_{\text{sp}}(R) = \max_{P} \sup_{0<\alpha\leq 1} \min_{\sigma} \left\{ \frac{1-\alpha}{\alpha} \left(\sum_{x} P(x) D_{\alpha}^{\flat}\left(W_x \| \sigma\right) - R \right) \right\}, \tag{5}$$

where σ is minimized over all density operators on some Hilbert space $\mathcal{H}$; W_x is the channel output state on $\mathcal{H}$; D_α is the (Petz) α-Rényi divergence [19]; and $D_\alpha^\flat$ is the *log-Euclidean* α-Rényi divergence.

Since $D_\alpha \geq D_\alpha^\flat$ for all $\alpha \in (0,1]$, as a simple consequence of the Golden-Thompson inequality [20, 21], the exponent $E_{\text{sp}}(R)$ in Eq. (4) is stronger than $\widetilde{E}_{\text{sp}}(R)$ in Eq. (5), i.e.

$$E(R) \leq E_{\text{sp}}(R) \leq \widetilde{E}_{\text{sp}}(R). \tag{6}$$

These two exponents coincide[2] only when all the channel output states commute (i.e. for classical channels). Thus, we call $E_{\text{sp}}(R)$ and $\widetilde{E}_{\text{sp}}(R)$ the *strong sphere-packing exponent* and the *weak sphere-packing exponent*, respectively. The lower bounds for the optimal error probability in terms of these two quantities are called the strong sphere-packing bound

$$\epsilon^*\left(n,R\right) \geq f(n) \exp\left\{-n\left[E_{\text{sp}}(R-g(n))\right]\right\}, \tag{7}$$

and the weak sphere-packing bound

$$\epsilon^*\left(n,R\right) \geq f(n) \exp\left\{-n\left[\widetilde{E}_{\text{sp}}(R-g(n))\right]\right\}, \tag{8}$$

where $f(n)$ is the pre-factor of the bound, and $g(n)$ is a rate back-off term. We note that $g(n) = 0$ in our main result, and hence we only study $f(n)$ in the following discussion.

The strong sphere-packing bound obtained by Dalai [17] had a pre-factor $f(n) = \mathrm{e}^{-O(\sqrt{n})}$, which is loose for small blocklength n or in the situation where the transmission rate is close to channel capacity. The main contribution of this paper is to establish a sphere-packing bound with a better pre-factor $f(n) = O(n^{-t})$ for some $t > 1/2$, which notably improves Dalai's bound [17] from the order of subexponential to polynomial (Corollary 10). When restricting to constant composition codes, we can be more explicit about the obtained pre-factor, namely, $f(n) = n^{-\frac{1}{2}\left(1+\left|E_{\text{sp}}'(R)\right|+o(1)\right)}$ (Theorem 9). Furthermore, this sphere-packing bound and the best known random coding upper bound [22, 23, 24, 25] in the classical case coincide up to the third-order term (see the discussion in Section 4)). Hence, our result yields an almost

[2] For the coding rates above channel capacity, these two exponents are both zero (α attains 1 in Eqs. (4) and (5)). We exclude this trivial case and only consider the rate being strictly below capacity.

exact asymptotics of the sphere-packing bound for constant composition codes. Our second contribution is to show that, for a class of symmetric c-q channels, the pre-factor $f(n) = O(n^{-\frac{1}{2}\left(1+\left|E_{\text{sp}}'(R)\right|\right)})$, holds for general codes. In other words, we are able to obtain an exact sphere-packing bound for general codes, by exploiting a symmetric property of the channel.

Our main ingredients are a tight concentration inequality in strong large deviation theory [26], [27, Theorem 3.7.4], [28, Section III.D] (Appendix B) and Blahut's approach of hypothesis testing reduction [10]. The strategy of the proof consists of three steps: (i) formulate the error probability of a certain codebook to a hypothesis testing problem; (ii) give a lower (or called the converse) bound to the type-I error in quantum hypothesis testing; and (iii) relate the error with the strong sphere-packing exponent. In Section 4.1, we provide two converse bounds for quantum hypothesis testing. The first bound generalizes Blahut's one-shot converse Hoeffding bound [10, Theorem 10] to the quantum case (Proposition 12). Unlike Blahut's result derived in the weak form, we establish a strong sphere-packing bound for c-q channels. For the second bound (Proposition 14), we employ Bahadur-Ranga Rao's inequality [26] to prove a sharp converse bound in step (ii). Finally, we combine these two results to obtain a refined strong sphere-packing bound with a polynomial pre-factor.

Table 1 collects major proof approaches of classical sphere-packing bounds, Eqs. (7) and (8), and discusses their generalizations to c-q channels. We remark that the established polynomial pre-factor is crucial for the analysis of coding performance in the medium error probability regime (more commonly known as moderate deviation analysis) [28, 29, 30].

The remaining part of the paper is organized as follows. Section 2 introduces the notation and necessary preliminaries. The relationship between the weak and strong sphere-packing exponents is proved in Section 3. In Section 4, we prove a refined sphere-packing bound for c-q channels. We consider a symmetric c-q channel and establish an exact sphere-packing bound in Section 5. Lastly, we conclude this paper in Section 6.

Bounds\Settings	Blocklength n	Composition dependent	Pre-factor $f(n)$	Rate back-off $g(n)$	Classical-quantum channels	Tightness
(a) Shannon-Gallager-Berlekamp [7]	Any n	Yes	$\mathrm{e}^{-O(\sqrt{n})}$	$O\left(\frac{\log n}{n}\right)$	Dalai [17]	Strong
(b) Haroutunian [8] Omura [31] Csisár-Korner [12]	Large n	Yes	$\mathrm{e}^{-o(n)}$	$o(1)$	Winter [16]	Weak
(c) Blahut [10]	Any n	No	$\mathrm{e}^{-O(\sqrt{n})}$	$O\left(n^{-\frac{1}{2}}\right)$	Eqs. (187) & (192)	Strong
(d) Altuğ-Wagner [32]	Large n	Yes	$n^{-\frac{1}{2}\left(1+\left\|E_{\text{sp}}'(R)\right\|+o(1)\right)}$	0	Theorem 9	Strong
(e) Elkayam-Feder [33]	Any n	Yes	$O\left(n^{-t}\right)$	$O\left(\frac{\log n}{n}\right)$	Unknown	Unknown
(f) Agustin-Nakiboğlu [34, 35, 36, 37]	Large n	No	$O\left(n^{-t}\right)$	0	Unknown	Unknown

Table 1. Different sphere-packing bounds are compared by (i) whether the bounds hold for any blocklength n or only for sufficiently large $n \in \mathbb{N}$; (ii) whether or not they are dependent on the constant composition codes; (iii) & (iv) the asymptotics $f(n)$ and $g(n)$; (v) the corresponding c-q generalizations. The parameter t in rows (e) and (f) is some value in the range $t > 1/2$; and (vi) whether their error exponent expressions for c-q channels are in the strong form (Eq. (2)) or weak form (Eq. (3)).

2. Notation and Preliminaries

Throughout this paper, we consider a finite-dimensional Hilbert space $\mathcal{H}$. The set of density operators (i.e. positive semi-definite operators with unit trace) and the set of full-rank density operators on $\mathcal{H}$ are defined as $\mathcal{S}(\mathcal{H})$ and $\mathcal{S}_{>0}(\mathcal{H})$, respectively. For $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, we write $\rho \ll \sigma$ if the support of ρ is contained in the support of σ. The identity operator on $\mathcal{H}$ is denoted by $\mathbb{1}_{\mathcal{H}}$. If there is no possibility of confusion, we will skip the subscript $\mathcal{H}$. We use $\operatorname{Tr}[\,\cdot\,]$ to denote the trace. Let $\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, and $\mathbb{R}_{>0}$ denote the set of integers, real numbers, non-negative real numbers, and positive real numbers, respectively. Define $[n] := \{1, 2, \ldots, n\}$ for $n \in \mathbb{N}$.

For a positive semi-definite operator A whose spectral decomposition is $A = \sum_i a_i P_i$, where $(a_i)_i$ and $(P_i)_i$ are the eigenvalues and eigenprojections of A, its power is defined as: $A^p := \sum_{i: a_i \neq 0} a_i^p P_i$. In particular, A^0 denotes the projection onto $\texttt{supp}(A)$, where we use $\texttt{supp}(A)$ to denote the support of the operator A. Further, $A \perp B$ means $\texttt{supp}(A) \cap \texttt{supp}(B) = \emptyset$. We denote by log the natural logarithm.

2.1. **Information Quantities and Error-Exponent Functions.** Given a pair of positive semi-definite operators $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, we define quantum relative entropy [38, 39] and relative variance [40, 41, 42], respectively as

$$D(\rho\|\sigma) := \operatorname{Tr}\left[\rho\left(\log\rho - \log\sigma\right)\right]; \tag{9}$$

$$V(\rho\|\sigma) := \operatorname{Tr}\left[\rho\left(\log\rho - \log\sigma\right)^2\right] - D(\rho\|\sigma)^2, \tag{10}$$

when $\rho \ll \sigma$, and $+\infty$ otherwise.

For density operators $\rho, \sigma \in \mathcal{S}_{>0}(\mathcal{H})$, and every $\alpha \in (0,1)$, we define the following two families of quantum Rényi divergences [19, 43, 44]:

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1}\log Q_\alpha(\rho\|\sigma), \quad Q_\alpha(\rho\|\sigma) := \operatorname{Tr}\left[\rho^\alpha \sigma^{1-\alpha}\right]; \tag{11}$$

$$D_\alpha^\flat(\rho\|\sigma) := \frac{1}{\alpha - 1}\log Q_\alpha^\flat(\rho\|\sigma), \quad Q_\alpha^\flat(\rho\|\sigma) := \operatorname{Tr}\left[\mathrm{e}^{\alpha \log\rho + (1-\alpha)\log\sigma}\right]. \tag{12}$$

We term the above quantities as the *(Petz) α-Rényi divergence*, and the *log-Euclidean α-Rényi divergence*, respectively. The log-Euclidean Rényi divergence arises from the *log-Euclidean operator mean* (also called the *chaotic mean*): $A \diamondsuit_\alpha B := \exp\left((1-\alpha)\log A + \alpha \log B\right)$ for $0 \leq \alpha \leq 1$. For general density operators $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, the above definitions can be extended as

$$Q_\alpha(\rho\|\sigma) := \lim_{\delta \downarrow 0} Q_\alpha(\rho + \delta\mathbb{1}\|\sigma + \delta\mathbb{1}) \quad \text{and} \quad Q_\alpha^\flat(\rho\|\sigma) := \lim_{\delta\downarrow 0} Q_\alpha^\flat(\rho + \delta\mathbb{1}\|\sigma + \delta\mathbb{1}). \tag{13}$$

Note that these two quantities are related by the Golden-Thompson inequality [20, 21]:

$$Q_\alpha^\flat(\rho\|\sigma) \leq Q_\alpha(\rho\|\sigma),\ \forall \alpha \in (0,1). \tag{14}$$

For $\alpha = 1$ and $\alpha = 0$, we define (see e.g. [44, Lemma III.4]):

$$D_1(\rho\|\sigma) := \lim_{\alpha\uparrow 1} D_\alpha(\rho\|\sigma) = D(\rho\|\sigma), \quad D_1^\flat(\rho\|\sigma) := \lim_{\alpha \uparrow 1} D_\alpha^\flat(\rho\|\sigma) = D(\rho\|\sigma); \tag{15}$$

$$D_0(\rho\|\sigma) := \lim_{\alpha\downarrow 0} D_\alpha(\rho\|\sigma), \quad D_0^\flat(\rho\|\sigma) := \lim_{\alpha\downarrow 0} D_\alpha^\flat(\rho\|\sigma). \tag{16}$$

We will need the following lemma in the next section.

Lemma 1 ([45], [44, Lemma III.3, Lemma III.11, Theorem III.14, Corollary III.25], [46, Corollary 2.2])**.** *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$. Then,*

$$\alpha \mapsto \log Q_\alpha(\rho\|\sigma) \ \textit{and}\ \alpha \mapsto \log Q_\alpha^\flat(\rho\|\sigma) \ \ \textit{are convex on}\ (0,1); \tag{17}$$

$$\alpha \mapsto D_\alpha\left(\rho\|\sigma\right) \ \ \textit{is continuous and monotone increasing on}\ [0,1]. \tag{18}$$

Moreover[3]*,*

$$\forall \alpha \in (0,1), \quad (\rho,\sigma) \mapsto Q_\alpha^\flat(\rho\|\sigma) \ \ \textit{is jointly concave on}\ \mathcal{S}(\mathcal{H}) \times \mathcal{S}(\mathcal{H}); \tag{19}$$

$$\forall \alpha \in [0,1], \quad \sigma \mapsto D_\alpha(\rho\|\sigma) \ \ \textit{is strictly convex and lower semi-continuous on}\ \mathcal{S}(\mathcal{H}). \tag{20}$$

[3] It was shown in [44, Lemma III.22] that the map $\sigma \mapsto D_\alpha(\rho\|\sigma)$ is lower semi-continuous on $\mathcal{S}(\mathcal{H})$ for all $\alpha \in (0,1)$. The argement can be extended to the range $\alpha \in [0,1]$ by the same method in [44, Lemma III.22].

Let $\mathcal{X} = \{1, 2, \ldots, |\mathcal{X}|\}$ be a finite alphabet, and let $\mathcal{P}(\mathcal{X})$ be the set of probability distributions on $\mathcal{X}$. A classical-quantum (c-q) channel $\mathcal{W}$ maps elements of the finite set $\mathcal{X}$ to density operators in $\mathcal{S}(\mathcal{H})$, i.e. $\mathcal{W} : x \mapsto W_x$. For a c-q channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ and $P \in \mathcal{P}(\mathcal{X})$, it is convenient to denote the corresponding c-q state:

$$P \circ \mathcal{W} := \sum_{x \in \mathcal{X}} P(x) |x\rangle\langle x| \otimes W_x. \tag{21}$$

We also express the input distribution $P \in \mathcal{P}(\mathcal{X})$ as a diagonal matrix with respect to the computational basis $\{|x\rangle\}_{x \in \mathcal{X}}$, i.e. $P = \sum_{x \in \mathcal{X}} P(x)|x\rangle\langle x|$. Denote the conditional relative entropy of two c-q channels $\mathcal{V}, \mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ with a prior distribution $P \in \mathcal{P}(\mathcal{X})$ by

$$D\left(\mathcal{V} \| \mathcal{W} | P\right) := \sum_{x \in \mathcal{X}} P(x) D\left(V_x \| W_x\right). \tag{22}$$

Similarly, we define the following conditional entropic quantities for $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$, $\sigma \in \mathcal{S}(\mathcal{H})$ and $P \in \mathcal{P}(\mathcal{X})$:

$$D\left(\mathcal{W} \| \sigma | P\right) := \sum_{x \in \mathcal{X}} P(x) D\left(W_x \| \sigma\right), \tag{23}$$

$$D_\alpha\left(\mathcal{W} \| \sigma | P\right) := \sum_{x \in \mathcal{X}} P(x) D_\alpha\left(W_x \| \sigma\right), \tag{24}$$

$$D_\alpha^\flat\left(\mathcal{W} \| \sigma | P\right) := \sum_{x \in \mathcal{X}} P(x) D_\alpha^\flat\left(W_x \| \sigma\right). \tag{25}$$

The *mutual information* of the prior distribution $P \in \mathcal{P}(\mathcal{X})$ and the c-q channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ is defined as

$$I(P, \mathcal{W}) := \inf_{\sigma \in \mathcal{S}(\mathcal{H})} D\left(\mathcal{W} \| \sigma | P\right) = D\left(\mathcal{W} \| P\mathcal{W} | P\right), \tag{26}$$

where $P\mathcal{W} := \sum_{x \in \mathcal{X}} P(x) W_x$. The (classical) *capacity* of the channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ is denoted by [47, 48]:

$$C_{\mathcal{W}} := \max_{P \in \mathcal{P}(\mathcal{X})} I(P, \mathcal{W}). \tag{27}$$

We define two related information quantities: for every $\alpha \in [0, 1]$,

$$I_\alpha^{(1)}(P, \mathcal{W}) := \inf_{\sigma \in \mathcal{S}(\mathcal{H})} D_\alpha\left(P \circ \mathcal{W} \| P \otimes \sigma\right); \tag{28}$$

$$I_\alpha^{(2)}(P, \mathcal{W}) := \inf_{\sigma \in \mathcal{S}(\mathcal{H})} D_\alpha\left(\mathcal{W} \| \sigma | P\right). \tag{29}$$

The term $I_\alpha^{(1)}(P, \mathcal{W})$ is called the *α-Rényi mutual information* [49, 50, 36] or the *generalized Holevo quantity*. The second term $I_\alpha^{(2)}(P, \mathcal{W})$ can be viewed as a variant of the α-Rényi mutual information. It can be verified that these two functions are related by Jensen's inequality:

$$I_\alpha^{(1)}(P, \mathcal{W}) \leq I_\alpha^{(2)}(P, \mathcal{W}). \tag{30}$$

For the case of $\alpha = 1$, they both equal conventional mutual information, i.e. $I_1^{(1)}(P, \mathcal{W}) = I_1^{(2)}(P, \mathcal{W}) = I(P, \mathcal{W})$. Mosonyi and Ogawa [44, Proposition IV.2] showed that for all $\alpha \in [0, 1]$,

$$C_{\alpha, \mathcal{W}} := \sup_{P \in \mathcal{P}(\mathcal{X})} I_\alpha^{(1)}(P, \mathcal{W}) = \sup_{P \in \mathcal{P}(\mathcal{X})} I_\alpha^{(2)}(P, \mathcal{W}), \tag{31}$$

and it is termed the *Rényi radius* or the *Rényi capacity* of order α. Moreover, Proposition 2 below and the compactness of $\mathcal{P}(\mathcal{X})$ show that the suprema in Eq. (31) can be replaced with maxima. The following proposition presents important properties of α-Rényi mutual information and radius. The proof is given in Appendix C.

Proposition 2 (Properties of α-Rényi Mutual Information and Radius)**.** *Given any classical-quantum channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$, the following holds:*

(a) The map $(\alpha, P) \mapsto I_\alpha^{(2)}(P, \mathcal{W})$ *is continuous on* $[0,1] \times \mathcal{P}(\mathcal{X})$*.*

(b) For every $P \in \mathcal{P}(\mathcal{X})$*,* $\alpha \mapsto I_\alpha^{(2)}(P, \mathcal{W})$ *is monotone increasing on* $[0,1]$*.*

(c) For every $P \in \mathcal{P}(\mathcal{X})$*,* $\alpha \mapsto \frac{1-\alpha}{\alpha} I_\alpha^{(2)}(P, \mathcal{W})$ *is strictly concave on* $(0,1)$*.*

(d) The map $\alpha \mapsto C_{\alpha,\mathcal{W}}$ *is continuous and monotone increasing on* $[0,1]$*.*

Items (a), (b), and (c) also hold for $I_\alpha^{(1)}(P, \mathcal{W})$*.*

The *strong sphere-packing exponent* [17] of a c-q channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ and a rate $R \geq 0$ is defined by

$$E_{\text{sp}}(R) := \max_{P \in \mathcal{P}(\mathcal{X})} E_{\text{sp}}(R, P), \tag{32}$$

where

$$E_{\text{sp}}(R, P) := \sup_{s \geq 0} \left\{ E_0(s, P) - sR \right\}, \tag{33}$$

and E_0 is the *auxiliary function* of the c-q channel $\mathcal{W}$ (see [13, 14, 15]):

$$E_0(s, P) := -\log \operatorname{Tr} \left[\left(\sum_{x \in \mathcal{X}} P(x) \cdot W_x^{1/(1+s)} \right)^{1+s} \right] \tag{34}$$

for all $P \in \mathcal{P}(\mathcal{X})$ and $s \geq 0$.

The *weak sphere-packing exponent* [16] is defined as

$$\widetilde{E}_{\text{sp}}(R) := \max_{P \in \mathcal{P}(\mathcal{X})} \widetilde{E}_{\text{sp}}(R, P), \tag{35}$$

where

$$\widetilde{E}_{\text{sp}}(R, P) := \min_{\mathcal{V}: \mathcal{X} \to \mathcal{S}(\mathcal{H})} \left\{ D\left(\mathcal{V} \| \mathcal{W} | P \right) : I(P, \mathcal{V}) \leq R \right\}. \tag{36}$$

We also need the following definitions: for any $R \geq 0$ and $P \in \mathcal{P}(\mathcal{X})$,

$$E_{\text{sp}}^{(1)}(R, P) := \sup_{0 < \alpha \leq 1} \frac{1-\alpha}{\alpha} \left(I_\alpha^{(1)}(P, \mathcal{W}) - R \right); \tag{37}$$

$$E_{\text{sp}}^{(2)}(R, P) := \sup_{0 < \alpha \leq 1} \frac{1-\alpha}{\alpha} \left(I_\alpha^{(2)}(P, \mathcal{W}) - R \right), \tag{38}$$

Eq. (30) implies that (see also Theorem 6) $E_{\text{sp}}^{(1)}(R, P) \leq E_{\text{sp}}^{(2)}(R, P)$. By quantum Sibson's identity [51], one finds

$$E_{\text{sp}}^{(1)}(R, P) = E_{\text{sp}}(R, P). \tag{39}$$

Proposition 2 and Eq. (31) imply that the two quantities given in Eqs. (37) and (38) are equal to the strong sphere-packing exponent by maximizing over the input distributions:

$$E_{\text{sp}}(R) = \max_{P \in \mathcal{P}(\mathcal{X})} E_{\text{sp}}^{(1)}(R, P) = \max_{P \in \mathcal{P}(\mathcal{X})} E_{\text{sp}}^{(2)}(R, P). \tag{40}$$

Further, we define [12, p. 152], [17, Theorem 6]:

$$R_\infty := C_{0,\mathcal{W}}. \tag{41}$$

From the definitions in Eqs. (27) and (41), it can be verified that $R_\infty \leq C_{\mathcal{W}}$ for all c-q channels $\mathcal{W}$. In Proposition 4 below, one has $E_{\text{sp}}(R) = +\infty$ for $R < R_\infty$, and $E_{\text{sp}}(R) = 0$ as $R > C_{\mathcal{W}}$. Throughout this paper, we further assume that the considered c-q channel $\mathcal{W}$ satisfies $R_\infty < C_{\mathcal{W}}$.

As we will show in Section 4, the quantity $E_{\text{sp}}^{(2)}(R, P)$ plays a significant role in the connection between hypothesis testing and channel coding. Moreover, Proposition 3 below shows that the the optimizer in Eqs. (29) and (38) forms a saddle-point. The proof closely follows Altuğ and Wagner [32, Proposition 1], and is given in Appendix D.

Proposition 3 (Saddle-Point)**.** *Consider a classical-quantum channel* $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$*, any* $R \in (R_\infty, C_\mathcal{W})$*, and* $P \in \mathcal{P}(\mathcal{X})$*. Let*

$$\mathcal{S}_{P,\mathcal{W}}(\mathcal{H}) := \{\sigma \in \mathcal{S}(\mathcal{H}) : \forall x \in \mathtt{supp}(P),\, W_x \not\perp \sigma\}\,. \tag{42}$$

Define

$$F_{R,P}(\alpha,\sigma) := \begin{cases} \frac{1-\alpha}{\alpha}\left(D_\alpha\left(\mathcal{W}\|\sigma|P\right) - R\right), & \alpha \in (0,1) \\ 0, & \alpha = 1 \end{cases}, \tag{43}$$

on $(0,1] \times \mathcal{S}(\mathcal{H})$*, and denote by*

$$\mathcal{P}_R(\mathcal{X}) := \left\{ P \in \mathcal{P}(X) : \sup_{0<\alpha\leq 1} \inf_{\sigma\in\mathcal{S}(\mathcal{H})} F_{R,P}(\alpha,\sigma) \in \mathbb{R}_{>0} \right\}. \tag{44}$$

The following holds

(a) For any $P \in \mathcal{P}(\mathcal{X})$*,* $F_{R,P}(\cdot,\cdot)$ *has a saddle-point on* $(0,1] \times \mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$ *with the saddle-value:*

$$\min_{\sigma\in\mathcal{S}(\mathcal{H})} \sup_{0<\alpha\leq 1} F_{R,P}(\alpha,\sigma) = \sup_{0<\alpha\leq 1} \min_{\sigma\in\mathcal{S}(\mathcal{H})} F_{R,P}(\alpha,\sigma) = E_{\text{sp}}^{(2)}(R,P). \tag{45}$$

(b) If $P \in \mathcal{P}_R(\mathcal{X})$*, the saddle-point is unique.*

(c) Fix $P \in \mathcal{P}_R(\mathcal{X})$*. Any saddle-point* $(\alpha_{R,P}^\star, \sigma_{R,P}^\star)$ *of* $F_{R,P}(\cdot,\cdot)$ *satisfies* $\alpha_{R,P}^\star \in (0,1)$ *and*

$$\sigma_{R,P}^\star \gg W_x, \quad \forall x \in \mathtt{supp}(P). \tag{46}$$

The following proposition discusses the continuity and differentiability of the error-exponent functions. The proof is shown in Appendix E.

Proposition 4 (Properties of Error-Exponent Functions)**.** *Consider a classical-quantum channel* $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ *with* $R_\infty < C_\mathcal{W}$*. We have*

(a) Given every $P \in \mathcal{P}(\mathcal{X})$*,* $E_{\text{sp}}^{(2)}(\cdot,P)$ *is convex and non-increasing on* $[0,+\infty]$*, and continuous on* $\left[I_0^{(2)}(P,\mathcal{W}), +\infty\right]$*. For every* $R > R_\infty$*,* $E_{\text{sp}}^{(2)}(R,\cdot)$ *is continuous on* $\mathcal{P}(\mathcal{X})$*. Further,*

$$E_{\text{sp}}^{(2)}(R,P) = \begin{cases} +\infty, & R < I_0^{(2)}(P,\mathcal{W}) \\ 0, & R \geq I_1^{(2)}(P,\mathcal{W}) \end{cases}. \tag{47}$$

(b) $E_{\text{sp}}(\cdot)$ *is convex and non-increasing on* $[0,+\infty]$*, and continuous on* $[R_\infty,+\infty]$*. Further,*

$$E_{\text{sp}}(R) = \begin{cases} +\infty, & R < R_\infty \\ 0, & R \geq C_\mathcal{W} \end{cases}. \tag{48}$$

(c) Consider any $R \in (R_\infty, C_\mathcal{W})$ *and* $P \in \mathcal{P}_R(\mathcal{X})$ *(see Eq. (44)). The function* $E_{\text{sp}}^{(2)}(\cdot,P)$ *is differentiable with*

$$s_{R,P}^\star = -\left.\frac{\partial E_{\text{sp}}^{(2)}(r,P)}{\partial r}\right|_{r=R} \in \mathbb{R}_{>0}, \tag{49}$$

where $s_{R,P}^\star := (1-\alpha_{R,P}^\star)/\alpha_{R,P}^\star$*, and* $\alpha_{R,P}^\star$ *is the optimizer in Eq. (38).*

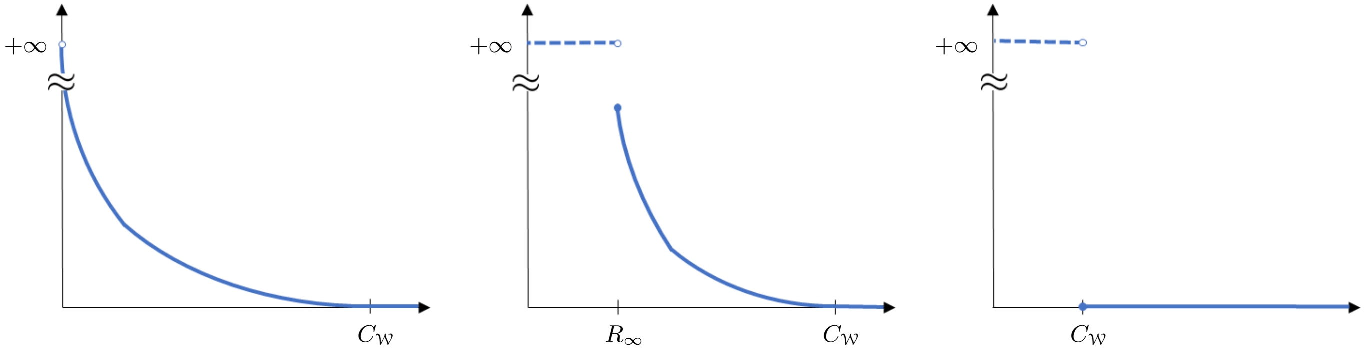
(d) $s_{R,(\cdot)}^\star$ *in Eq. (49) is continuous on* $\mathcal{P}_R(\mathcal{X})$*.*

Given any $R \in (R_\infty, C_\mathcal{W})$ and $P \in \mathcal{P}_R(\mathcal{X})$, we denote a *maximum absolute value subgradient* of the sphere-packing exponent at R by

$$\left|E_{\text{sp}}'(R)\right| := \max_{P: E_{\text{sp}}^{(2)}(R,P) = E_{\text{sp}}(R)} s_{R,P}^\star. \tag{50}$$

Note that the term $\left|E_{\text{sp}}'(R)\right|$ in Eq. (50) is well-defined and finite by item (d) in Proposition 4.

Figure 1 below depicts different cases of the $E_{\text{sp}}(R)$ over rate R.



(A) $0 = R_\infty < C_{\mathcal{W}}$.

(B) $0 < R_\infty < C_{\mathcal{W}}$.

(C) $0 < R_\infty = C_{\mathcal{W}}$.

FIGURE 1. This figure illustrates three cases of the strong sphere-packing exponent $E_{\mathrm{sp}}(R)$ over $R \geq 0$. In the first case $0 = R_\infty < C_{\mathcal{W}}$ (the left figure), $E_{\mathrm{sp}}(R)$ is only infinite at $R = 0$ and finite otherwise. In the second case $0 < R_\infty < C_{\mathcal{W}}$ (the central figure), $E_{\mathrm{sp}}(R) = +\infty$ for $R < R_\infty$, and $E_{\mathrm{sp}}(R) < +\infty$ for $R \geq R_\infty$. In the third case $0 < R_\infty = C_{\mathcal{W}}$ (the right figure), $E_{\mathrm{sp}}(R) = +\infty$ for $R < C_{\mathcal{W}}$, and $E_{\mathrm{sp}}(R) = 0$ for $R \geq C_{\mathcal{W}}$. Without loss of generality, we assume $R_\infty < C_{\mathcal{W}}$ to exclude the last case throughout this paper.

2.2. **Quantum Hypothesis Testing and Channel Coding.** Consider a binary hypothesis whose null and alternative hypotheses are $\rho \in \mathcal{S}(\mathcal{H})$ and $\sigma \in \mathcal{S}(\mathcal{H})$, respectively. The *type-I error* and *type-II error* of the hypothesis testing, for an operator $0 \leq Q \leq \mathbb{1}$, are defined as:

$$\alpha\left(Q;\rho\right) := \mathrm{Tr}\left[(\mathbb{1} - Q)\rho\right], \tag{51}$$

$$\beta\left(Q;\sigma\right) := \mathrm{Tr}\left[Q\sigma\right]. \tag{52}$$

There is a trade-off relation between these two errors. Thus we can define the minimum Type-I error when the type-II error is below $\mu \in (0,1)$ as

$$\widehat{\alpha}_\mu\left(\rho\|\sigma\right) := \min_{0 \leq Q \leq \mathbb{1}} \left\{\alpha\left(Q;\rho\right) : \beta\left(Q;\sigma\right) \leq \mu\right\}. \tag{53}$$

We define an error-exponent function [52, 53, 54] for two sequences of states

$$\mathsf{H}_0 : \rho^n = \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n, \tag{54}$$

$$\mathsf{H}_1 : \sigma^n = \sigma_1 \otimes \sigma_2 \otimes \cdots \otimes \sigma_n, \tag{55}$$

by

$$\phi_n\left(r|\rho^n\|\sigma^n\right) := \sup_{\alpha \in (0,1]} \left\{\frac{1-\alpha}{\alpha}\left(\frac{1}{n}D_\alpha\left(\rho^n\|\sigma^n\right) - r\right)\right\}, \quad r \geq 0. \tag{56}$$

It is known that [54, Lemma 4]

$$\phi_n\left(r|\rho^n\|\sigma^n\right) = +\infty, \quad \forall r \in \left[0, -\frac{1}{n}D_0\left(\rho^n\|\sigma^n\right)\right). \tag{57}$$

Let $\mathcal{M}$ be a finite alphabetical set with size $M = |\mathcal{M}|$. An (n-block) *encoder* is a map $f_n : \mathcal{M} \to \mathcal{X}^n$ that encodes each message $m \in \mathcal{M}$ to a codeword $\mathbf{x}^n(m) := x_1(m)x_2(m)\dots x_n(m) \in \mathcal{X}^n$. The codeword $\mathbf{x}^n(m)$ is then mapped to a state

$$W_{\mathbf{x}^n(m)}^{\otimes n} = W_{x_1(m)} \otimes W_{x_2(m)} \otimes \cdots \otimes W_{x_n(m)} \in \mathcal{S}(\mathcal{H}^{\otimes n}). \tag{58}$$

The *decoder* is described by a positive operator-valued measurement (POVM) $\Pi_n = \{\Pi_{n,1}, \dots, \Pi_{n,M}\}$ on $\mathcal{H}^{\otimes n}$, where $\Pi_{n,i} \geq 0$ and $\sum_{i=1}^M \Pi_{n,i} = \mathbb{1}$. The pair $(f_n, \Pi_n) =: \mathcal{C}_n$ is called a *code* with *rate* $R = \frac{1}{n}\log|\mathcal{C}_n| = \frac{1}{n}\log M$. The error probability of sending a message m with the code $\mathcal{C}_n$ is $\epsilon_m(\mathcal{C}_n) := 1 - \mathrm{Tr}\left(\Pi_{n,m}W_{\mathbf{x}^n(m)}\right)$. We use $\epsilon_{\max}(\mathcal{C}_n) = \max_{m \in \mathcal{M}} \epsilon_m(\mathcal{C}_n)$ and $\bar{\epsilon}(\mathcal{C}_n) = \frac{1}{M}\sum_{m \in \mathcal{M}} \epsilon_m(\mathcal{C}_n)$ to denote the *maximal* error probability and the *average* error probability, respectively. Given a sequence $\mathbf{x}^n \in \mathcal{X}^n$, we denote by

$$P_{\mathbf{x}^n}(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\left\{x = x_i\right\} \tag{59}$$

the empirical distribution of $\mathbf{x}^n$, where x_i is the i-th position of $\mathbf{x}^n$. A constant composition code with a composition $P_{\mathbf{x}^n}$ refers to a codebook whose codewords all have the same distribution $P_{\mathbf{x}^n}$.

Denote by $\epsilon^*\left(n, R\right)$ the smallest average probability of error among all the coding strategies with a blocklengh n and coding rate R. The reliability function of the channel $\mathcal{W}$ and the coding rate R is defined by[4]

$$E(R) := \limsup_{n\to+\infty} -\frac{1}{n}\log \epsilon^*\left(n, R\right). \tag{60}$$

Winter [16] and Dalai [17] showed that the reliability function of a c-q channel can be upper bounded by $E(R) \leq \widetilde{E}_{\text{sp}}(R)$ and $E(R) \leq E_{\text{sp}}(R)$, respectively.

2.3. **Nussbaum-Szkoła Distributions.** Assume the dimension of the Hilbert space $\mathcal{H}$ is d. Given density operators $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ with spectral decompositions

$$\rho = \sum_{i\in[d]} \lambda_i |x_i\rangle\langle x_i|, \quad \text{and} \quad \sigma = \sum_{j\in[d]} \gamma_j |y_j\rangle\langle y_j|, \tag{61}$$

we define the *Nussbaum-Szkoła distributions* [55] $p^{\rho,\sigma}, q^{\rho,\sigma}$ as

$$p^{\rho,\sigma}(i,j) := \lambda_i |\langle x_i|y_j\rangle|^2, \quad q^{\rho,\sigma}(i,j) := \gamma_j |\langle x_i|y_j\rangle|^2. \tag{62}$$

The distributions $p^{\rho,\sigma}, q^{\rho,\sigma}$ have the same mathematical properties as the density operators ρ, σ in some cases, and thus are useful in the sequel. First, one can verify that [55, 40],

$$D_\alpha\left(\rho\|\sigma\right) = D_\alpha\left(p^{\rho,\sigma}\|q^{\rho,\sigma}\right), \quad \forall \alpha \in [0,1]. \tag{63}$$

Second, for product states $\rho_1 \otimes \rho_2$ and $\sigma_1 \otimes \sigma_2$, we have

$$p^{\rho_1\otimes\rho_2, \sigma_1\otimes\sigma_2} = p^{\rho_1,\sigma_1} \otimes p^{\rho_2,\sigma_2}, \quad \text{and} \quad q^{\rho_1\otimes\rho_2, \sigma_1\otimes\sigma_2} = q^{\rho_1,\sigma_1} \otimes q^{\rho_2,\sigma_2}. \tag{64}$$

Third, $\rho \ll \sigma$ if and only if $p^{\rho,\sigma} \ll q^{\rho,\sigma}$. Moreover, we will use ω to represent the pair of indices (i,j) in Eq. (62), and view the distributions $p^{\rho,\sigma}, q^{\rho,\sigma}$ as diagonal matrices, e.g. $\operatorname{Tr}\left[p^{\rho,\sigma}\right] = \sum_{\omega\in[d]\times[d]} p^{\rho,\sigma}(\omega)$.

3. Relation between the Strong and Weak Sphere-Packing Exponents

This section derives alternative formulations of the strong and weak sphere-packing exponents of Eqs. (2)-(3), and provides a relation between these two exponents. As we will show later, the derived formulations are essentially optimization problems in the primal domain, while the expressions in Eqs. (2) and (3) are corresponding dual representations.

We first consider the following convex optimization problem and then exploit it to establish variational formulations of the sphere-packing exponents. Let $\rho, \tau \in \mathcal{S}(\mathcal{H})$ be two density operators. Consider the following convex optimization problem:

$$\begin{aligned} \text{(P)} \quad e(r) := & \inf_{\sigma\in\mathcal{S}(\mathcal{H})} D\left(\sigma\|\rho\right), \\ & \text{subject to} \quad D\left(\sigma\|\tau\right) \leq r. \end{aligned} \tag{65}$$

The above primal problem is interpreted as finding the optimal operator $\sigma^\star$ that achieves the minimum relative entropy $e(r)$ to ρ, within r-radius to τ. The following result shows the dual representation of problem (P) via Lagrangian duality.

Lemma 5 ([52, Section 3.7], [56], [44, Theorem III.5])**.** *The dual problem of* (P) *is given by*

$$\text{(D)} \quad \sup_{s\geq 0}\left\{-(1+s)\log Q^\flat_{\frac{1}{1+s}}\left(\rho\|\tau\right) - sr\right\}. \tag{66}$$

[4] Throughout this paper, we skip the dependence of the channel $\mathcal{W}$ in the reliability function and error-exponent functions.

Proof. By the method of Lagrange multipliers, the primal problem in Eq. (65) can be rewritten as

$$\sup_{s\geq 0} \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \{D(\sigma\|\rho) + s\,(D(\sigma\|\tau) - r)\} \tag{67}$$

$$= \sup_{s\geq 0} \left\{(1+s) \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \left\{\frac{1}{1+s}D(\sigma\|\rho) + \frac{s}{1+s}D(\sigma\|\tau)\right\} - sr\right\} \tag{68}$$

$$= \sup_{s\geq 0} \left\{-(1+s)\log Q^{\flat}_{\frac{1}{1+s}}(\rho\|\tau) - sr\right\}, \tag{69}$$

where the last equality follows from [44, Theorem III.5]. □

Theorem 6 (Variational Representations of the Sphere-Packing Exponents)**.** *Let $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ be a classical-quantum channel. For any $R > R_\infty$, we have*

$$\widetilde{E}_{\mathrm{sp}}(R,P) = \sup_{0<\alpha\leq 1} \min_{\sigma\in\mathcal{S}(\mathcal{H})} \left\{\frac{1-\alpha}{\alpha}\left(D^{\flat}_{\alpha}\left(\mathcal{W}\|\sigma|P\right) - R\right)\right\}, \quad and \tag{70}$$

$$E_{\mathrm{sp}}(R,P) \leq \sup_{0<\alpha\leq 1} \min_{\sigma\in\mathcal{S}(\mathcal{H})} \left\{\frac{1-\alpha}{\alpha}\left(D_{\alpha}\left(\mathcal{W}\|\sigma|P\right) - R\right)\right\}, \tag{71}$$

where $\widetilde{E}_{\mathrm{sp}}(R,P)$ and $E_{\mathrm{sp}}(R,P)$ are defined in Eqs. (36) and (33), respectively.

Moreover, equality in Eq. (71) is attained when maximizing over all prior distributions, i.e.,

$$E_{\mathrm{sp}}(R) = \max_{P\in\mathcal{P}(\mathcal{X})} E_{sp}(R,P) = \max_{P\in\mathcal{P}(\mathcal{X})} \sup_{0<\alpha\leq 1} \min_{\sigma\in\mathcal{S}(\mathcal{H})} \left\{\frac{1-\alpha}{\alpha}\left(D_{\alpha}\left(\mathcal{W}\|\sigma|P\right) - R\right)\right\}. \tag{72}$$

Proof. We start with the proof of Eq. (70). Observe that

$$\min_{\sigma\in\mathcal{S}(\mathcal{H})} D\left(\mathcal{V}\|\sigma|P\right) = \min_{\sigma\in\mathcal{S}(\mathcal{H})} \sum_{x\in\mathcal{X}} P(x)\operatorname{Tr}\left[V_x\left(\log V_x - \log\sigma\right)\right] \tag{73}$$

$$= I(P,\mathcal{V}). \tag{74}$$

We find

$$\widetilde{E}_{\mathrm{sp}}(R,P) = \min_{\mathcal{V}:\mathcal{X}\to\mathcal{S}(\mathcal{H})} \left\{D\left(\mathcal{V}\|\mathcal{W}|P\right) : I(P,\mathcal{V}) \leq R\right\} \tag{75}$$

$$= \min_{\mathcal{V}:\mathcal{X}\to\mathcal{S}(\mathcal{H})} \left\{D\left(\mathcal{V}\|\mathcal{W}|P\right) : \min_{\sigma\in\mathcal{S}(\mathcal{H})} D\left(\mathcal{V}\|\sigma|P\right) \leq R\right\} \tag{76}$$

$$= \sup_{s\geq 0} \min_{\mathcal{V}:\mathcal{X}\to\mathcal{S}(\mathcal{H})} \left\{D\left(\mathcal{V}\|\mathcal{W}|P\right) + s\left(\min_{\sigma\in\mathcal{S}(\mathcal{H})} D\left(\mathcal{V}\|\sigma|P\right) - R\right)\right\} \tag{77}$$

$$= \sup_{s\geq 0} \min_{\sigma\in\mathcal{S}(\mathcal{H})} \min_{\mathcal{V}:\mathcal{X}\to\mathcal{S}(\mathcal{H})} \left\{-sR + \sum_{x\in\mathcal{X}} P(x)D\left(V_x\|W_x\right) + s\cdot D\left(V_x\|\sigma\right)\right\} \tag{78}$$

$$= \sup_{s\geq 0} \min_{\sigma\in\mathcal{S}(\mathcal{H})} \left\{\sum_{x\in\mathcal{X}} P(x) \min_{V_x\in\mathcal{S}(\mathcal{H})} \left[D\left(V_x\|W_x\right) + s\cdot D\left(V_x\|\sigma\right) - sR\right]\right\} \tag{79}$$

$$= \min_{\sigma\in\mathcal{S}(\mathcal{H})} \left\{\sum_{x\in\mathcal{X}} P(x) \min_{V_x\in\mathcal{S}(\mathcal{H})} \left\{D\left(V_x\|W_x\right) : D\left(V_x\|\sigma\right) \leq R\right\}\right\}. \tag{80}$$

In Eq. (77) we introduced the constraint into the objective function via the Lagrange multiplier $s \geq 0$; and Eq. (79) follows from the linearity of the convex combination. By Lemma 5, the inner minimum over

$V_x \in \mathcal{S}(\mathcal{H})$ can be represented as its dual problem:

$$\widetilde{E}_{\mathrm{sp}}(R,P) = \min_{\sigma\in\mathcal{S}(\mathcal{H})} \sup_{s\geq 0} \left\{ -(1+s) \sum_{x\in\mathcal{X}} P(x) \log\left[Q^\flat_{\frac{1}{1+s}}\left(W_x \| \sigma\right) \right] - sR \right\} \tag{81}$$

$$= \min_{\sigma\in\mathcal{S}(\mathcal{H})} \sup_{0<\alpha\leq 1} \left\{ \frac{-\sum_{x\in\mathcal{X}} P(x) \log\left[Q^\flat_{\alpha}\left(W_x \| \sigma\right) \right] - (1-\alpha)R}{\alpha} \right\}, \tag{82}$$

where we substitute $\alpha = 1/(1+s)$. From Lemma 1, the numerator in the bracket of Eq. (82) is a concave-convex saddle function for every $\sigma \in \mathcal{S}(\mathcal{H})$ and every $\alpha \in (0,1]$. Hence, we invoke the minimax theorem, Proposition 7 below, to exchange the order of min-sup in Eq. (82):

$$\widetilde{E}_{\mathrm{sp}}(R,P) = \sup_{0<\alpha\leq 1} \min_{\sigma\in\mathcal{S}(\mathcal{H})} \left\{ \frac{-\sum_{x\in\mathcal{X}} P(x) \log\left[Q^\flat_{\alpha}\left(W_x \| \sigma\right) \right] - (1-\alpha)R}{\alpha} \right\} \tag{83}$$

$$= \sup_{0<\alpha\leq 1} \min_{\sigma\in\mathcal{S}(\mathcal{H})} \left\{ \frac{1-\alpha}{\alpha} \left(D^\flat_\alpha\left(\mathcal{W} \| \sigma | P\right) - R \right) \right\}, \tag{84}$$

where in (84) we recall the definition of the log-Euclidean α-Rényi divergence, Eq. (12), and hence prove the first claim in Eq. (70).

Next, we will prove Eq. (71). From Jensen's inequality and the concavity of the logarithm, the right-hand side of Eq. (71) implies that

$$\sup_{0<\alpha\leq 1} \min_{\sigma\in\mathcal{S}(\mathcal{H})} \left\{ \frac{1-\alpha}{\alpha} \left(\sum_{x\in\mathcal{X}} P(x) D_\alpha\left(W_x \| \sigma\right) - R \right) \right\} \tag{85}$$

$$= \sup_{0<\alpha\leq 1} \min_{\sigma\in\mathcal{S}(\mathcal{H})} \left\{ -\frac{1}{\alpha} \sum_{x\in\mathcal{X}} P(x) \log \operatorname{Tr}\left[W_x^\alpha \sigma^{1-\alpha} \right] - \frac{1-\alpha}{\alpha} R \right\} \tag{86}$$

$$\geq \sup_{0<\alpha\leq 1} \min_{\sigma\in\mathcal{S}(\mathcal{H})} \left\{ -\frac{1}{\alpha} \log \operatorname{Tr}\left[\sum_{x\in\mathcal{X}} P(x) \left[W_x^\alpha \sigma^{1-\alpha} \right] \right] - \frac{1-\alpha}{\alpha} R \right\} \tag{87}$$

$$= E_{\mathrm{sp}}(R,P), \tag{88}$$

where the last equality follows from Eq. (39).

Finally, Eq. (72) follows from the following identity proved by Mosonyi and Ogawa [44, Proposition IV.2]:

$$\max_{P\in\mathcal{P}(\mathcal{X})} \min_{\sigma\in\mathcal{S}(\mathcal{H})} D_\alpha\left(\mathcal{W} \| \sigma | P\right) = \max_{P\in\mathcal{P}(\mathcal{X})} \min_{\sigma\in\mathcal{S}(\mathcal{H})} D_\alpha\left(P \circ \mathcal{W} \| P \otimes \sigma\right), \tag{89}$$

Note that the above relation also holds for $D^\flat_\alpha$.

Proposition 7 ([49, Proposition 21])**.** *Let $\mathcal{A} \subset \mathbb{R}_{\geq 0}$ be a convex set and let $\mathcal{B}$ be a compact Hausdorff space. Further, let $f : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$ be concave on $\mathcal{A}$ as well as convex on $\mathcal{B}$. Then*

$$\sup_{x\in\mathcal{A}} \inf_{y\in\mathcal{B}} \frac{f(x,y)}{x} = \inf_{y\in\mathcal{B}} \sup_{x\in\mathcal{A}} \frac{f(x,y)}{x}. \tag{90}$$

□

The following corollary is a simple consequence of the variational representations of the sphere-packing exponents in Theorem 6 and Eq. (14) .

Corollary 8. *For any classical-quantum channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$, $R > R_\infty$, and $P \in \mathcal{P}(\mathcal{X})$, it holds that*

$$E_{\mathrm{sp}}(R,P) \leq \widetilde{E}_{\mathrm{sp}}(R,P). \tag{91}$$

4. The Refined Strong Sphere-Packing Bound

The main result in the section is a refined strong sphere-packing bound for c-q channels with a polynomial pre-factor (Theorem 9), improving upon a subexponential pre-factor obtained in [17]. To establish this result, we combine Blahut's insight of relating a channel coding problem to binary hypothesis testing [10, 57] with a sharp concentration inequality employed in Ref. [32]. Our proof consists of three major steps: (i) reduce the channel coding problem to binary hypothesis testing (Lemma 11); (ii) bound its type-I error from below (Propositions 12 and 14); (iii) employ Theorem 6 to relate the derived bound to the strong sphere-packing exponent.

Theorem 9 (Refined Strong Sphere-Packing Bound of Constant Composition Codes)**.** *Consider a classical-quantum channel $\mathcal{W}:\mathcal{X}\to\mathcal{S}(\mathcal{H})$ and $R\in(R_\infty, C_\mathcal{W})$. For every $\gamma>0$, there exist an $N_0\in\mathbb{N}$ and a constant $A>0$ such that for all constant composition codes $\mathcal{C}_n$ of length $n\geq N_0$ with message size $|\mathcal{C}_n|\geq\exp\{nR\}$, we have*

$$\bar{\epsilon}\left(\mathcal{C}_n\right)\geq\frac{A}{n^{\frac{1}{2}\left(1+|E_{\text{sp}}'(R)|+\gamma\right)}}\exp\left\{-nE_{\text{sp}}(R)\right\}. \tag{92}$$

The following corollary generalizes the refined sphere-packing bound for constant composition codes to arbitrary codes via a standard argument [7, p. 95].

Corollary 10 (Refined Strong Sphere-Packing Bound for General Codes)**.** *Consider a classical-quantum channel $\mathcal{W}:\mathcal{X}\to\mathcal{S}(\mathcal{H})$ and $R\in(R_\infty, C_\mathcal{W})$. There exist some $t>1/2$ and $N_0\in\mathbb{N}$ such that for all codes of length $n\geq N_0$, we have*

$$\epsilon^*\left(n,R\right)\geq n^{-t}\exp\left\{-nE_{\text{sp}}(R)\right\}. \tag{93}$$

Proofs for Theorem 9 and Corollary 10 are provided in Section 4.2.

Theorem 9 yields

$$\log\frac{1}{\bar{\epsilon}(\mathcal{C}_n)}\leq nE_{\text{sp}}(R)+\frac{1}{2}\left(1+\left|E_{\text{sp}}'(R)\right|\right)\log n+o(\log n), \tag{94}$$

where the term $\frac{1}{2}\left(1+\left|E_{\text{sp}}'(R)\right|\right)$ can be viewed as a second-order term (see the discussions in [58, Section 4.4]). On the other hand, for the case of classical *non-singular* channels[5], it was shown that [24, Theorem 3.6], for all constant composition codes $\mathcal{C}_n$ and rate $R\in(R_{\text{crit}}, C_\mathcal{W})$,

$$\log\frac{1}{\bar{\epsilon}(\mathcal{C}_n)}\geq nE_{\text{r}}(R)+\frac{1}{2}\left(1+\left|E_{\text{r}}'(R)\right|\right)\log n+\Omega(1), \tag{95}$$

where $E_\text{r}(R)$ is the *random coding exponent*, and R_crit is the critical rate such that $E_\text{r}(R)=E_\text{sp}(R)$ for all $R\geq R_\text{crit}$ [6, p. 160], [15]. Hence our result, Theorem 9, matches the achievability up to the logarithmic order. We note that whether the third order $o(\log n)$ in Eq. (94) can be improved to $O(1)$ is still unknown even for the classical case.

4.1. **Converse Bounds for Quantum Hypothesis Testing.** This section contains the hypothesis testing reduction method (Lemma 11) and two converse bounds (Propositions 12 and 14). We first present a proof that relates the decoding error of a code to binary hypothesis testing. We note that Lemma 11 below is similar to the meta-converse in Ref. [60]. However, the idea dates back to Blahut [10].

Lemma 11. *For any classical-quantum channel $W:\mathcal{X}\to\mathcal{S}(\mathcal{H})$ and any code $\mathcal{C}_n$ with message size M, it follows that*

$$\epsilon_{\max}\left(\mathcal{C}_n\right)\geq\max_{\sigma\in\mathcal{S}(H)}\min_{\mathbf{x}^n\in\mathcal{C}_n}\widehat{\alpha}_{\frac{1}{M}}\left(W_{\mathbf{x}^n}^{\otimes n}\|\sigma^{\otimes n}\right). \tag{96}$$

[5] For classical *singular* channels, one has $\log\frac{1}{\bar{\epsilon}(\mathcal{C}_n)}\geq nE_\text{r}(R)+\frac{1}{2}\log n+\Omega(1)$ [24]. Further, it was conjectured that [59] that $\log\frac{1}{\bar{\epsilon}(\mathcal{C}_n)}\leq nE_\text{sp}(R)+\frac{1}{2}\log n+o(\log n)$, for all asymmetric classical singular channels and constant composition codes. However, such a result remains open.

Proof. Let $\mathbf{x}^n(m)$ be the codeword encoding the message $m \in \{1, \dots, M\}$. Define a binary hypothesis testing problem:

$$\mathsf{H}_0 : W^{\otimes n}_{\mathbf{x}^n(m)}, \tag{97}$$

$$\mathsf{H}_1 : \sigma^n := \bigotimes_{i=1}^{n} \sigma_i, \tag{98}$$

where $\sigma^n \in \mathcal{S}\left(\mathcal{H}^{\otimes n}\right)$ can be viewed as a dummy channel output. Since $\sum_{m=1}^{M} \beta\left(\Pi_{n,m}; \sigma^n\right) = 1$ for any POVM $\Pi_n = \{\Pi_{n,1}, \dots, \Pi_{n,M}\}$, and $\beta\left(\Pi_{n,m}; \sigma^n\right) \geq 0$ for every $m \in \mathcal{M}$, there must exist a message $m \in \mathcal{M}$ for any code $\mathcal{C}_n$ such that $\beta\left(\Pi_{n,m}; \sigma^n\right) \leq \frac{1}{M}$. Fix $\mathbf{x}^n := \mathbf{x}^n\left(m\right)$. Then

$$\epsilon_{\max}\left(\mathcal{C}_n\right) \geq \epsilon_m\left(\mathcal{C}_n\right) = \alpha\left(\Pi_{n,m}; W^{\otimes n}_{\mathbf{x}^n}\right) \geq \widehat{\alpha}_{\frac{1}{M}}\left(W^{\otimes n}_{\mathbf{x}^n} \| \sigma^n\right). \tag{99}$$

Since the above inequality (99) holds for every $\sigma^n \in \mathcal{S}\left(\mathcal{H}^{\otimes n}\right)$, it follows that

$$\epsilon_{\max}\left(\mathcal{C}_n\right) \geq \max_{\sigma \in \mathcal{S}(H)} \min_{\mathbf{x}^n \in \mathcal{C}_n} \widehat{\alpha}_{\frac{1}{M}}\left(W^{\otimes n}_{\mathbf{x}^n} \| \sigma^{\otimes n}\right). \tag{100}$$

□

In the following Proposition, we generalize Blahut's one-shot converse Hoeffding bound [10, Theorem 10] to the quantum setting. This result is essentially a Chebyshev-type bound. We will employ it to lower bound the error of "bad sequences" that yield smaller error exponent in Section 4.2.

Proposition 12 (One-Shot Converse Hoeffding Bound)**.** *Consider the following binary hypothesis testing problem:* $\mathsf{H}_0 : \rho$ *versus* $\mathsf{H}_1 : \sigma$*, where* $\rho, \sigma \in \mathcal{S}(\mathcal{H})$*. For every* $r \geq 0$ *and* $\nu > 0$*, we have*

$$\widehat{\alpha}_{\frac{1}{4}\exp\{-(r+\nu)\}}\left(\rho \| \sigma\right) \geq \frac{1}{2}\left(\frac{1}{2} - \frac{K(\rho,\sigma)}{\nu^2}\right) \exp\left\{-\nu - \phi\left(r|\rho\|\sigma\right)\right\} \tag{101}$$

where

$$\phi\left(r|\rho\|\sigma\right) := \sup_{\alpha \in (0,1]} \left\{ \frac{1-\alpha}{\alpha}\left(D_\alpha\left(\rho\|\sigma\right) - r\right)\right\}, \tag{102}$$

and

$$K(\rho,\sigma) := V\left(\hat{q}_t \| q\right) + V\left(\hat{q}_t \| p\right) \in \mathbb{R}_{\geq 0}, \tag{103}$$

where (p,q) *are the Nussbaum-Szkoła distributions of* (ρ,σ)*, and*

$$\hat{q}_t(\omega) = \frac{p^{1-t}(\omega) q^t(\omega)}{\sum_{\omega \in \mathsf{supp}(p) \cap \mathsf{supp}(q)} p^{1-t}(\omega) q^t(\omega)}, \quad \omega \in \mathsf{supp}(p) \cap \mathsf{supp}(q) \tag{104}$$

for some $t \in [0,1]$*.*

Proof. If ρ and σ have disjoint supports, then Eq. (101) trivially holds since $D_\alpha(\rho\|\sigma) = +\infty$ for all $\alpha \in [0,1]$. Hence, we assume ρ and σ have non-disjoint support in the following. Let $\mathcal{B} := \mathsf{supp}(p) \cap \mathsf{supp}(q)$ be the intersection of the joint support of p and q. Fix $\phi(r) := \phi(r|\rho\|\sigma) = \phi(r|p\|q)$ since $D_\alpha(\rho\|\sigma) = D_\alpha(p\|q)$.

For any test $0 \leq Q \leq \mathbb{1}$, Nagaoka showed that [56, Lemma 1] (see also [54, Proposition 2], [55]):

$$\alpha\left(Q;\rho\right) + \delta\beta\left(Q;\sigma\right) \geq \frac{1}{2}\left(\sum_{\omega : p(\omega) \leq \delta q(\omega)} p(\omega) + \sum_{\omega : p(\omega) > \delta q(\omega)} \delta q(\omega)\right), \quad \forall \delta \geq 0. \tag{105}$$

Let $r > 0$, $\delta = e^{r-\phi(r)}$, and $\mu \geq 0$ that will be specified later. Eq. (105) implies that

$$\begin{aligned} \widehat{\alpha}_\mu\left(\rho\|\sigma\right) &\geq \frac{1}{2}\left(\sum_{\omega : p(\omega)\mathrm{e}^{\phi(r)} \leq q(\omega)\mathrm{e}^{r}} p(\omega) + \sum_{\omega : p(\omega)\mathrm{e}^{\phi(r)} > q(\omega)\mathrm{e}^{r}} \mathrm{e}^{r-\phi(r)} q(\omega)\right) - \mathrm{e}^{r-\phi(r)}\mu \qquad (106) \\ &\geq \frac{1}{2}\left(\sum_{\omega \in \mathcal{U}_1(\nu)} p(\omega) + \sum_{\omega \in \mathcal{U}_2(\nu)} \mathrm{e}^{r-\phi(r)} q(\omega)\right) - \mathrm{e}^{r-\phi(r)}\mu, \qquad (107) \end{aligned}$$

where in the last line we introduce the decision regions for some $\nu > 0$:

$$\mathcal{U}_1(\nu) := \left\{\omega : \hat{q}_t(\omega)\mathrm{e}^{-\nu} < p(\omega)\mathrm{e}^{\phi(r)} \le q(\omega)\mathrm{e}^{r}\right\}, \quad \mathcal{U}_2(\nu) := \left\{\omega : \hat{q}_t(\omega)\mathrm{e}^{-\nu} < q(\omega)\mathrm{e}^{r} < p(\omega)\mathrm{e}^{\phi(r)}\right\}, \tag{108}$$

and $\hat{q}_t$ is the *tilted distribution* (see [10, Theorem 4]):

$$\hat{q}_t(\omega) = \frac{p^{1-t}(\omega)q^t(\omega)}{\sum_{\omega\in\mathcal{B}} p^{1-t}(\omega)q^t(\omega)}, \quad \omega \in \mathcal{B} \tag{109}$$

for some $t \in [0,1]$ such that $\hat{q}_t$ satisfies

$$D\left(\hat{q}_t\|p\right) = \phi\left(r\right) \quad \text{and} \quad D\left(\hat{q}_t\|q\right) = r. \tag{110}$$

In the following, we are going to lower bound the right-hand side of Eq. (107) in terms of $\hat{q}_t$. From Eq. (108), we find

$$\begin{aligned} \sum_{\omega\in\mathcal{U}_1(\nu)} p(\omega) &\ge \mathrm{e}^{-(\phi(r)+\nu)} \sum_{\omega\in\mathcal{U}_1(\nu)} \hat{q}_t(\omega); \\ \sum_{\omega\in\mathcal{U}_2(\nu)} q(\omega) &\ge \mathrm{e}^{-(r+\nu)} \sum_{\omega\in\mathcal{U}_2(\nu)} \hat{q}_t(\omega). \end{aligned} \tag{111}$$

Next, we estimate the error in the union: $\sum_{\omega\in\mathcal{U}_1(\nu)\cup\mathcal{U}_2(\nu)} \hat{q}_t(\omega)$. Let

$$\mathcal{U}_A := \left\{\omega : \hat{q}_t(\omega)\mathrm{e}^{-\nu} < q(\omega)\mathrm{e}^{r}\right\}, \quad \mathcal{U}_B := \left\{\omega : \hat{q}_t(\omega)\mathrm{e}^{-\nu} < p(\omega)\mathrm{e}^{\phi(r)}\right\}. \tag{112}$$

Observe that $\mathcal{U}_1(\nu) \cup \mathcal{U}_2(\nu) = \mathcal{U}_A \cap \mathcal{U}_B$ and

$$\sum_{\omega\in\mathcal{U}_A\cap\mathcal{U}_B} \hat{q}_t(\omega) \ge 1 - \sum_{\omega\in\mathcal{U}_A^{\mathrm{c}}} \hat{q}_t(\omega) - \sum_{\omega\in\mathcal{U}_B^{\mathrm{c}}} \hat{q}_t(\omega). \tag{113}$$

Denote by

$$\begin{aligned} \mathcal{U}_T &:= \left\{\omega : \left|\log\frac{\hat{q}_t(\omega)}{q(\omega)}\mathrm{e}^{-r}\right| \ge \nu\right\} \qquad (114) \\ &= \left\{\omega : \left|\log\frac{\hat{q}_t(\omega)}{q(\omega)} - \sum_{\omega\in\mathcal{B}} \hat{q}_t(\omega)\log\frac{\hat{q}_t(\omega)}{q(\omega)}\right| \ge \nu\right\}, \qquad (115) \end{aligned}$$

where the last equality follows from Eq. (110). Since $\mathcal{U}_A^{\mathrm{c}} \subseteq \mathcal{U}_T$, we apply Chebyshev's inequality to obtain

$$\sum_{\omega\in\mathcal{U}_A^{\mathrm{c}}} \hat{q}_t(\omega) \le \sum_{\omega\in\mathcal{U}_T} \hat{q}_t(\omega) \le \frac{V\left(\hat{q}_t\|q\right)}{\nu^2}. \tag{116}$$

Similarly,

$$\sum_{\omega\in\mathcal{U}_B^{\mathrm{c}}} \hat{q}_t(\omega) \le \frac{V\left(\hat{q}_t\|p\right)}{\nu^2}. \tag{117}$$

Let $K = K(\rho,\sigma) := V\left(\hat{q}_t\|q\right) + V\left(\hat{q}_t\|p\right)$. Equation (113), along with (116) and (117) yields that

$$\sum_{\omega\in\mathcal{U}_1(\nu)\cup\mathcal{U}_2(\nu)} \hat{q}_t(\omega) = \sum_{\omega\in\mathcal{U}_A\cap\mathcal{U}_B} \hat{q}_t(\omega) \ge 1 - \frac{K}{\nu^2}. \tag{118}$$

Hence, from Eqs. (107), (111), and (118), we obtain the lower bound of the type-I error:

$$\widehat{\alpha}_{\mu}\left(\rho\|\sigma\right) \geq \frac{1}{2}\left(\sum_{\omega\in\mathcal{U}_1(\nu)} p(\omega) + \sum_{\omega\in\mathcal{U}_2(\nu)} \mathrm{e}^{r-\phi(r)}q(\omega)\right) - \mathrm{e}^{r-\phi(r)}\mu, \tag{119}$$

$$\geq \frac{1}{2}\mathrm{e}^{-(\phi(r)+\nu)}\left(\sum_{\omega\in\mathcal{U}_1(\nu)} \hat{q}_t(\omega) + \sum_{\omega\in\mathcal{U}_2(\nu)} \hat{q}_t(\omega)\right) - \mathrm{e}^{r-\phi(r)}\mu \tag{120}$$

$$\geq \frac{1}{2}\mathrm{e}^{-(\phi(r)+\nu)}\left(\sum_{\omega\in\mathcal{U}_1(\nu)\cup\mathcal{U}_2(\nu)} \hat{q}_t(\omega)\right) - \mathrm{e}^{r-\phi(r)}\mu \tag{121}$$

$$\geq \frac{1}{2}\mathrm{e}^{-(\phi(r)+\nu)}\left(1-\frac{K}{\nu^2}\right) - \mathrm{e}^{r-\phi(r)}\mu. \tag{122}$$

Choose $\mu = \frac{1}{4}\exp\{-(r+\nu)\}$. Eq. (122) further gives

$$\widehat{\alpha}_{\frac{1}{4}\exp\{-(r+\nu)\}}\left(\rho\|\sigma\right) \geq \frac{1}{2}\mathrm{e}^{-(\phi(r)+\nu)}\left(1-\frac{K}{\nu^2}\right) - \frac{1}{4}\mathrm{e}^{-(\phi(r)+\nu)} \tag{123}$$

$$= \frac{1}{2}\left(\frac{1}{2}-\frac{K}{\nu^2}\right)\mathrm{e}^{-(\phi(r)+\nu)}, \tag{124}$$

which completes the proof. □

Applying Proposition 12 to product states yields the following result.

Proposition 13 (Chebyshev-Type Converse Hoeffding Bound)**.** *Let $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ be a classical-quantum channel, and let $R \in (R_\infty, C_{\mathcal{W}})$. Consider the binary hypothesis testing with sequences*

$$\mathsf{H}_0 : \rho^n = W_{\mathbf{x}^n}^{\otimes n}; \tag{125}$$

$$\mathsf{H}_1 : \sigma^n = \left(\sigma^{\star}_{R,P_{\mathbf{x}^n}}\right)^{\otimes n}, \tag{126}$$

where $\mathbf{x}^n \in \mathcal{X}^n$ and $\sigma^{\star}_{R,P} \in \arg\min_{\sigma\in\mathcal{S}(\mathcal{H})}\sup_{0<\alpha\leq 1}\frac{1-\alpha}{\alpha}\left(D_\alpha\left(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}\right)-R\right)$. Then, for every $c>0$, there exist $N_0 \in \mathbb{N}$ and $\kappa_1, \kappa_2 \in \mathbb{R}_{>0}$ such that for all $n \geq N_0$ we have

$$\widehat{\alpha}_{c\exp\{-nR\}}\left(\rho^n\|\sigma^n\right) \geq \kappa_1 \exp\left\{-\kappa_2\sqrt{n} - nE_{\mathrm{sp}}^{(2)}\left(R, P_{\mathbf{x}^n}\right)\right\}, \tag{127}$$

Remark 4.1. Consider independent and identically distributed (i.i.d.) extensions $\mathsf{H}_0 : \rho^{\otimes n}$ and $\mathsf{H}_1 : \sigma^{\otimes n}$. Proposition 13 then recovers the converse proof of the *quantum Hoeffding bound* (see [56] and [61, Section 5.4]): for $r \in (0, D(\rho\|\sigma))$,

$$\lim_{n\to+\infty} -\frac{1}{n}\log\widehat{\alpha}_{\exp\{-nr\}}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) \leq \sup_{0<\alpha\leq 1}\frac{1-\alpha}{\alpha}\left(D_\alpha(\rho\|\sigma)-r\right). \tag{128}$$

Proof. Denote by $p^n = \bigotimes_{i=1}^n p_{x_i}$, $q^n = \bigotimes_{i=1} q_{x_i}$ Nussbaum-Szkoła distributions of ρ^n and σ^n [55] with joint supports $\mathcal{B}_{x_i} := \mathsf{supp}(p_{x_i}) \cap \mathsf{supp}(q_{x_i})$, $i \in [n]$. Let $R_n := R - \gamma_n$, where $\gamma_n := \frac{\nu+\log 4c}{n}$. Fix an arbitrary $R_0 \in (R_\infty, R)$. Choose an $N_0 \in \mathbb{N}$ such that $R_n \geq R_0$ for all $n \geq N_0$. Consider $n \geq N_0$ onwards. Then, Proposition 12 implies that

$$\widehat{\alpha}_{c\exp\{-nR\}}\left(\rho^n\|\sigma^n\right) \geq \frac{1}{2}\left(\frac{1}{2}-\frac{K(\rho^n,\sigma^n)}{\nu^2}\right)\exp\left\{-\nu - n\phi_n\left(R_n|\rho^n\|\sigma^n\right)\right\} \tag{129}$$

$$= \frac{1}{2}\left(\frac{1}{2}-\frac{K(\rho^n,\sigma^n)}{\nu^2}\right)\exp\left\{-\nu - nE_{\mathrm{sp}}^{(2)}\left(R_n, P_{\mathbf{x}^n}\right)\right\}, \tag{130}$$

where the second equality (130) follows from the saddle-point property, item (a) in Proposition 3. Since the coefficient $K(\rho^n, \sigma^n)$ in Eq. (103) is additive for product states, one has

$$K\left(\rho^n, \sigma^n\right) = V\left(\hat{q}_t^n \| p^n\right) + V\left(\hat{q}_t^n \| q^n\right) \tag{131}$$

$$= n \sum_{x \in \mathcal{X}} P_{\mathbf{x}^n}(x) \left[V\left(\hat{q}_{x,t} \| p_x\right) + V\left(\hat{q}_{x,t} \| q_x\right)\right], \tag{132}$$

where $P_{\mathbf{x}^n}$ is the empirical distribution for the sequence $\mathbf{x}^n$, and $\hat{q}_t^n := \bigotimes_{i=1}^n \hat{q}_{x_i,t}$ is the tilted distribution (see Eqs. (104) and (109)). Note that $\hat{q}_t^n \ll p^n$ and $\hat{q}_t^n \ll q^n$ for all $t \in [0,1]$. This guarantees that the quantity $K(\rho^n, \sigma^n)$ is finite.

Let

$$V_{\max} := \max_{t \in [0,1],\, P_{\mathbf{x}^n} \in \mathcal{P}(\mathcal{X})} \sum_{x \in \mathcal{X}} P_{\mathbf{x}^n}(x) \left[V\left(\hat{q}_{x,t} \| p_x\right) + V\left(\hat{q}_{x,t} \| q_x\right)\right] \in \mathbb{R}_{>0}, \tag{133}$$

we obtain

$$K\left(\rho^n, \sigma^n\right) \leq n V_{\max}. \tag{134}$$

By choosing $\nu = \sqrt{4nV_{\max}}$, Eqs. (130) and (134) give

$$\widehat{\alpha}_{c\exp\{-nR\}}\left(\rho^n \| \sigma^n\right) \geq \frac{1}{8} \exp\left\{-\sqrt{4nV_{\max}} - nE_{\text{sp}}^{(2)}\left(R - \gamma_n, P_{\mathbf{x}^n}\right)\right\}. \tag{135}$$

Finally, we will remove the rate back-off term γ_n in Eq. (135). Recall item (a) in Proposition 4 that the map $r \mapsto E_{\text{sp}}^{(2)}(r, P_{\mathbf{x}^n})$ is convex and monotone decreasing. Further, we assume $E_{\text{sp}}^{(2)}(R_0, P_{\mathbf{x}^n}) > 0$ and thus the $E_{\text{sp}}^{(2)}(\cdot, P_{\mathbf{x}^n})$ is differentiable at R_0 by item (c) in Proposition 4. Otherwise, the monotone decreases imply that $E_{\text{sp}}^{(2)}(R, P_{\mathbf{x}^n}) = E_{\text{sp}}^{(2)}(R_0, P_{\mathbf{x}^n}) = 0$, which already completes the proof. Denoting by ∂_- the left derivative, the convexity then implies that

$$E_{\text{sp}}^{(2)}(R - \gamma_n, P_{\mathbf{x}^n}) \leq E_{\text{sp}}^{(2)}(R, P_{\mathbf{x}^n}) - \gamma_n \partial_- E_{\text{sp}}^{(2)}(R - \gamma_n, P_{\mathbf{x}^n}), \tag{136}$$

$$\leq E_{\text{sp}}^{(2)}(R, P_{\mathbf{x}^n}) - \gamma_n \left.\frac{\partial E_{\text{sp}}^{(2)}(r, P_{\mathbf{x}^n})}{\partial r}\right|_{r=R_0}, \tag{137}$$

where the last inequality (137) follows from the monotone decreases. Let

$$\Upsilon := \max_{P_{\mathbf{x}^n} \in \mathcal{P}(\mathcal{X})} \left| \left.\frac{\partial E_{\text{sp}}^{(2)}(r, P_{\mathbf{x}^n})}{\partial r}\right|_{r=R_0} \right|. \tag{138}$$

Note that $\Upsilon \in \mathbb{R}_{\geq 0}$ due to $R_0 > R_\infty$ and item (d) of Proposition 4. Then, Eqs. (135), (137), and (138) lead to

$$\widehat{\alpha}_{c\exp\{-nR\}}\left(\rho^n \| \sigma^n\right) \geq \frac{1}{8} \exp\left\{-\sqrt{4nV_{\max}} - \gamma_n \Upsilon - nE_{\text{sp}}^{(2)}\left(R, P_{\mathbf{x}^n}\right)\right\}. \tag{139}$$

Setting $\kappa_1 = 1/8$ and choosing a constant $\kappa_2 \in \mathbb{R}_{>0}$ such that $\sqrt{4nV_{\max}} + \gamma_n \Upsilon \leq \kappa_2 \sqrt{n}$ for all $n \geq N_0$ conclude this corollary. □

The following Proposition 14 is a sharp converse bound from Bahadur-Ranga Rao's inequality (see Appendix B). In Section 4.2, we will exploit this result to bound the error of "good sequences" with a polynomial pre-factor.

Proposition 14 (Sharp Converse Hoeffding Bound)**.** *Let $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ be a classical-quantum channel, and let $R \in (R_\infty, C_\mathcal{W})$. Consider the following binary hypothesis testing problem with sequences*

$$\mathsf{H}_0 : \rho^n = W_{\mathbf{x}^n}^{\otimes n}; \tag{140}$$

$$\mathsf{H}_1 : \sigma^n = \left(\sigma_{R,P_{\mathbf{x}^n}}^\star\right)^{\otimes n}, \tag{141}$$

where $\mathbf{x}^n \in \mathcal{X}^n$, and $\sigma_{R,P}^\star := \arg\min_{\sigma \in \mathcal{S}(\mathcal{H})} \sup_{0<\alpha\leq 1} \frac{1-\alpha}{\alpha}\left(D_\alpha\left(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}\right) - R\right)$ satisfying

$$E_{\text{sp}}^{(2)}\left(R, P_{\mathbf{x}^n}\right) \in [\nu, +\infty) \tag{142}$$

for some positive $\nu > 0$. For every $c > 0$, there exists a constant $N_0 \in \mathbb{N}$, independent of the sequences ρ^n and σ^n, such that for all $n \geq N_0$ we have

$$\widehat{\alpha}_{c\exp\{-nR\}}\left(\rho^n \| \sigma^n\right) \geq \frac{A}{n^{\frac{1}{2}\left(1+s^\star_{R,P_{\mathbf{x}^n}}\right)}} \exp\left\{-nE_{\text{sp}}^{(2)}\left(R, P_{\mathbf{x}^n}\right)\right\}, \tag{143}$$

where $s^\star_{R,P} := -\left.\frac{\partial E_{\text{sp}}^{(2)}(r,P)}{\partial r}\right|_{r=R}$, and $A \in \mathbb{R}_{>0}$ is a finite constant depending on R, ν and $\mathcal{W}$.

Proof. Fix an arbitrary $R_0 \in (R_\infty, R)$. Let $\gamma_n := \frac{\log n}{2n} + \frac{x}{n}$ and $R_n := R - \gamma_n$ for some $x \in \mathbb{R}$. The choice of x and the rate back-off term γ_n will become evident later. Let $N_1 \in \mathbb{N}$ such that $R_n \in [R_0, R]$ for all $n \geq N_1$. Subsequently, we choose such $n \geq N_1$ onwards.

Let $p^n := \bigotimes_{i=1}^n p_{x_i}$ and $q^n := \bigotimes_{i=1}^n q_{x_i}$, where (p_{x_i}, q_{x_i}) are Nussbaum-Szkoła distributions [55] of $(W_{x_i}, \sigma^\star)$ for every $i \in [n]$. Since $D_\alpha(\rho_{x_i} \| \sigma_{x_i}) = D_\alpha(p_{x_i} \| q_{x_i})$, for $\alpha \in (0,1]$, again we shorthand

$$\phi_n(R_n) := \phi_n\left(R_n | \rho^n \| \sigma^n\right) = \phi_n(R_n | p^n \| q^n) = E_{\text{sp}}^{(2)}\left(R_n, P_{\mathbf{x}^n}\right), \tag{144}$$

where the last equality in Eq. (144) follows from the saddle-point property, item (a) in Proposition 3. Moreover, item (c) in Proposition 3 implies that the state $\sigma^\star$ dominants all the states: $\sigma^\star \gg W_x$, for all $x \in \texttt{supp}(P_{\mathbf{x}^n})$, Hence, we have $p^n \ll q^n$. Without loss of generality, we set zero all elements of q_{x_i} that do not lie in the support of p_{x_i}, i.e. $q_{x_i}(\omega) = 0$, $\omega \notin \texttt{supp}(p_{x_i})$, $i \in [n]$, because those elements do not contribute in $\phi_n(R_n)$.

Repeating Nagaoka's argument [56] in Eq. (105) for any $0 \leq Q_n \leq \mathbb{1}$ and choosing $\delta = \exp\{nR_n - n\phi_n(R_n)\}$ yields

$$\alpha\left(Q_n; \rho^n\right) + \delta\beta\left(Q_n; \sigma^n\right) \geq \frac{1}{2}\left(\alpha\left(\mathcal{U}; p^n\right) + \mathrm{e}^{nR_n - n\phi_n(R_n)}\beta\left(\mathcal{U}; q^n\right)\right), \tag{145}$$

where $\alpha\left(\mathcal{U}; p^n\right) := \sum_{\omega \in \mathcal{U}^c} p^n(\omega)$, $\beta\left(\mathcal{U}; q^n\right) := \sum_{\omega \in \mathcal{U}} q^n(\omega)$, and

$$\mathcal{U} := \left\{\omega : p^n(\omega)\mathrm{e}^{n\phi_n(R_n)} > q^n(\omega)\mathrm{e}^{nR_n}\right\}. \tag{146}$$

In the following, we will employ Bahadur-Ranga Rao's concentration inequality, Theorem 18, in Appendix B, to further lower bound $\alpha\left(\mathcal{U}; p^n\right)$ and $\beta\left(\mathcal{U}; q^n\right)$. Before proceeding, we need to introduce some notation. Let

$$\begin{aligned} \Lambda_{0,P_{\mathbf{x}^n}}(t) &:= \sum_{x \in \mathcal{X}} P_{\mathbf{x}^n}(x)\Lambda_{0,x_i}(t), \quad \Lambda_{1,P_{\mathbf{x}^n}}(t) := \sum_{x \in \mathcal{X}} P_{\mathbf{x}^n}(x)\Lambda_{0,x_i}(t); \\ \Lambda_{0,x_i}(t) &:= \log \mathbb{E}_{p_{x_i}}\left[\mathrm{e}^{t\log\frac{q_{x_i}}{p_{x_i}}}\right], \quad \Lambda_{1,x_i}(t) := \log \mathbb{E}_{q_{x_i}}\left[\mathrm{e}^{t\log\frac{p_{x_i}}{q_{x_i}}}\right], \end{aligned} \tag{147}$$

and the *Lengendre-Fenchel transform*:

$$\Lambda^*_{j,P_{\mathbf{x}^n}}(z) := \sup_{t \in \mathbb{R}}\left\{tz - \Lambda_{j,P_{\mathbf{x}^n}}(t)\right\}, \quad j \in \{0,1\}. \tag{148}$$

The quantities $\Lambda^*_{j,P_{\mathbf{x}^n}}(z)$ would appear in the lower bounds of $\alpha\left(\mathcal{U}; p^n\right)$ and $\beta\left(\mathcal{U}; q^n\right)$ obtained by Bahadur-Randga Rao's inequality as shown later.

Note that Eqs. (144), (142) and item (a) in Proposition 4 imply that, for all $r \in [R_0, R]$,

$$\phi_n(r) \geq \phi_n(R) \geq \nu > 0. \tag{149}$$

Lemma 17 in Appendix A thus relates the Legendre-Fenchel transform $\Lambda^*_{j,P_{\mathbf{x}^n}}(z)$ to the desired error-exponent function $\phi_n(R_n)$: for all $r \in [R_0, R]$:

$$\Lambda''_{0,P_{\mathbf{x}^n}}(t) > 0, \quad \forall t \in [0,1]; \tag{150}$$

$$\Lambda^*_{0,P_{\mathbf{x}^n}}\left(\phi_n(r) - r\right) = \phi_n(r); \tag{151}$$

$$\Lambda^*_{1,P_{\mathbf{x}^n}}\left(r - \phi_n(r)\right) = r, \tag{152}$$

and there exists a unique optimizer $t^\star := t^\star_{r,P_{\mathbf{x}^n}}$ to the Legendre-Fenchel transform $\Lambda^*_{0,P_{\mathbf{x}^n}}(z)$ with

$$t^\star = \frac{s^\star_{r,P_{\mathbf{x}^n}}}{1+s^\star_{r,P_{\mathbf{x}^n}}} \in (0,1), \tag{153}$$

$$s^\star_{r,P_{\mathbf{x}^n}} = -\frac{\partial \phi_n(r)}{\partial r}. \tag{154}$$

Next, we show that the optimizer $t^\star$ in Eq. (153) can be further bounded in the following region:

$$t^\star \in \left[\frac{\frac{\nu}{\Psi(R,\nu)}}{1+\frac{\nu}{\Psi(R,\nu)}}, 1\right] =: H, \tag{155}$$

where

$$\Psi(R,\nu) := \max_{P_{\mathbf{x}^n}\,:\,\nu\leq\phi_n(R)<+\infty} I_1^{(2)}\left(P_{\mathbf{x}^n}, \mathcal{W}\right) \in \mathbb{R}_{>0}. \tag{156}$$

Owing to $t^\star = \frac{s^\star_{r,P_{\mathbf{x}^n}}}{1+s^\star_{r,P_{\mathbf{x}^n}}}$ in Eq. (153), proving Eq. (155) is equivalent to showing that, whenever $\phi_n(R) \in [\nu, +\infty)$ and $r \in [R_0, R]$:

$$s^\star_{r,P_{\mathbf{x}^n}} \geq \frac{\nu}{\Psi(R,\nu)}. \tag{157}$$

Item (a) in Proposition 4 gives $\phi_n(\Psi(R,\nu)) = 0$ because $I_1^{(2)}(P_{\mathbf{x}^n}, \mathcal{W}) \leq \Psi(R,\nu)$. Continuing from Eq. (154) leads to

$$s^\star_{r,P_{\mathbf{x}^n}} = -\frac{\partial \phi_n(r)}{\partial r} \geq -\frac{\partial \phi_n(r)}{\partial r}\bigg|_{r=R} \geq \frac{\nu - 0}{\Psi(R,\nu)-R} \geq \frac{\nu}{\Psi(R,\nu)}, \tag{158}$$

where the first and second inequalities follow from the fact that $\phi_n(r)$ is convex and non-increasing in r.

Since Eq. (155) shows that the optimizer $t^\star$ always lies in the compact set H, we can define the following quantities:

$$V_{\max}(R,\nu) := \max_{t\in H,\, P_{\mathbf{x}^n}\in\mathcal{P}_{R,\nu}(\mathcal{X})} \Lambda''_{0,P_{\mathbf{x}^n}}(t); \tag{159}$$

$$V_{\min}(R,\nu) := \min_{t\in H,\, P_{\mathbf{x}^n}\in\mathcal{P}_{R,\nu}(\mathcal{X})} \Lambda''_{0,P_{\mathbf{x}^n}}(t); \tag{160}$$

$$K_{\max}(R,\nu) := 15\sqrt{2\pi} \max_{t\in H,\, P_{\mathbf{x}^n}\in\mathcal{P}_{R,\nu}(\mathcal{X})} \frac{T_{0,P_{\mathbf{x}^n}}(t)}{\Lambda''_{0,P_{\mathbf{x}^n}}(t)}; \tag{161}$$

$$T_{0,P_{\mathbf{x}^n}}(t) := \sum_{x\in\mathcal{X}} P_{\mathbf{x}^n}(x)\mathbb{E}_{\hat{q}_{x,t}}\left[\left|\log\frac{q_x}{p_x} - \Lambda'_{0,x}(t)\right|^3\right], \tag{162}$$

where

$$\mathcal{P}_{R,\nu}(\mathcal{X}) := \left\{P_{\mathbf{x}^n} \in \mathcal{P}(\mathcal{X}) : \nu \leq E_{\text{sp}}^{(2)}(R, P_{\mathbf{x}^n}) \leq E_{\text{sp}}(R) < +\infty\right\} \tag{163}$$

is a compact set owing to the continuity of $r \mapsto \phi_n(r)$. Also note that the maximization and minimization in the above definitions are well-defined and finite because $\Lambda''_{0,(\cdot)}(\cdot)$ and $T_{0,(\cdot)}(\cdot)$ are continuous functions in $(0,1] \times \mathcal{P}_R(\mathcal{X})$ [32, Lemma 6], where $\mathcal{P}_R(\mathcal{X})$ is defined in Eq. (44). Further, the quantity $V_{\min}(R,\nu)$ is bounded away from zero because of the positivity in Eq. (150).

Now, we are ready to derive the lower bounds for $\alpha\left(\mathcal{U}; p^n\right)$ and $\beta\left(\mathcal{U}; q^n\right)$. Let $N_2 \in \mathbb{N}$ be sufficiently large such that for all $n \geq N_2$,

$$\sqrt{n} \geq \frac{1 + (1 + K_{\max}(R,\nu))^2}{\sqrt{V_{\min}(R,\nu)}}. \tag{164}$$

Letting $Z_i = \log q_i - \log p_i$ with probability measure $\mu_i = p_i$, and $z = R_n - \phi_n(R_n)$ in Theorem 18, the Bahadur-Randga Rao's inequality gives

$$\begin{aligned}
\alpha\left(\mathcal{U}; p^n\right) &:= \sum_{\omega \in \mathcal{U}^c} p^n(\omega) && (165)\\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^n Z_i \geq R_n - \phi_n(R_n) \right\} && (166)\\
&\geq \frac{2A(R,\nu)}{\sqrt{n}} \exp\left\{ -n\Lambda^*_{0,P_{\mathbf{x}^n}}\left(\phi_n(R_n) - R_n\right)\right\} && (167)
\end{aligned}$$

where

$$A(R,\nu) := \frac{\mathrm{e}^{-K_{\max}(R,\nu)}}{\sqrt{4\pi V_{\max}(R,\nu)}}. \tag{168}$$

Similarly, applying Theorem 18 with $Z_i = \log p_i - \log q_i$, $\mu_i = q_i$, and $z = \phi_n(R_n) - R_n$ yields

$$\begin{aligned}
\beta\left(\mathcal{U}; q^n\right) &:= \sum_{\omega \in \mathcal{U}} q^n(\omega) && (169)\\
&= \Pr\left\{ \frac{1}{n}\sum_{i=1}^n Z_i \geq \phi_n(R_n) - R_n \right\} && (170)\\
&\geq \frac{2A(R,\nu)}{\sqrt{n}} \exp\left\{ -n\Lambda^*_{1,P_{\mathbf{x}^n}}\left(R_n - \phi_n(R_n)\right)\right\}. && (171)
\end{aligned}$$

Continuing from Eq. (167) and item (b) in Lemma 17 gives

$$\alpha\left(\mathcal{U}; p^n\right) \geq \frac{2A(R,\nu)}{\sqrt{n}} \exp\left\{-n\phi_n\left(R_n\right)\right\}. \tag{172}$$

Eq. (171) together with item (c) in Lemma 17 yields

$$\beta\left(\mathcal{U}; q^n\right) \geq \frac{2A(R,\nu)}{\sqrt{n}} \exp\left\{-nR_n\right\} = 2c\exp\left\{-nR\right\}, \tag{173}$$

where we choose $x = -\log A(R,\nu) + \log c$ in the rate back-off term $\gamma_n = \frac{\log n}{2n} + \frac{x}{n}$. Thus we can bound the left-hand side of Eq. (145) from below by $\frac{A(R,\nu)}{\sqrt{n}}\exp\{-n\phi_n(R_n)\}$. For any test $0 \leq Q_n \leq \mathbb{1}$ such that

$$\beta(Q_n; \sigma^n) \leq c\exp\left\{-nR\right\}, \tag{174}$$

we have,

$$\alpha(Q_n; \rho^n) \geq \frac{A(R,\nu)}{\sqrt{n}} \exp\left\{-n\phi_n\left(R_n\right)\right\}. \tag{175}$$

Hence, by choosing Q_n in Eqs.(174) and (175) that attains $\widehat{\alpha}_{c\exp\{-nR\}}$, we have

$$\widehat{\alpha}_{c\exp\{-nR\}}\left(\rho^n \| \sigma^n\right) \geq \frac{A(R,\nu)}{\sqrt{n}} \exp\left\{-n\phi_n\left(R_n\right)\right\} = \frac{A(R,\nu)}{\sqrt{n}} \exp\left\{-nE^{(2)}_{\text{sp}}\left(R - \gamma_n, P_{\mathbf{x}^n}\right)\right\}. \tag{176}$$

It remains to remove the rate back-off term γ_n in Eq. (176). By Taylor's theorem, one has

$$E^{(2)}_{\text{sp}}\left(R - \gamma_n, P_{\mathbf{x}^n}\right) = E^{(2)}_{\text{sp}}\left(R, P_{\mathbf{x}^n}\right) - \gamma_n \left.\frac{\partial E^{(2)}_{\text{sp}}\left(r, P_{\mathbf{x}^n}\right)}{\partial r}\right|_{r=R} + \frac{\gamma_n^2}{2}\left.\frac{\partial^2 E^{(2)}_{\text{sp}}\left(r, P_{\mathbf{x}^n}\right)}{\partial r^2}\right|_{r=\bar{R}}, \tag{177}$$

for some $\bar{R} \in (R_0, R)$. Recalling item (d) in Lemma 17, one can show that

$$\begin{aligned} - \left.\frac{\partial E_{\mathrm{sp}}^{(2)}\left(r, P_{\mathbf{x}^n}\right)}{\partial r}\right|_{r=R} &= s^\star_{R,P_{\mathbf{x}^n}} \in \mathbb{R}_{>0}, \\ \left.\frac{\partial^2 E_{\mathrm{sp}}^{(2)}\left(r, P_{\mathbf{x}^n}\right)}{\partial r^2}\right|_{r=\bar{R}} &= \frac{(1+\bar{s})^3}{\Lambda''_{0,P_{\mathbf{x}^n}}\left(\frac{\bar{s}}{1+\bar{s}}\right)} \leq \frac{(1+\bar{s})^3}{V_{\min}(R,\nu)} =: \Upsilon \in \mathbb{R}_{>0}, \end{aligned} \tag{178}$$

where $\bar{s} := -\left. \partial E_{\mathrm{sp}}^{(2)}(r, P_{\mathbf{x}^n})/\partial r \right|_{r=\bar{R}} \in \mathbb{R}_{>0}$, and the inequality follows from Eq. (160). Then, Eqs. (176), (177) and (178) lead to

$$\begin{aligned} \widehat{\alpha}_{c\exp\{-nR\}}\left(\rho^n \| \sigma^n\right) &\geq \frac{A(R,\nu)}{\sqrt{n}} \exp\left\{ -nE_{\mathrm{sp}}^{(2)}\left(R, P_{\mathbf{x}^n}\right) - n\left[\gamma_n \left(s^\star_{R,P_{\mathbf{x}^n}} + \frac{\gamma_n}{2}\Upsilon \right)\right]\right\} && (179) \\ &= \frac{A(R,\nu)}{n^{\frac{1}{2}\left(1+s^\star_{R,P_{\mathbf{x}^n}}\right)}} \exp\left\{ -nE_{\mathrm{sp}}^{(2)}\left(R, P_{\mathbf{x}^n}\right) - \ell_n \right\}, && (180) \end{aligned}$$

where we denote by

$$\ell_n := -\left(s^\star_{R,P_{\mathbf{x}^n}} + \frac{\gamma_n}{2}\Upsilon \right) \log A(R,\nu) + \frac{\gamma_n \Upsilon}{4} \log n. \tag{181}$$

Since $s^\star_{R,P_{\mathbf{x}^n}} \in \mathbb{R}_{>0}$ and $\gamma_n \log n = o(1)$, we choose a constant $L \in \mathbb{R}_{>0}$ and $N_3 \in \mathbb{N}$ such that

$$\ell_n \leq L, \quad \forall N \geq N_3. \tag{182}$$

Hence, Eqs. (180) and (182) lead to

$$\widehat{\alpha}_{c\exp\{-nR\}}\left(\rho^n \| \sigma^n\right) = \frac{A(R,\nu)\exp\{-L\}}{n^{\frac{1}{2}\left(1+s^\star_{R,P_{\mathbf{x}^n}}\right)}} \exp\left\{ -nE_{\mathrm{sp}}^{(2)}\left(R, P_{\mathbf{x}^n}\right) \right\}. \tag{183}$$

By letting $N_0 := \max\{N_1, N_2, N_3\}$ and $A' := A(R,\nu)\exp\{-L\}$, we conclude the proof. □

4.2. **Proofs of Theorem 9 and Corollary 10.** We are ready to prove our main result—the refined strong sphere-packing bound in Theorem 9 for constant composition codes and Corollary 10 for general codes.

Proof of Theorem 9. Fix any rate $R_\infty < R < C_{\mathcal{W}}$. First note that by Ref. [15, Proposition 10], we find

$$E_{\mathrm{sp}}(R) \in \mathbb{R}_{>0}. \tag{184}$$

By Lemma 11 and the standard expurgation method (see e.g. [7, p. 96], [10, Theorem 20], [57, p. 395]), it holds for every constant composition code $\mathcal{C}_n$ with a common composition $P_{\mathbf{x}^n}$ that

$$\begin{aligned} \bar{\epsilon}\left(\mathcal{C}_n\right) \geq \frac{1}{2}\epsilon_{\max}\left(\mathcal{C}'_n\right) &\geq \max_{\sigma \in \mathcal{S}(\mathcal{H})} \frac{1}{2} \widehat{\alpha}_{1/|\mathcal{C}'_n|}\left(W_{\mathbf{x}^n}^{\otimes n} \| \sigma^{\otimes n}\right) && (185) \\ &\geq \max_{\sigma \in \mathcal{S}(\mathcal{H})} \frac{1}{2} \widehat{\alpha}_{2\exp\{-nR\}}\left(W_{\mathbf{x}^n}^{\otimes n} \| \sigma^{\otimes n}\right) && (186) \\ &\geq \frac{1}{2} \widehat{\alpha}_{2\exp\{-nR\}}\left(W_{\mathbf{x}^n}^{\otimes n} \| (\sigma^\star)^{\otimes n}\right), && (187) \end{aligned}$$

where $\mathcal{C}'_n$ is an expurgated code with message size $|\mathcal{C}'_n| = \lceil |\mathcal{C}_n|/2 \rceil \geq \frac{1}{2}\exp\{nR\}$. Inequality (186) holds because the map $\mu \mapsto \widehat{\alpha}_\mu$ is monotone decreasing. In the last line (187) we denote by

$$\sigma^\star = \sigma^\star_{R,P_{\mathbf{x}^n}} := \operatorname*{arg\,min}_{\sigma \in \mathcal{S}(\mathcal{H})} \sup_{0<\alpha\leq 1} \left\{ \frac{1-\alpha}{\alpha}\left(D_\alpha\left(\mathcal{W} \| \sigma | P_{\mathbf{x}^n}\right) - R \right) \right\} \tag{188}$$

a channel output state that depends on the coding rate R and the composition $P_{\mathbf{x}^n}$.

In the following, we deal with sequences of inputs that will yield different lower bounds. Fix an arbitrary $\delta \in (0, E_{\mathrm{sp}}(R))$. Let $\nu := E_{\mathrm{sp}}(R) - \delta > 0$, and recall the definition in Eq. (163):

$$\mathcal{P}_{R,\nu}(\mathcal{X}) := \left\{ P_{\mathbf{x}^n} \in \mathcal{P}(\mathcal{X}) : \nu \leq E_{\mathrm{sp}}^{(2)}(R, P_{\mathbf{x}^n}) \leq E_{\mathrm{sp}}(R) < +\infty \right\}. \tag{189}$$

The set $\mathcal{P}_{R,\nu}(\mathcal{X})$ ensures that the error exponents of the input sequences $\mathbf{x}^n$ with composition $P_{\mathbf{x}^n} \in \mathcal{P}_{R,\nu}(\mathcal{X})$ are close to the sphere-packing exponent $E_{\text{sp}}(R)$.

For sequences $\mathbf{x}^n$ with $P_{\mathbf{x}^n} \notin \mathcal{P}_{R,\nu}(\mathcal{X})$, we infer that

$$E_{\text{sp}}(R) - E_{\text{sp}}^{(2)}\left(R, P_{\mathbf{x}^n}\right) = \delta > 0. \tag{190}$$

We then apply the Chebyshev-type bound, Proposition 13, with $c = 2$ to obtain, $\forall P_{\mathbf{x}^n} \notin \mathcal{P}_{R,\nu}(\mathcal{X})$,

$$\begin{aligned} \widehat{\alpha}_{2\exp\{-nR\}}\left(W_{\mathbf{x}^n}^{\otimes n} \| (\sigma^\star)^{\otimes n}\right) &\geq \kappa_1 \exp\left\{-\kappa_2\sqrt{n} - nE_{\text{sp}}^{(2)}\left(R, P_{\mathbf{x}^n}\right)\right\}, & (191)\\ &\geq \kappa_1 \exp\left\{-\kappa_2\sqrt{n} - n\left[E_{\text{sp}}\left(R\right) - \delta\right]\right\}, & (192)\end{aligned}$$

for all sufficiently large n, say $n \geq N_1 \in \mathbb{N}$. The equality in Eq. (191) follows from the saddle-point property, item (a) in Proposition 3, and the constants κ_1, κ_2 are positive and finite constants.

Next, we consider sequences $\mathbf{x}^n$ with $P_{\mathbf{x}^n} \in \mathcal{P}_{R,\nu}(\mathcal{X})$. Since such sequences satisfy Eq. (142), we apply the sharp lower bound, Proposition 14, with $c = 2$ to obtain, $\forall P_{\mathbf{x}^n} \in \mathcal{P}_{R,\nu}(\mathcal{X})$,

$$\widehat{\alpha}_{2\exp\{-nR\}}\left(W_{\mathbf{x}^n}^{\otimes n} \| (\sigma^\star)^{\otimes n}\right) \geq \frac{2A}{n^{\frac{1}{2}\left(1+s^\star_{R,P_{\mathbf{x}^n}}\right)}} \exp\left\{-nE_{\text{sp}}^{(2)}\left(R, P_{\mathbf{x}^n}\right)\right\}, \tag{193}$$

for all sufficiently large n, say $n \geq N_2 \in \mathbb{N}$, and some $A \in \mathbb{R}_{>0}$. In the following, we will relate the term $s^\star_{R,P_{\mathbf{x}^n}}$ in Eq. (193) to $\left|E'_{\text{sp}}(R)\right|$. The idea follows similar from [32, Eqs. (111)–(114)]. Let

$$\mathcal{P}^\star_R(\mathcal{X}) := \left\{P \in \mathcal{P}(\mathcal{X}) : E_{\text{sp}}^{(2)}(R, P) = E_{\text{sp}}(R)\right\}, \tag{194}$$

$$\mathcal{P}_\theta(\mathcal{X}) := \left\{P \in \mathcal{P}_{R,\nu}(\mathcal{X}) : \min_{Q \in \mathcal{P}^\star_R(\mathcal{X})} \|P - Q\|_1 \geq \theta\right\}. \tag{195}$$

Since $s^\star_{R,(\cdot)}$ is uniformly continuous on the compact set $P \in \mathcal{P}_{R,\nu}(\mathcal{X})$ (see item (d) of Proposition 4), one has

$$\forall \gamma \in \mathbb{R}_{>0},\ \exists f(\gamma) \in \mathbb{R}_{>0},\ \text{such that}\ \forall P, Q \in \mathcal{P}_{R,\nu}(\mathcal{X}),\ \|P - Q\|_1 < f(\gamma) \Rightarrow \left|s^\star_{R,P} - s^\star_{R,Q}\right| < \gamma. \tag{196}$$

By choosing $\gamma \in \mathbb{R}_{>0}$ that satisfies Eq. (196), it follows that

$$s^\star_{R,P_{\mathbf{x}^n}} \leq \left|E'_{\text{sp}}(R)\right| + \gamma, \quad \forall P_{\mathbf{x}^n} \in \mathcal{P}_{R,\nu}(\mathcal{X}) \backslash \mathcal{P}_{f(\gamma)}(\mathcal{X}). \tag{197}$$

Hence, Eqs. (193) and (197) further lead to, $\forall P_{\mathbf{x}^n} \in \mathcal{P}_{R,\nu}(\mathcal{X}) \backslash \mathcal{P}_{f(\gamma)}(\mathcal{X})$,

$$\widehat{\alpha}_{2\exp\{-nR\}}\left(W_{\mathbf{x}^n}^{\otimes n} \| (\sigma^\star)^{\otimes n}\right) \geq \frac{2A}{n^{\frac{1}{2}\left(1+|E'_{\text{sp}}(R)|+\gamma\right)}} \exp\left\{-nE_{\text{sp}}\left(R\right)\right\}. \tag{198}$$

For the case $P_{\mathbf{x}^n} \in \mathcal{P}_{R,\nu}(\mathcal{X}) \cap \mathcal{P}_{f(\gamma)}(\mathcal{X})$, we have

$$E_{\text{sp}}(R) - \max_{P \in \mathcal{P}_{f(\gamma)}(\mathcal{X})} E_{\text{sp}}^{(2)}(R, P_{\mathbf{x}^n}) =: \delta' > 0. \tag{199}$$

Then, Eqs. (193) and (199) give, $\forall P_{\mathbf{x}^n} \in \mathcal{P}_{R,\nu}(\mathcal{X}) \cap \mathcal{P}_{f(\gamma)}(\mathcal{X})$,

$$\widehat{\alpha}_{2\exp\{-nR\}}\left(W_{\mathbf{x}^n}^{\otimes n} \| (\sigma^\star)^{\otimes n}\right) \geq \frac{2A}{n^{\frac{1}{2}\left(1+s^\star_{R,P_{\mathbf{x}^n}}\right)}} \exp\left\{-n\left[E_{\text{sp}}\left(R\right) - \delta'\right]\right\}. \tag{200}$$

Finally, by comparing the bounds in Eqs. (192), (198) and (200), the first-order leading term in the right-hand side of Eq. (198) decays faster than that of Eqs. (192) and (200). Thus, for sufficiently large n, say $n \geq N_3 \in \mathbb{N}$, we combine the bounds to obtain, for all compositions $P_{\mathbf{x}^n} \in \mathcal{P}(\mathcal{X})$,

$$\widehat{\alpha}_{2\exp\{-nR\}}\left(W_{\mathbf{x}^n}^{\otimes n} \| (\sigma^\star)^{\otimes n}\right) \geq \frac{2A}{n^{\frac{1}{2}\left(1+|E'_{\text{sp}}(R)|+\gamma\right)}} \exp\left\{-nE_{\text{sp}}\left(R\right)\right\}. \tag{201}$$

By combining Eqs. (187), (201), we conclude our result: for any $\gamma > 0$ and every n-blocklength constant composition code $\mathcal{C}_n$,

$$\bar{\epsilon}\left(\mathcal{C}_n\right) \geq \frac{A}{n^{\frac{1}{2}\left(1+|E'_{\text{sp}}(R)|+\gamma\right)}} \exp\left\{-nE_{\text{sp}}\left(R\right)\right\}, \tag{202}$$

for all sufficiently large $n \geq N_0 := \max\left\{N_1, N_2, N_3\right\}$. □

Proof of Corollary 10. For an n-blocklength code, there are at most $\binom{n+|\mathcal{X}|-1}{|\mathcal{X}|-1} < n^{|\mathcal{X}|}$ different compositions. Hence, for any code with $M = \exp\{nR\}$ codewords, there exists some codewords M' of the same composition such that $M' \geq M/n^{|\mathcal{X}|}$. Denote by $\mathcal{C}_n'$ such constant composition codes with composition $P_{\mathbf{x}^n}$.

Fix an arbitrary $R_0 \in (R_\infty, R)$, and choose N_1 be an integer such that $R - \frac{|\mathcal{X}|}{n}\log n \geq R_0$ for all $n \geq N_1$. Consider such $n \geq N_1$ onwards. By following the similar steps in Theorem 9, we obtain

$$\epsilon^*\left(n,R\right) \geq \bar{\epsilon}\left(\mathcal{C}_n'\right) \geq \frac{A}{n^{\frac{1}{2}\left(1+s^\star_{R,P^n_{\mathbf{x}}}\right)}} \exp\left\{-nE_{\text{sp}}^{(2)}\left(R - \frac{|\mathcal{X}|}{n}\log n, P_{\mathbf{x}^n}\right)\right\}, \tag{203}$$

for all sufficiently large n, say $n \geq N_2 \in \mathbb{N}$, and some $s^\star_{R,P_{\mathbf{x}^n}} \in \mathbb{R}_{>0}$. Let

$$\Upsilon := \max_{P \in \mathcal{P}(\mathcal{X}) : E_{\text{sp}}^{(2)}(\bar{R},P) = E_{\text{sp}}(\bar{R})} \left| \left.\frac{\partial E_{\text{sp}}^{(2)}(r,P)}{\partial r}\right|_{r=R_0} \right|. \tag{204}$$

Then, item (a) in Proposition 4 implies that

$$\begin{aligned} E_{\text{sp}}^{(2)}\left(R - \frac{|\mathcal{X}|}{n}\log n, P_{\mathbf{x}^n}\right) &\leq E_{\text{sp}}^{(2)}(R, P_{\mathbf{x}^n}) + \Upsilon \cdot \frac{|\mathcal{X}|}{n}\log n \qquad (205)\\ &\leq E_{\text{sp}}(R) + \Upsilon \cdot \frac{|\mathcal{X}|}{n}\log n, \quad \forall n \geq N_2 \qquad (206)\end{aligned}$$

Combining Eqs. (203) and (206) gives

$$\epsilon^*\left(n,R\right) \geq \frac{A}{n^{\frac{1}{2}\left(1+s^\star_{R,P_{\mathbf{x}^n}}\right)+\Upsilon|\mathcal{X}|}} \exp\left\{-nE_{\text{sp}}(R)\right\}, \quad \forall n \geq \max\{N_1, N_2\}. \tag{207}$$

By choosing $t \in \mathbb{R}_{>0}$ such that $n^{-t} \leq An^{-\frac{1}{2}\left(1+s^\star_{R,P_{\mathbf{x}^n}}\right)-\Upsilon|\mathcal{X}|}$, and letting $N_0 := \max\{N_1, N_2\}$, we conclude our claim. □

5. Symmetric Classical-Quantum Channels

In this section, we consider a symmetric c-q channels. By using the symmetric property of the channels, we show that the uniform distribution, denoted by $U_\mathcal{X}$, achieves the maximum of $E_{\text{sp}}^{(1)}(R,\cdot)$ and $E_{\text{sp}}^{(2)}(R,\cdot)$. Then, by choosing the optimal output state $\sigma_R^\star = \sigma_{R,U_\mathcal{X}}^\star$, every input sequence in the codebook is a good codeword and attains the sphere-packing exponent $E_{\text{sp}}(R)$. Hence, we can remove the assumption of constant composition codes and apply Theorem 9 in Section 4 to obtain the optimal pre-factor for the sphere-packing bound (Theorem 15).

A c-q channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ is *symmetric* if it satisfies

$$W_x := V^{x-1} W_1 (V^\dagger)^{x-1}, \quad \forall x \in \mathcal{X}, \tag{208}$$

where $W_1 \in \mathcal{S}(\mathcal{H})$ is an arbitrary density operator, and V satisfies $V^\dagger V = VV^\dagger = V^{|\mathcal{X}|} = \mathbb{1}_\mathcal{H}$.

Theorem 15 (Exact Sphere-packing Bound for Symmetric Classical-Quantum Channels)**.** *For any rate $R \in (R_\infty, C_\mathcal{W})$, there exist $A > 0$ and $N_0 \in \mathbb{N}$ such that for all codes $\mathcal{C}_n$ of length $n \geq N_0$ with message size $|\mathcal{C}_n| \geq \exp\{nR\}$, we have*

$$\epsilon_{\max}\left(\mathcal{C}_n\right) \geq \frac{A}{n^{\frac{1}{2}\left(1+|E_{\text{sp}}'(R)|\right)}} \exp\left\{-nE_{\text{sp}}(R)\right\}. \tag{209}$$

Proof. The proof consists of the following steps. First, we show that the distribution $U_\mathcal{X}$ satisfies $E_{\text{sp}}^{(1)}(R,U_\mathcal{X}) = E_{\text{sp}}^{(2)}(R,U_\mathcal{X}) = E_{\text{sp}}(R)$. Second, we show that $E_{\text{sp}}^{(2)}(R,P) = E_{\text{sp}}(R)$ for all $P \in \mathcal{P}(\mathcal{X})$, which means that any codeword attains the sphere-packing exponent. Finally, we follow Theorem 9 to complete the proof.

Fix any $R \in (R_\infty, C_{\mathcal{W}})$. From the definition of the symmetric channels in Eq. (208), it is not hard to verify that $U_{\mathcal{X}}\mathcal{W}^\alpha = V U_{\mathcal{X}}\mathcal{W}^\alpha V^\dagger$ for all $\alpha \in (0,1]$, where we denote by $P\mathcal{W}^\alpha := \sum_{x\in\mathcal{X}} P(x) W_x^\alpha$ for all $\alpha \in (0,1]$. Hence, it follows that

$$\operatorname{Tr}[W_x^\alpha (U_{\mathcal{X}}\mathcal{W}^\alpha)^{\frac{1-\alpha}{\alpha}}] = \operatorname{Tr}[V^{x-1} W_1^\alpha V^{\dagger\, x-1} (U_{\mathcal{X}}\mathcal{W}^\alpha)^{\frac{1-\alpha}{\alpha}}] \tag{210}$$

$$= \operatorname{Tr}[W_1^\alpha (U_{\mathcal{X}}\mathcal{W}^\alpha)^{\frac{1-\alpha}{\alpha}}] \tag{211}$$

for all $x \in \mathcal{X}$ and $\alpha \in (0,1]$. Summing Eq. (211) over all $x \in \mathcal{X}$ and dividing by M yields that

$$\operatorname{Tr}[W_x^\alpha (U_{\mathcal{X}}\mathcal{W}^\alpha)^{\frac{1-\alpha}{\alpha}}] = \operatorname{Tr}[(U_{\mathcal{X}}\mathcal{W}^\alpha)^{\frac{1}{\alpha}}], \tag{212}$$

for all $x \in \mathcal{X}$ and $\alpha \in (0,1]$. Recalling Proposition 16 below, the above equation shows that the distribution $U_{\mathcal{X}}$ indeed maximizes $E_0(s,P)$, $\forall s \in \mathbb{R}_{\geq 0}$. Then we have

$$E_{\text{sp}}^{(1)}(R, U_{\mathcal{X}}) = \sup_{s\geq 0} \left\{ \max_{P\in\mathcal{P}(\mathcal{X})} E_0(s,P) - sR \right\} = E_{\text{sp}}(R).$$

Further, Jensen's inequality shows that $E_{\text{sp}}^{(2)}(R, U_{\mathcal{X}}) \geq E_{\text{sp}}^{(1)}(R, U_{\mathcal{X}}) = E_{\text{sp}}(R)$, and thus, $E_{\text{sp}}^{(2)}(R, U_{\mathcal{X}}) = E_{\text{sp}}(R)$.

Next, let $(\alpha_R^\star, \sigma_R^\star)$ be the saddle-point of $F_{R,U_{\mathcal{X}}}(\cdot,\cdot)$ (see Eq. (43)). One can observe from the definition of $E_{\text{sp}}^{(2)}$ and Eq. (212) that all the quantities $D_{\alpha_R^\star}(W_x \| \sigma_R^\star)$, $x \in \mathcal{X}$, are equal. By item (c) of Proposition 3 and Eq. (300), we obtain

$$\sigma_R^\star = \frac{\left(U_{\mathcal{X}}\mathcal{W}^{\alpha_R^\star}\right)^{1/\alpha_R^\star}}{\operatorname{Tr}\left[\left(U_{\mathcal{X}}\mathcal{W}^{\alpha_R^\star}\right)^{1/\alpha_R^\star}\right]}, \tag{213}$$

which, in turn, implies that

$$E_{\text{sp}}^{(2)}(R,P) = \sup_{\alpha\in(0,1]} F_{R,P}(\alpha, \sigma_R^\star) = \sup_{s\geq 0} \left\{ E_0(s, U_{\mathcal{X}}) - sR \right\} = E_{\text{sp}}(R), \quad \forall P \in \mathcal{P}(\mathcal{X}). \tag{214}$$

Further, we have

$$\left| E_{\text{sp}}'(R) \right| = \frac{1-\alpha_R^\star}{\alpha_R^\star} = \left| \frac{\partial E_{\text{sp}}^{(2)}(R,P)}{\partial R} \right|, \quad \forall P \in \mathcal{P}(\mathcal{X}). \tag{215}$$

Since Eqs. (214) and (5) indicates that every input sequence attains the sphere-packing exponent, we apply the same arguments in the proof of Theorem 9 to conclude this theorem.

Proposition 16 ([14, Eq. (38)]). *Let $s \in \mathbb{R}_{\geq 0}$ be arbitrary. The Necessary and sufficient condition for the distribution $P^\star$ to maximize $E_0(s,P)$ is*

$$\operatorname{Tr}\left[W_x^{1/(1+s)} \cdot \left(\sum_{x\in\mathcal{X}} P^\star(x) W_x^{1/(1+s)} \right)^s \right] \geq \operatorname{Tr}\left[\left(\sum_{x\in\mathcal{X}} P^\star(x) W_x^{1/(1+s)} \right)^{1+s} \right], \forall x \in \mathcal{X} \tag{216}$$

with equality if $P^\star(x) \neq 0$.

□

6. Conclusions

In this paper, we provided an exposition of sphere-packing bounds in classical and quantum channel coding. Unlike classical results, there are two different quantum sphere-packing exponents, one being stronger than the other. We provided variational representations for these two exponents, and showed that they are ordered by the Golden-Thompson inequality. Our proof strategy was inspired by Blahut's approach of hypothesis testing reduction [10] and Altuğ-Wagner's technique in strong large deviation theory [32]. Specifically, the pre-factor of the bound, that is akin to the converse Hoeffding bound in quantum hypothesis testing, can be improved by Bahadur-Ranga Rao's sharp concentration inequality [26, 27]. Consequently, we obtained a refined strong sphere-packing bound for c-q channels and constant composition codes with a polynomial pre-factor $f(n) = n^{-\frac{1}{2}\left(1+|E'_{\mathrm{sp}}(R)|+o(1)\right)}$. Moreover, the established result matches the best known random coding bound (i.e. achievability) up to the logarithmic order [32, 23, 24, 25]. For the case of general codes, the derived pre-factor is of the polynomial order, i.e. $f(n) = O(n^{-t})$ for some $t > 1/2$. We are able to obtain the exact pre-factor without the assumption of constant composition codes for a class of symmetric c-q channels. We note that the exact pre-factor for general codes is still open even in the classical case. Finally, our refinement enables a moderate deviation analysis in c-q channels [29] (see also [30]).

Appendix A. Lengendre-Fenchel Transform and Error-Exponent Functions

In this section, we will see that the Lengendre-Fenchel transform is closely related to the error-exponent function of hypothesis testing and channel coding. Consider the following binary hypotheses:

$$\begin{aligned} \mathsf{H}_0 &: p^n := p_{x_1} \otimes p_{x_2} \otimes \cdots p_{x_n}, \\ \mathsf{H}_1 &: q^n := q_{x_1} \otimes q_{x_2} \otimes \cdots q_{x_n}, \end{aligned} \tag{217}$$

where p_{x_i}, q_{x_i} are probability mass functions; and x_i belongs to some finite alphabet $\mathcal{X}$ and $n \in \mathbb{N}$ be fixed. Given any $r \geq 0$, recall the definition of the error-exponent function in Eq. (56):

$$\phi_n(r) = \phi_n(r|p^n\|q^n) = \sup_{\alpha\in(0,1]} \left\{ \frac{1-\alpha}{\alpha} \left(\frac{1}{n} D_\alpha\left(p^n\|q^n\right) - r \right) \right\}. \tag{218}$$

Without loss of generality, we assume that $p^n \ll q^n$ have the same support since elements of q_{x_i}, that do not lie in the support of p_{x_i}, do not contribute to $\phi_n(r)$.

Let Z be a random variable with probability measure μ. Further, we assume Z is finite on $\mathtt{supp}(\mu)$. The cumulant generating function (c.g.f.) of Z is defined as

$$\Lambda(t) := \log \mathbb{E}_\mu\left[\mathrm{e}^{tZ}\right], \quad t \in \mathbb{R}. \tag{219}$$

The *Lengendre-Fenchel transform* of $\Lambda(t)$ is

$$\Lambda^*(z) := \sup_{t\in\mathbb{R}} \left\{ zt - \Lambda(t) \right\}. \tag{220}$$

Such a transform plays a significant role in concentration inequalities, convex analysis, and large deviation theory [27].

Let $P_{\mathbf{x}^n}$ be the empirical distribution of the sequence $\mathbf{x}^n = x_1x_2\ldots x_n$. Let $Z_0 = \log\frac{q^n}{p^n}$ with probability measure p^n, $Z_1 = \log\frac{p^n}{q^n}$ with probability measure q^n, and denote

$$\begin{aligned} \Lambda_{0,P_{\mathbf{x}^n}}(t) &:= \frac{1}{n} \log \mathbb{E}_{p^n}\left[\mathrm{e}^{tZ_0}\right] = \sum_{x\in\mathcal{X}} P_{\mathbf{x}^n}(x)\Lambda_{0,x_i}(t), \\ \Lambda_{1,P_{\mathbf{x}^n}}(t) &:= \frac{1}{n} \log \mathbb{E}_{q^n}\left[\mathrm{e}^{tZ_1}\right] = \sum_{x\in\mathcal{X}} P_{\mathbf{x}^n}(x)\Lambda_{1,x_i}(t); \end{aligned} \tag{221}$$

where

$$\Lambda_{0,x_i}(t) := \log \mathbb{E}_{p_{x_i}}\left[\mathrm{e}^{t\log\frac{q_{x_i}}{p_{x_i}}}\right], \quad \Lambda_{1,x_i}(t) := \log \mathbb{E}_{q_{x_i}}\left[\mathrm{e}^{t\log\frac{p_{x_i}}{q_{x_i}}}\right]. \tag{222}$$

Rewrite the right-hand side of Eq. (218) with $\alpha = \frac{1}{1+s}$, and observe that

$$\sum_{x\in\mathcal{X}} P_{\mathbf{x}^n}(x) s D_{\frac{1}{1+s}}(p_x\|q_x) = -(1+s)\Lambda_{0,P_{\mathbf{x}^n}}\left(\frac{s}{1+s}\right) \tag{223}$$

$$=: E_0^{(2)}(s, P_{\mathbf{x}^n}). \tag{224}$$

Then the error-exponent function in Eq. (218) can also be viewed as a Lengendre-Fenchel transform of $E_0^{(2)}(s, P_{\mathbf{x}^n})$:

$$\phi_n(r) = \sup_{s\geq 0} \left\{ E_0^{(2)}(s, P_{\mathbf{x}^n}) - sr \right\}. \tag{225}$$

The following lemma relates $\phi_n(r)$ to $\Lambda^*_{j,P_{\mathbf{x}^n}}(z)$, the Lengendre-Fenchel transform of Eq. (221):

$$\Lambda^*_{j,P_{\mathbf{x}^n}}(z) := \sup_{t\in\mathbb{R}} \left\{ tz - \Lambda_{j,P_{\mathbf{x}^n}}(t) \right\}, \quad j \in \{0,1\}. \tag{226}$$

Lemma 17. *Let p^n and q^n, $n \in \mathbb{N}$, be described as above. Assume $r > \frac{1}{n} D_0\left(p^n \| q^n\right)$ and $\phi_n(r) > 0$. The following hold:*

(a) $\Lambda''_{0,P_{\mathbf{x}^n}}(t) > 0$ for all $t \in [0,1]$.
(b) $\Lambda^*_{0,P_{\mathbf{x}^n}}\left(\phi_n(r) - r\right) = \phi_n(r)$.
(c) $\Lambda^*_{1,P_{\mathbf{x}^n}}\left(r - \phi_n(r)\right) = r$.
(d) Let $t^\star := t^\star_{r,P_{\mathbf{x}^n}}$ be the optimizer of $\Lambda^*_{0,P_{\mathbf{x}^n}}(z)$ in Eq. (226), and $s^\star := s^\star_{r,P_{\mathbf{x}^n}}$ be the optimizer of $\phi_n(r)$ in Eq. (225). The optimizer $t^\star \in (0,1)$ is unique, and satisfies $\Lambda'_{0,P_{\mathbf{x}^n}}(t^\star) = \phi_n(r) - r$. In particular, one has $t^\star = \frac{s^\star}{1+s^\star}$; $s^\star = -\frac{\partial \phi_n(r)}{\partial r}$; and $\frac{\partial^2 \phi_n(r)}{\partial r^2} = -\left(\left. \frac{\partial^2 E_0^{(2)}(s, P_{\mathbf{x}^n})}{\partial s^2} \right|_{s=s^\star} \right)^{-1} = \frac{\left(1+s^\star_{r,P_{\mathbf{x}^n}}\right)^3}{\Lambda''_{0,P_{\mathbf{x}^n}}(t^\star)} > 0$.

Before proving Lemma 17, we will need the following partial derivatives with respect to t:

$$\Lambda'_{0,x_i}(t) = \mathbb{E}_{\hat{q}_{x_i,t}}\left[\log \frac{q_{x_i}}{p_{x_i}}\right], \quad \Lambda'_{1,x_i}(t) = \mathbb{E}_{\hat{q}_{x_i,1-t}}\left[\log \frac{p_{x_i}}{q_{x_i}}\right]; \tag{227}$$

$$\Lambda''_{0,x_i}(t) = \mathrm{Var}_{\hat{q}_{x_i,t}}\left[\log \frac{q_{x_i}}{p_{x_i}}\right], \quad \Lambda''_{1,x_i}(t) = \mathrm{Var}_{\hat{q}_{x_i,1-t}}\left[\log \frac{p_{x_i}}{q_{x_i}}\right], \tag{228}$$

where we denote the *tilted distributions* for every $i \in [n]$ and $t \in [0,1]$ by

$$\hat{q}_{x_i,t}(\omega) := \frac{p_{x_i}(\omega)^{1-t} q_{x_i}(\omega)^t}{\sum_{\omega\in\mathsf{supp}(p_{x_i})} p_{x_i}(\omega)^{1-t} q_{x_i}(\omega)^t}, \quad \omega \in \mathsf{supp}(p_{x_i}). \tag{229}$$

It is also easy to verify that

$$\Lambda_{0,x_i}(t) = \Lambda_{1,x_i}(1-t), \quad \Lambda'_{0,x_i}(t) = -\Lambda'_{1,x_i}(1-t), \quad \Lambda''_{0,x_i}(t) = \Lambda''_{1,x_i}(1-t). \tag{230}$$

This lemma closely follows Ref. [32, Lemma 9]; however, the major difference is that we prove the claim using $\phi_n(r|\rho^n\|\sigma^n)$ in Eq. (56) instead of the discrimination function: $\min\left\{D\left(\tau\|\rho\right) : D\left(\tau\|\sigma\right) \le r\right\}$ in Eq. (65). This expression is crucial to obtaining the sphere-packing bound in Theorem 9 in the strong from, cf. Eq. (2), instead of the weak form, cf. Eq. (3).

Proof of Lemma 17-(a). We will prove this statement by contradiction. Let $t \in [0,1]$, Assuming that $\Lambda''_{0,P_{\mathbf{x}^n}}(t) = 0$, implies $\Lambda''_{0,x}(t) = 0$, $\forall x \in \mathsf{supp}(P_{\mathbf{x}^n})$. Recall from Eq. (228)

$$0 = \Lambda''_{0,x}(t) = \mathrm{Var}_{\hat{q}_{x,t}}\left[\log \frac{q_x}{p_x}\right], \tag{231}$$

which is equivalent to

$$p_x(\omega) = q_x(\omega) \cdot \mathrm{e}^{-\Lambda'_{0,x}(t)}, \quad \forall \omega \in \mathsf{supp}(p_x). \tag{232}$$

Summing both sides of Eq. (232) over $\omega \in \mathsf{supp}(p_x)$ gives

$$1 = \mathrm{Tr}\left[p_x^0 q_x\right] \mathrm{e}^{-\Lambda'_{0,x}(t)}. \tag{233}$$

Then, Eqs. (232) and (233) imply that

$$\begin{aligned}
\phi_n(r) &= \sup_{0<\alpha\le 1} \frac{\alpha - 1}{\alpha}\left(r - \sum_{x\in\mathcal{X}} P_{\mathbf{x}^n}(x) D_\alpha\left(p_x \| q_x\right)\right) &(234)\\
&= \sup_{0<\alpha\le 1} \frac{\alpha - 1}{\alpha}\left(r + \sum_{x\in\mathcal{X}} P_{\mathbf{x}^n}(x) \log \mathrm{Tr}\left[p_x^0 q_x\right]\right) &(235)\\
&= 0, &(236)
\end{aligned}$$

where Eq. (236) follows since $r > \frac{1}{n} D_0(p^n\|q^n) = -\frac{1}{n}\sum_{x\in\mathcal{X}} P_{\mathbf{x}^n}(x) \log \mathrm{Tr}\left[p_x^0 q_x\right]$ by assumption. However, this contradicts with the assumption $\phi_n(r) > 0$. Hence, we conclude item (a). □

Proof of Lemma 17-(b). Observe that $E_0^{(2)}(s, P_{\mathbf{x}^n}) - sr$ in Eq. (225) is strictly concave in $s \in \mathbb{R}_{\geq 0}$ since

$$\frac{\partial^2 E_0^{(2)}(s, P_{\mathbf{x}^n})}{\partial s^2} = -\frac{1}{(1+s)^3} \Lambda''_{0,P_{\mathbf{x}^n}} \left(\frac{s}{1+s} \right) < 0, \tag{237}$$

owing to Eqs. (224), (228), and Lemma (a). Moreover, $s = 0$ cannot be an optimum in Eq. (225); otherwise, it will violate the assumption $\phi_n(r) \geq 0$. Thus a unique maximizer $s^\star \in \mathbb{R}_{>0}$ exists such that

$$\phi_n(r) = -s^\star r + E_0^{(2)}(s^\star, P_{\mathbf{x}^n}) \tag{238}$$

$$= \frac{s^\star}{1+s^\star} \Lambda'_{0,P_{\mathbf{x}^n}} \left(\frac{s^\star}{1+s^\star} \right) - \Lambda_{0,P_{\mathbf{x}^n}} \left(\frac{s^\star}{1+s^\star} \right). \tag{239}$$

where in the second equality we use Eq. (224) and

$$r = \left. \frac{\partial E_0^{(2)}(s, P_{\mathbf{x}^n})}{\partial s} \right|_{s=s^\star} \tag{240}$$

$$= -\frac{1}{1+s^\star} \Lambda'_{0,P_{\mathbf{x}^n}} \left(\frac{s^\star}{1+s^\star} \right) - \Lambda_{0,P_{\mathbf{x}^n}} \left(\frac{s^\star}{1+s^\star} \right). \tag{241}$$

Comparing Eq. (239) with (241) gives

$$\Lambda'_{0,P_{\mathbf{x}^n}} \left(\frac{s^\star}{1+s^\star} \right) = \phi_n(r) - r, \tag{242}$$

which is exactly the optimum solution to $\Lambda^*_{0,P_{\mathbf{x}^n}}(z)$ in Eq. (226) with

$$t^\star = \frac{s^\star}{1+s^\star} \in (0,1), \tag{243}$$

$$z = \phi_n(r) - r. \tag{244}$$

Hence, we obtain

$$\Lambda^*_{0,P_{\mathbf{x}^n}} (\phi_n(r) - r) = t^\star z - \Lambda_{0,P_{\mathbf{x}^n}}(t^\star) \tag{245}$$

$$= \frac{s^\star}{1+s^\star} (\phi_n(r) - r) - \Lambda_{0,P_{\mathbf{x}^n}} \left(\frac{s^\star}{1+s^\star} \right) \tag{246}$$

$$= \frac{s^\star}{1+s^\star} \Lambda'_{0,P_{\mathbf{x}^n}} \left(\frac{s^\star}{1+s^\star} \right) - \Lambda_{0,P_{\mathbf{x}^n}} \left(\frac{s^\star}{1+s^\star} \right) \tag{247}$$

$$= \phi_n(r), \tag{248}$$

where Eqs. (242) and (239) are used in the third and last equalities. □

Proof of Lemma 17-(c). This proof follows from similar arguments in item (b) and Eq. (230). Eqs. (242) and (230) lead to

$$\Lambda'_{1,P_{\mathbf{x}^n}} \left(\frac{1}{1+s^\star} \right) = r - \phi_n(r), \tag{249}$$

which satisfies the optimum solution to $\Lambda_{1,P_{\mathbf{x}^n}}(z)$ in Eq. (226) with $t^\star = \frac{1}{1+s^\star} \in (0,1)$ and $z = r - \phi_n(r)$. Then,

$$\Lambda^*_{1,P_{\mathbf{x}^n}} (r - \phi_n(r)) = t^\star z - \Lambda_{1,P_{\mathbf{x}^n}}(t^\star) \tag{250}$$

$$= \frac{1}{1+s^\star} (r - \phi_n(r)) - \Lambda_{1,P_{\mathbf{x}^n}} \left(\frac{s^\star}{1+s^\star} \right) \tag{251}$$

$$= \frac{1}{1+s^\star} \Lambda'_{1,P_{\mathbf{x}^n}} \left(\frac{1}{1+s^\star} \right) - \Lambda_{1,P_{\mathbf{x}^n}} \left(\frac{1}{1+s^\star} \right) \tag{252}$$

$$= r, \tag{253}$$

where the third equality is due to Eq. (249), and the last equality follows from Eqs. (230) and (241).

□

Proof of Lemma 17-(d). The fact that a unique optimizer $t^\star \in (0,1)$ exists such that $\Lambda'_{0,P_{\mathbf{x}^n}}(t^\star) = \phi_n(r) - r$ follows directly from Eqs. (242), (243) and $\Lambda''_{0,P_{\mathbf{x}^n}}(t) > 0$, for $t \in [0,1]$.

Moreover, Eqs. (238), (240), and (237) yield

$$-\frac{\partial \phi_n(r)}{\partial r} = s^\star, \tag{254}$$

$$\frac{\partial^2 \phi_n(r)}{\partial r^2} = -\frac{\partial s^\star}{\partial r} = -\left(\frac{\partial^2 E_0^{(2)}(s, P_{\mathbf{x}^n})}{\partial s^2} \right)^{-1} \Bigg|_{s=s^\star} = \frac{(1+s^\star)^3}{\Lambda_{0,P_{\mathbf{x}^n}}\left(\frac{s^\star}{1+s^\star}\right)}, \tag{255}$$

which completes the claim in item (d). □

Appendix B. A Tight Large Deviation Inequality

Let $(Z_i)_{i=1}^n$ be a sequence of independent, real-valued random variables with probability measures $(\mu_i)_{i=1}^n$. Let $\Lambda_i(t) := \log \mathbb{E}\left[\mathrm{e}^{tZ_i}\right]$ and define the Legendre-Fenchel transform of $\frac{1}{n}\sum_{i=1}^n \Lambda_i(\cdot)$ to be:

$$\Lambda_n^*(z) := \sup_{t\in\mathbb{R}} \left\{ zt - \frac{1}{n}\sum_{i=1}^n \Lambda_i(t) \right\}, \quad \forall z \in \mathbb{R}. \tag{256}$$

Then there exists a real number $t^\star \in (0,1]$ for every $z \in \mathbb{R}$ such that

$$z = \frac{1}{n}\sum_{i=1}^n \Lambda_i'(t^\star); \tag{257}$$

$$\Lambda_n^*(z) = zt^\star - \frac{1}{n}\sum_{i=1}^n \Lambda_i(t^\star). \tag{258}$$

Define the probability measure $\tilde{\mu}_i$ via

$$\frac{\mathrm{d}\tilde{\mu}_i}{\mathrm{d}\mu_i}(z_i) := \mathrm{e}^{t^\star z_i - \Lambda_i(t^\star)}, \tag{259}$$

and let $\bar{Z}_i := Z_i - \mathbb{E}_{\tilde{\mu}_i}[Z_i]$. Furthermore, define $m_{2,n} := \sum_{i=1}^n \mathrm{Var}_{\tilde{\mu}_i}\left[\bar{Z}_i\right]$, $m_{3,n} := \sum_{i=1}^n \mathbb{E}_{\tilde{\mu}_i}\left[\left|\bar{Z}_i\right|^3\right]$, and $K_n(t^\star) := \frac{15\sqrt{2\pi}m_{3,n}}{m_{2,n}}$. With these definitions, we can now state the following sharp concentration inequality for $\frac{1}{n}\sum_{i=1}^n Z_i$:

Theorem 18 (Bahadur-Ranga Rao's Concentration Inequality [32, Proposition 5], [26])**.** *Provided that $\sqrt{m_{2,n}} \geq 1 + (1 + K_n(t^\star))^2$, then*

$$\Pr\left\{ \frac{1}{n}\sum_{i=1}^n Z_i \geq z \right\} \geq \mathrm{e}^{-n\Lambda_n^*(z)} \frac{\mathrm{e}^{-K_n(t^\star)}}{2\sqrt{2\pi m_{2,n}}}. \tag{260}$$

Appendix C. Proof of Proposition 2

Proposition 2 (Properties of α-Rényi Mutual Information and Radius)**.** *Given any classical-quantum channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$, the following holds:*

(a) The map $(\alpha, P) \mapsto I_\alpha^{(2)}(P, \mathcal{W})$ is continuous on $[0,1] \times \mathcal{P}(\mathcal{X})$.

(b) For every $P \in \mathcal{P}(\mathcal{X})$, $\alpha \mapsto I_\alpha^{(2)}(P, \mathcal{W})$ is monotone increasing on $[0,1]$.

(c) For every $P \in \mathcal{P}(\mathcal{X})$, $\alpha \mapsto \frac{1-\alpha}{\alpha} I_\alpha^{(2)}(P, \mathcal{W})$ is strictly concave on $(0,1]$.

(d) The map $\alpha \mapsto C_{\alpha,\mathcal{W}}$ is continuous and monotone increasing on $[0,1]$.

Items (a), (b), and (c) also hold for $I_\alpha^{(1)}(P, \mathcal{W})$.

Proof of Proposition 2-(a). Fix an arbitrary sequence $(\alpha_k, P_k)_{k\in\mathbb{N}}$ such that $\alpha_k \in [0,1]$, $P_k \in \mathcal{P}(\mathcal{X})$, and $\lim_{k\to+\infty}(\alpha_k, P_k) = (\alpha_\infty, P_\infty) \in [0,1] \times \mathcal{P}(\mathcal{X})$. Let

$$\sigma_k^\star \in \arg\min_{\sigma\in\mathcal{S}(\mathcal{H})} D_{\alpha_k}\left(\mathcal{W}\|\sigma|P_k\right), \quad \forall k \in \mathbb{N}\cup\{+\infty\}. \tag{261}$$

The definition in Eq. (29) implies that

$$\begin{aligned}
\liminf_{k\to+\infty} I_{\alpha_k}^{(2)}(P_k, \mathcal{W}) &= \liminf_{k\to+\infty} D_{\alpha_k}\left(\mathcal{W}\|\sigma_k^\star|P_k\right) && (262)\\
&\geq D_{\alpha_\infty}\left(\mathcal{W}\middle\|\lim_{k\to+\infty}\sigma_k^\star\middle| P_\infty\right) && (263)\\
&\geq \min_{\sigma\in\mathcal{S}(\mathcal{H})} D_{\alpha_\infty}\left(\mathcal{W}\|\sigma|P_\infty\right) && (264)\\
&= I_{\alpha_\infty}^{(2)}(P_\infty, \mathcal{W}), && (265)
\end{aligned}$$

where, in order to establish (263), we used the lower semi-continuity of the map $\sigma \mapsto D_{\alpha_k}(\mathcal{W}\|\sigma|P_k)$ in Eq. (23) and the continuity of $(\alpha, P) \mapsto D_\alpha\left(\mathcal{W}\|\sigma_k^\star|P\right)$ (Eq. (18) in Lemma 1).

Next, we let

$$\begin{aligned}
\sigma_k &:= (1-\epsilon_k)\,\sigma_\infty^\star + \epsilon_k\frac{\mathbb{1}}{d}, \quad \forall k \in \mathbb{N}; && (266)\\
\epsilon_k &:= \frac{\|P_k - P_\infty\|_1}{2}. && (267)
\end{aligned}$$

Then, it follows that

$$\begin{aligned}
\limsup_{k\to+\infty} I_{\alpha_k}^{(2)}(P_k, \mathcal{W}) &\leq \limsup_{k\to+\infty}\left\{D_{\alpha_k}\left(\mathcal{W}\|\sigma_k|P_k\right)\right\} && (268)\\
&= \limsup_{k\to+\infty}\left\{D_{\alpha_k}\left(\mathcal{W}\|\sigma_k|P_\infty\right) + \sum_{x\in\mathcal{X}}\left[P_k(x) - P_\infty(x)\right] D_{\alpha_k}\left(W_x\|\sigma_k\right)\right\} && (269)\\
&\leq \limsup_{k\to+\infty}\left\{D_{\alpha_k}\left(\mathcal{W}\|\sigma_k|P_\infty\right)\right\} + \limsup_{k\to+\infty}\left\{\sum_{x\in\mathcal{X}}\left[P_k(x) - P_\infty(x)\right] D_{\alpha_k}\left(W_x\|\sigma_k\right)\right\} && (270)\\
&= D_{\alpha_\infty}\left(\mathcal{W}\|\sigma_\infty^\star|P_\infty\right) + \limsup_{k\to+\infty}\left\{\sum_{x\in\mathcal{X}}\left[P_k(x) - P_\infty(x)\right] D_{\alpha_k}\left(W_x\|\sigma_k\right)\right\} && (271)\\
&= I_{\alpha_\infty}^{(2)}(P_\infty, \mathcal{W}) + \limsup_{k\to+\infty}\left\{\sum_{x\in\mathcal{X}}\left[P_k(x) - P_\infty(x)\right] D_{\alpha_k}\left(W_x\|\sigma_k\right)\right\}. && (272)
\end{aligned}$$

Here, Eq. (269) follows from the definition in Eq. (23). Inequality (270) holds because the superior limit of sum is smaller than the sum of superior limits. Equality (271) holds because $\sigma_k \gg W_x$ for all $x \in \texttt{supp}(P_\infty)$ and $k \in \mathbb{N}\cup\{+\infty\}$. Thus, the map $(\alpha_k, \sigma_k) \mapsto D_{\alpha_k}(\mathcal{W}\|\sigma_k|P_\infty)$ is continuous for $k \in \mathbb{N}\cup\{+\infty\}$.

It remains to show the second term in Eq. (272) is actually zero. The definition in Eq. (23) and direct calculation show that

$$\limsup_{k\to+\infty}\left\{\sum_{x\in\mathcal{X}}\left[P_k(x)-P_\infty(x)\right]D_{\alpha_k}\left(W_x\|\sigma_k\right)\right\} \tag{273}$$

$$\leq \limsup_{k\to+\infty}\left\{\epsilon_k\cdot\max_{x\in\mathcal{X}}D_{\alpha_k}(W_x\|\sigma_k)\right\} \tag{274}$$

$$\leq \limsup_{k\to+\infty}\left\{\epsilon_k\cdot\max_{x\in\mathcal{X}}D_{\alpha_k}\left(W_x\middle\|\epsilon_k\frac{\mathbb{1}}{d}\right)\right\} \tag{275}$$

$$= \limsup_{k\to+\infty}\left\{\epsilon_k\cdot\left[\log\epsilon_k+\max_{x\in\mathcal{X}}D_{\alpha_k}\left(W_x\middle\|\frac{\mathbb{1}}{d}\right)\right]\right\} \tag{276}$$

$$= \limsup_{k\to+\infty}\epsilon_k\log\epsilon_k \tag{277}$$

$$= 0, \tag{278}$$

where Eq. (275) follows from the dominance of α-Rényi divergence [62, Section 4]; in the last equality (278) we use the convention $\lim_{\epsilon_k\downarrow 0}\epsilon_k\log\epsilon_k=0$ and $\lim_{k\to+\infty}P_k=P_\infty$. Hence, item (a) is proven. □

Proof of Proposition 2-(b). Recall the definition in Eq. (29). The statement immediately follows from Eq. (18) (see also [44, Lemma IV.5]) because the minimization over $\sigma\in\mathcal{S}(\mathcal{H})$ preserves the monotonicity. □

Proof of Proposition 2-(c). The claim was proven by Mosonyi and Ogawa [44, Appendix B]. □

Proof of Proposition 2-(d). The map $\alpha\mapsto C_{\alpha,\mathcal{W}}$ is continuous and monotone increasing on $[0,1]$. Berge's maximum theorem [63, Section IV.3], [64, Lemma 3.1] shows that the continuous map $(\alpha,P)\mapsto I_\alpha^{(2)}(P,\mathcal{W})$ maximized over the compact set $P\in\mathcal{P}(\mathcal{X})$ is still continuous for $\alpha\in[0,1]$.

Lastly, we show the the assertions for $I_\alpha^{(1)}(P,\mathcal{W})$. Quantum Sibson's identity [51] implies that $I_\alpha^{(1)}(P,\mathcal{W})=\frac{\alpha}{1-\alpha}E_0((1-\alpha)/\alpha,P)$ for $\alpha\in[0,1)$, where E_0 is defined in Eq. (34). Items (a) and (b) hold directly. Item (c) follows from the concavity of $s\mapsto E_0(s,P)$ for all $s\geq 0$ [15]. □

Appendix D. Proof of Proposition 3

Proposition 3 (Saddle-Point)**.** *Consider a classical-quantum channel $\mathcal{W}:\mathcal{X}\to\mathcal{S}(\mathcal{H})$, any $R\in(R_\infty,C_\mathcal{W})$, and $P\in\mathcal{P}(\mathcal{X})$. Let*

$$\mathcal{S}_{P,\mathcal{W}}(\mathcal{H}):=\left\{\sigma\in\mathcal{S}(\mathcal{H}):\forall x\in\texttt{supp}(P),\ W_x\not\perp\sigma\right\}. \tag{279}$$

Define

$$F_{R,P}(\alpha,\sigma):=\begin{cases}\dfrac{1-\alpha}{\alpha}\left(D_\alpha\left(\mathcal{W}\|\sigma|P\right)-R\right), & \alpha\in(0,1)\\ 0, & \alpha=1\end{cases}, \tag{280}$$

on $(0,1]\times\mathcal{S}(\mathcal{H})$, and denote by

$$\mathcal{P}_R(\mathcal{X}):=\left\{P\in\mathcal{P}(X):\sup_{0<\alpha\leq 1}\inf_{\sigma\in\mathcal{S}(\mathcal{H})}F_{R,P}(\alpha,\sigma)\in\mathbb{R}_{>0}\right\}. \tag{281}$$

The following holds

(a) *For any $P\in\mathcal{P}(\mathcal{X})$, $F_{R,P}(\cdot,\cdot)$ has a saddle-point on $(0,1]\times\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$ with the saddle-value:*

$$\min_{\sigma\in\mathcal{S}(\mathcal{H})}\sup_{0<\alpha\leq 1}F_{R,P}(\alpha,\sigma)=\sup_{0<\alpha\leq 1}\min_{\sigma\in\mathcal{S}(\mathcal{H})}F_{R,P}(\alpha,\sigma)=E_{\text{sp}}^{(2)}(R,P). \tag{282}$$

(b) *If $P\in\mathcal{P}_R(\mathcal{X})$, the saddle-point is unique.*

(c) *Fix $P\in\mathcal{P}_R(\mathcal{X})$. Any saddle-point $(\alpha_{R,P}^\star,\sigma_{R,P}^\star)$ of $F_{R,P}(\cdot,\cdot)$ satisfies $\alpha_{R,P}^\star\in(0,1)$ and*

$$\sigma_{R,P}^\star\gg W_x,\quad\forall x\in\texttt{supp}(P). \tag{283}$$

Proof of Proposition 3-(a). Fix arbitrary $R > R_\infty$ and $P \in \mathcal{P}(\mathcal{X})$. In the following, we prove the existence of a saddle-point of $F_{R,P}(\cdot,\cdot)$ on $(0,1] \times \mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$. Ref. [65, Lemma 36.2] states that $(\alpha^\star, \sigma^\star)$ is a saddle point of $F_{R,P}(\cdot,\cdot)$ if and only if the supremum in

$$\sup_{\alpha\in(0,1]} \inf_{\sigma\in\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})} F_{R,P}(\alpha,\sigma) \tag{284}$$

is attained at $\alpha^\star \in (0,1]$, the infimum in

$$\inf_{\sigma\in\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})} \sup_{\alpha\in(0,1]} F_{R,P}(\alpha,\sigma) \tag{285}$$

is attained at $\sigma^\star \in \mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$, and the two extrema in Eqs. (284), (285) are equal and finite. We first claim that, $\forall\alpha \in (0,1]$,

$$\inf_{\sigma\in\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})} F_{R,P}(\alpha,\sigma) = \inf_{\sigma\in\mathcal{S}(\mathcal{H})} F_{R,P}(\alpha,\sigma). \tag{286}$$

To see this, observe that for any $\alpha \in (0,1)$, Eqs. (11) and (24) yield

$$\forall\sigma \in \mathcal{S}(\mathcal{H})\backslash\mathcal{S}_{P,\mathcal{W}}(\mathcal{H}), \quad D_\alpha\left(\mathcal{W}\|\sigma|P\right) = +\infty, \tag{287}$$

which, in turn, implies

$$\forall\sigma \in \mathcal{S}(\mathcal{H})\backslash\mathcal{S}_{P,\mathcal{W}}(\mathcal{H}), \quad F_{R,P}(\alpha,\sigma) = +\infty. \tag{288}$$

Further, Eq. (286) holds trivially when $\alpha = 1$. Hence, Eq. (286) yields

$$\sup_{\alpha\in(0,1]} \inf_{\sigma\in\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})} F_{R,P}(\alpha,\sigma) = \sup_{\alpha\in(0,1]} \inf_{\sigma\in\mathcal{S}(\mathcal{H})} F_{R,P}(\alpha,\sigma) \tag{289}$$

Owing to the fact $R > R_\infty$ and Eq. (38), we have

$$E_{\text{sp}}^{(2)}(R,P) = \sup_{\alpha\in(0,1]} \inf_{\sigma\in\mathcal{S}(\mathcal{H})} F_{R,P}(\alpha,\sigma) < +\infty, \tag{290}$$

which guarantees the supremum in the right-hand side of Eq. (290) is attained at some $\alpha \in (0,1]$. Namely, there exists some $\bar{\alpha}_{R,P} \in (0,1]$ such that

$$\sup_{\alpha\in(0,1]} \inf_{\sigma\in\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})} F_{R,P}(\alpha,\sigma) = \max_{\alpha\in[\bar{\alpha}_{R,P},1]} \inf_{\sigma\in\mathcal{S}(\mathcal{H})} F_{R,P}(\alpha,\sigma) < +\infty. \tag{291}$$

Thus, we complete our claim in Eq. (284). It remains to show that the infimum in Eq.(285) is attained at some $\sigma^\star \in \mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$ and the supremum and infimum are exchangeable. To achieve this, we will show that $\left([\bar{\alpha}_{R,P},1], \mathcal{S}_{P,\mathcal{W}}(\mathcal{H}), F_{R,P}\right)$ is a closed saddle-element (see Definition 19 below) and employ the boundness of $[\bar{\alpha}_{R,P},1] \times \mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$ to conclude our claim.

Definition 19 (Closed Saddle-Element [65])**.** We denote by $\texttt{ri}$ and $\texttt{cl}$ the relative interior and the closure of a set, respectively. Let $\mathcal{A}, \mathcal{B}$ be subsets of a real vector space, and $F : \mathcal{A}\times\mathcal{B} \to \mathbb{R}\cup\{\pm\infty\}$. The triple $(\mathcal{A},\mathcal{B},F)$ is called a closed saddle-element if for any $x \in \texttt{ri}\,(\mathcal{A})$ (resp. $y \in \texttt{ri}\,(\mathcal{B})$),

(i) $\mathcal{B}$ (resp. $\mathcal{A}$) is convex.

(ii) $F(x,\cdot)$ (resp. $F(\cdot,y)$) is convex (resp. concave) and lower (resp. upper) semi-continuous.

(iii) Any accumulation point of $\mathcal{B}$ (resp. $\mathcal{A}$) that does not belong to $\mathcal{B}$ (resp. $\mathcal{A}$), say y_o (resp. x_o) satisfies $\lim_{y\to y_o} F(x,y) = +\infty$ (resp. $\lim_{x\to x_o} F(x,y) = -\infty$).

Fix an arbitrary $\alpha \in \texttt{ri}\,([\bar{\alpha}_{R,P},1]) = (\bar{\alpha}_{R,P},1)$. We check that $\left(\mathcal{S}_{P,\mathcal{W}}(\mathcal{H}), F_{R,P}(\alpha,\cdot)\right)$ fulfills the three items in Definition 19. (i) The set $\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$ is clearly convex. (ii) Eq. (20) in Lemma 1 implies that $\sigma \mapsto D_\alpha(W_x\|\sigma)$ is convex and lower semi-continuous. Since convex combination preservers the convexity and the lower semi-continuity, Eq. (280) yields that $\sigma \mapsto F_{R,P}(\alpha,\sigma)$ is convex and lower semi-continuous on $\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$. (iii) Due to the compactness of $\mathcal{S}(\mathcal{H})$, any accumulation point of $\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$ that does not belong to $\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$, say σ_o, satisfies $\sigma_o \in \mathcal{S}(\mathcal{H})\backslash\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$. Eqs. (287) and (288) then show that $F_{R,P}(\alpha,\sigma_o) = +\infty$.

Next, fix an arbitrary $\sigma \in \mathtt{ri}\left(\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})\right)$. Owing to the convexity of $\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$, it follows that $\mathtt{ri}\left(\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})\right) = \mathtt{ri}\left(\mathtt{cl}\left(\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})\right)\right)$ (see e.g. [66, Theorem 6.3]). We first claim $\mathtt{cl}\left(\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})\right) = \mathcal{S}(\mathcal{H})$. To see this, observe that $\mathcal{S}_{>0}(\mathcal{H}) \subseteq \mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$ since a full-rank density operator is not orthogonal with every W_x, $x \in \mathcal{X}$. Hence,

$$\mathcal{S}(\mathcal{H}) = \mathtt{cl}\left(\mathcal{S}_{>0}(\mathcal{H})\right) \subseteq \mathtt{cl}\left(\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})\right). \tag{292}$$

On the other hand, the fact $\mathcal{S}_{P,\mathcal{W}}(\mathcal{H}) \subseteq \mathcal{S}(\mathcal{H})$ leads to

$$\mathtt{cl}\left(\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})\right) \subseteq \mathtt{cl}\left(\mathcal{S}(\mathcal{H})\right) = \mathcal{S}(\mathcal{H}). \tag{293}$$

By Eqs. (292) and (293), we deduce that

$$\mathtt{ri}\left(\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})\right) = \mathtt{ri}\left(\mathtt{cl}\left(\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})\right)\right) = \mathtt{ri}\left(\mathcal{S}(\mathcal{H})\right) = \mathcal{S}_{>0}(\mathcal{H}), \tag{294}$$

where the last equality in Eq. (294) follows from [67, Proposition 2.9]. Hence, we obtain

$$\forall \sigma \in \mathtt{ri}\left(\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})\right) \quad \text{and} \quad \forall x \in \mathcal{X}, \quad \sigma \gg W_x. \tag{295}$$

Now we verify that $([\bar{\alpha}_{R,P}, 1], F_{R,P}(\cdot,\sigma))$ satisfies the three items in Definition 19. Fix an arbitrary $\sigma \in \mathtt{ri}\left(\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})\right)$. (i) The set $(0,1]$ is obviously convex. (ii) From Eq. (18) in Lemma 1, the map $\alpha \mapsto F_{R,P}(\alpha,\sigma)$ is continuous on $(0,1)$. Further, it is not hard to verify that $F_{R,P}(1,\sigma) = 0 = \lim_{\alpha\uparrow 1} F_{R,P}(\alpha,\sigma)$ from Eqs. (295), (280), and (11). Item (c) in Proposition 2 implies that $\alpha \mapsto F_{R,P}(\alpha,\sigma)$ on $[\bar{\alpha}_R, 1)$ is concave. Moreover, the continuity of $\alpha \mapsto F_{R,P}(\alpha,\sigma)$ on $[\bar{\alpha}_{R,P}, 1)$ guarantees the concavity of $\alpha \mapsto F_{R,P}(\alpha,\sigma)$ on $[\bar{\alpha}_{R,P}, 1]$. (iii) Since $[\bar{\alpha}_{R,P}, 1]$ is closed, there is no accumulation point of $[\bar{\alpha}_{R,P}, 1]$ that does not belong to $[\bar{\alpha}_{R,P}, 1]$.

We are at the position to prove item (a) of Proposition 3. The closed saddle-element, along with the boundness of $\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$ and Rockafellar's saddle-point result [65, Theorem 8], [66, Theorem 37.3] imply that

$$-\infty < \sup_{\alpha\in[\bar{\alpha}_{R,P},1]} \inf_{\sigma\in\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})} F_{R,P}(s,\sigma) = \min_{\sigma\in\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})} \sup_{\alpha\in[\bar{\alpha}_{R,P},1]} F_{R,P}(s,\sigma). \tag{296}$$

Then Eqs. (291) and (296) lead to the existence of a saddle-point of $F_{R,P}(\cdot,\cdot)$ on $(0,1]\times\mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$. Hence, item (a) is proved.

□

Proof of Proposition 3-(b). Fix arbitrary $R \in (R_\infty, C_\mathcal{W})$ and $P \in \mathcal{P}_R(\mathcal{X})$. We have

$$\sup_{0<\alpha\leq 1} \min_{\sigma\in\mathcal{S}(\mathcal{H})} F_{R,P}(\alpha,\sigma) \in \mathbb{R}_{>0}. \tag{297}$$

First note that $\alpha^\star = 1$ will not be a saddle point of $F_{R,P}(\cdot,\sigma)$ because $F_{R,P}(1,\sigma) = 0$, $\forall \sigma \in \mathcal{S}(\mathcal{H})$, contradicting Eq. (297).

Now, fix $\alpha^\star \in (0,1)$ to be a saddle-point of $F_{R,P}(\cdot,\cdot)$. Eq. (20) in Lemma 1 implies that the map $\sigma \mapsto D_{\alpha^\star}(\mathcal{W}\|\sigma|P)$ is strictly convex, and thus the minimizer of Eq. (297) is unique. Next, let $\sigma^\star \in \mathcal{S}_{P,\mathcal{W}}(\mathcal{H})$ be a saddle-point of $F_{R,P}(\cdot,\cdot)$. Then,

$$F_{R,P}(\alpha,\sigma^\star) = \frac{1-\alpha}{\alpha}\left(I_\alpha^{(2)}(P,\mathcal{W}) - R\right). \tag{298}$$

Item (c) in Proposition 2 then shows that $\frac{1-\alpha}{\alpha} I_\alpha^{(2)}(P,\mathcal{W})$ is strictly concave on $(0,1)$, which in turn implies that $F_{R,P}(\cdot,\sigma^\star)$ is also strictly concave on $(0,1)$. Hence, the maximizer of Eq. (297) is unique. □

Proof of Proposition 3-(c). As shown in the proof of item (b), $\alpha^\star = 1$ is not a saddle point of $F_{R,P}(\cdot,\cdot)$ for any $R > R_\infty$ and $P \in \mathcal{P}_R(\mathcal{X})$. We assume $(\alpha^\star,\sigma^\star)$ is a saddle-point of $F_{R,P}(\cdot,\cdot)$ with $\alpha^\star \in (0,1)$, it holds that

$$F_{R,P}(\alpha^\star,\sigma^\star) = \min_{\sigma\in\mathcal{S}(\mathcal{H})} F_{R,P}(\alpha^\star,\sigma) = \frac{\alpha^\star - 1}{\alpha^\star} R + \frac{1-\alpha^\star}{\alpha^\star} \min_{\sigma\in\mathcal{S}(\mathcal{H})} D_{\alpha^\star}(\mathcal{W}\|\sigma|P). \tag{299}$$

We claim that the minimizer of Eq. (299) must satisfy

$$\sigma^\star = \frac{\left(\sum_{x\in\mathcal{X}} P(x)\frac{W_x^{\alpha^\star}}{\operatorname{Tr}\left[W_x^{\alpha^\star}(\sigma^\star)^{1-\alpha^\star}\right]}\right)^{\frac{1}{\alpha^\star}}}{\operatorname{Tr}\left[\left(\sum_{x\in\mathcal{X}} P(x)\frac{W_x^{\alpha^\star}}{\operatorname{Tr}\left[W_x^{\alpha^\star}(\sigma^\star)^{1-\alpha^\star}\right]}\right)^{\frac{1}{\alpha^\star}}\right]}. \tag{300}$$

Our approach follows closely from Hayashi and Tomamichel [49, Lemma 5]. Observe that

$$\arg\min_{\sigma\in\mathcal{S}(\mathcal{H})} D_\alpha\left(\mathcal{W}\|\sigma|P\right) = \arg\max_{\sigma\in\mathcal{S}(\mathcal{H})} g_\alpha(\sigma), \quad \forall\alpha\in(0,1), \tag{301}$$

where

$$g_\alpha(\sigma) := \sum_{x\in\mathcal{X}} P(x)\log\operatorname{Tr}\left[W_x^\alpha\sigma^{1-\alpha}\right]. \tag{302}$$

Note that the map $\sigma\mapsto g_\alpha(\sigma)$ is strictly concave for every $\alpha\in(0,1)$ by Eq. (20) in Lemma 1. A sufficient and necessary condition for σ to be an optimizer of Eq. (301) is

$$\partial_\omega g_\alpha(\sigma) := \mathsf{D}g_\alpha(\sigma)[\omega-\sigma] = 0, \tag{303}$$

for all $\omega\in\mathcal{S}(\mathcal{H})$, where $\mathsf{D}g_\alpha(\sigma)$ denotes the Fréchet derivative of the map g_α (see e.g. [49, Appendix C], [68, 69, 70, 71]). Direct calculation shows that

$$\partial_\omega g_\alpha(\sigma) = \operatorname{Tr}\left[\sum_{x\in\mathcal{X}} P(x)\frac{W_x^\alpha}{\operatorname{Tr}\left[W_x^\alpha\sigma^{1-\alpha}\right]}\partial_\omega\sigma^{1-\alpha}\right]. \tag{304}$$

Next, we check that the fixed-points of the following map attains Eq. (303):

$$\sigma\mapsto\frac{\left(\sum_{x\in\mathcal{X}} P(x)\frac{W_x^\alpha}{\operatorname{Tr}\left[W_x^\alpha\sigma^{1-\alpha}\right]}\right)^{\frac{1}{\alpha}}}{\kappa_\alpha(\sigma)}, \tag{305}$$

where $\kappa_\alpha(\sigma)$ denotes a finite normalization constant. Let $\bar\sigma$ be a fix-point of the map in Eq. (305). Then Eqs. (304) and (305) yield

$$\begin{aligned}\partial_\omega g_\alpha(\bar\sigma) &= \operatorname{Tr}\left[\kappa_\alpha(\bar\sigma)^\alpha\bar\sigma^\alpha\partial_\omega\bar\sigma^{1-\alpha}\right] = \operatorname{Tr}\left[\kappa_\alpha(\bar\sigma)^\alpha\bar\sigma^\alpha(1-\alpha)\bar\sigma^{-\alpha}(\omega-\bar\sigma)\right]\\ &= (1-\alpha)\kappa_\alpha(\bar\sigma)^\alpha\operatorname{Tr}\left[\omega-\bar\sigma\right] = 0.\end{aligned} \tag{306}$$

By Brouwer's fixed-point theorem, the map in Eq. (305) is indeed the optimizer for Eq. (301). Further, it is clear from Eq. (300) that

$$\sigma^\star \gg W_x, \quad \forall x\in\mathsf{supp}(P), \tag{307}$$

and thus item (c) is proved.

□

Appendix E. Proof of Proposition 4

Proposition 4 (Properties of Error-Exponent Functions)**.** *Consider a classical-quantum channel $\mathcal{W}:\mathcal{X}\to\mathcal{S}(\mathcal{H})$ with $R_\infty<C_\mathcal{W}$. We have*

(a) *Given every $P\in\mathcal{P}(\mathcal{X})$, $E_{\text{sp}}^{(2)}(\cdot,P)$ is convex and non-increasing on $[0,+\infty]$, and continuous on $\left[I_0^{(2)}(P,\mathcal{W}),+\infty\right]$. For every $R>R_\infty$, $E_{\text{sp}}^{(2)}(R,\cdot)$ is continuous on $\mathcal{P}(\mathcal{X})$. Further,*

$$E_{\text{sp}}^{(2)}(R,P) = \begin{cases}+\infty, & R<I_0^{(2)}(P,\mathcal{W})\\ 0, & R\geq I_1^{(2)}(P,\mathcal{W})\end{cases}. \tag{308}$$

(b) *$E_{\mathrm{sp}}(\cdot)$ is convex and non-increasing on $[0, +\infty]$, and continuous on $[R_\infty, +\infty]$. Further,*

$$E_{\mathrm{sp}}(R) = \begin{cases} +\infty, & R < R_\infty \\ 0, & R \geq C_{\mathcal{W}} \end{cases}. \tag{309}$$

(c) *Consider any $R \in (R_\infty, C_{\mathcal{W}})$ and $P \in \mathcal{P}_R(\mathcal{X})$ (see Eq. (44)). The function $E_{\mathrm{sp}}^{(2)}(\cdot, P)$ is differentiable with*

$$s^\star_{R,P} = - \left. \frac{\partial E_{\mathrm{sp}}^{(2)}(r, P)}{\partial r} \right|_{r=R} \in \mathbb{R}_{>0}, \tag{310}$$

where $s^\star_{R,P} := (1 - \alpha^\star_{R,P})/\alpha^\star_{R,P}$, and $\alpha^\star_{R,P}$ is the optimizer in Eq. (38).

(d) *$s^\star_{R,(\cdot)}$ in Eq. (310) is continuous on $\mathcal{P}_R(\mathcal{X})$.*

Proof of Proposition 4-(a). Fix any arbitrary $P \in \mathcal{P}(\mathcal{X})$. Item (b) in Proposition 2 shows that the map $\alpha \mapsto I_\alpha^{(2)}(P, \mathcal{W})$ is monotone increasing on $[0, 1]$. Hence, from the definition in Eq. (38), it is not hard to verify that $E_{\mathrm{sp}}^{(2)}(R, P) = +\infty$ for all $R \in (0, I_0^{(2)}(P, \mathcal{W}))$; finite for all $R > I_0^{(2)}(P, \mathcal{W})$; and $E_{\mathrm{sp}}^{(2)}(R, P) = 0$, for all $R \geq I_1^{(2)}(P, \mathcal{W})$.

For every $\alpha \in (0, 1]$, the function $\frac{1-\alpha}{\alpha}(I_\alpha^{(2)}(P, \mathcal{W}) - R)$ in Eq. (38) is an non-increasing, convex, and continuous function in $R \in \mathbb{R}_{>0}$. Since $E_{\mathrm{sp}}^{(2)}(R, P)$ is the pointwise supremum of the above function, $E_{\mathrm{sp}}^{(2)}(R, P)$ is non-increasing, convex, and lower semi-continuous function for all $R \geq 0$. Furthermore, since a convex function is continuous on the interior of the interval if it is finite [72, Corollary 6.3.3], thus $E_{\mathrm{sp}}^{(2)}(R, P)$ is continuous for all $R > I_0^{(2)}(P, \mathcal{W})$, and continuous from the right at $R = I_0^{(2)}(P, \mathcal{W})$.

To establish the continuity of $E_{\mathrm{sp}}^{(2)}(R, P)$ in $P \in \mathcal{P}(\mathcal{X})$, we first claim that there exists some $\bar{\alpha}_R \in (0, 1]$ such that for every $P \in \mathcal{P}(\mathcal{X})$,

$$\sup_{\alpha \in (0,1]} \frac{1-\alpha}{\alpha} \left(I_\alpha^{(2)}(P, \mathcal{W}) - R \right) = \sup_{\alpha \in [\bar{\alpha}_R, 1]} \frac{1-\alpha}{\alpha} \left(I_\alpha^{(2)}(P, \mathcal{W}) - R \right). \tag{311}$$

Recall that $R > R_\infty = \max_{P \in \mathcal{P}(\mathcal{X})} I_0^{(2)}(P, \mathcal{W})$. The continuity, item (a) in Proposition 2, implies that there is an $\bar{\alpha}_R > 0$ such that

$$R \geq I_{\bar{\alpha}_R}^{(2)}(P, \mathcal{W}), \quad \forall P \in \mathcal{P}(\mathcal{X}). \tag{312}$$

Then, Eq. (312) and the monotone increases of the map $\alpha \mapsto I_\alpha^{(2)}(P, \mathcal{W})$ yield that,

$$\frac{1-\alpha}{\alpha} \left(I_\alpha^{(2)}(P, \mathcal{W}) - R \right) < 0, \quad \forall P \in \mathcal{P}(\mathcal{X}), \text{ and } \alpha \in (0, \bar{\alpha}_R). \tag{313}$$

The non-negativity of $E_{\mathrm{sp}}^{(2)}(R, P) \geq 0$ ensures that the maximizer $\alpha^\star$ will not happen in the region $(0, \bar{\alpha}_R)$, and thus Eq. (311) is evident. Finally, Berge's maximum theorem [63, Section IV.3], [64, Lemma 3.1] coupled with the compactness of $[\bar{\alpha}_R, 1]$ and item (a) in Proposition 2 complete our claim:

$$P \mapsto E_{\mathrm{sp}}^{(2)}(R, P) = \sup_{\alpha \in [\bar{\alpha}_R, 1]} \frac{1-\alpha}{\alpha} \left(I_\alpha^{(2)}(P, \mathcal{W}) - R \right) \text{ is continuous on } \mathcal{P}(\mathcal{X}). \tag{314}$$

□

Proof of Proposition 4-(b). The statement follows since item (a) holds for any $P \in \mathcal{P}(\mathcal{X})$. □

Proof of Proposition 4-(c). For any $R \in (R_\infty, C_{\mathcal{W}})$ and $P \in \mathcal{P}_R(\mathcal{X})$, item (b) in Proposition 3 shows that the optimizer $\alpha^\star_{R,P}$ is unique. Moreover, Eq. (310) follows from item (d) in Lemma 17.

□

Proof of Proposition 4-(d). The proof of this item is similar to [32, Proposition 3.4]. Fix any $P_o \in \mathcal{P}_R(\mathcal{X})$ and consider arbitrary $\{P_k\}_{k\in\mathbb{N}}$ such that $P_k \in \mathcal{P}_R(\mathcal{X})$, $\forall k \in \mathbb{N}$, and $\lim_{n\to+\infty} P_k = P_o$. Following from Eq. (310), we have

$$s^\star_{R,P_k} = -\left.\frac{\partial E^{(2)}_{\text{sp}}(r, P_k)}{\partial r}\right|_{r=R}. \tag{315}$$

Given any $R \in (R_\infty, C_\mathcal{W})$, the continuity of $E^{(2)}_{\text{sp}}(R, \cdot)$ (see item (a)) implies that

$$\lim_{k\to+\infty} E^{(2)}_{\text{sp}}(R, P_k) = E^{(2)}_{\text{sp}}(R, P_o). \tag{316}$$

Then, continuity of the first-order derivative in [73, Corollary VI.6.2.8], we have

$$\lim_{k\to+\infty} s^\star_{R,P_k} = \lim_{k\to+\infty} -\left.\frac{\partial E^{(2)}_{\text{sp}}(r, P_k)}{\partial r}\right|_{r=R} = -\left.\frac{\partial E^{(2)}_{\text{sp}}(r, P_o)}{\partial r}\right|_{r=R} = s^\star_{R,P_o}, \tag{317}$$

which completes the proof. □

References

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[2] ——, "Probability of error for optimal codes in a Gaussian channel," *Bell System Technical Journal*, vol. 38, no. 3, pp. 611–656, may 1959.

[3] A. Feinstein, "Error bounds in noisy channels without memory," *IEEE Transactions on Information Theory*, vol. 1, no. 2, pp. 13–14, sep 1955.

[4] R. M. Fano, *Transmission of Information, A Statistical Theory of Communications.* The MIT Press, 1961.

[5] R. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Transaction on Information Theory*, vol. 11, no. 1, pp. 3–18, jan 1965.

[6] ——, *Information Theory and Reliable Communication.* Wiley, 1968. [Online]. Available: http://as.wiley.com/WileyCDA/WileyTitle/productCd-0471290483.html

[7] C. Shannon, R. Gallager, and E. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. I," *Information and Control*, vol. 10, no. 1, pp. 65–103, jan 1967.

[8] E. A. Haroutunian, "Estimates of the error exponents for the semicontinuous memoryless channel," *Problemy Peredachi Informatsii*, vol. 4, no. 4, pp. 37–48, 1968, (in Russian). [Online]. Available: http://mi.mathnet.ru/eng/ppi1871

[9] E. A. Haroutunian, M. E. Haroutunian, and A. N. Harutyunyan, "Reliability criteria in information theory and in statistical hypothesis testing," *Foundations and Trends® in Communications and Information Theory*, vol. 4, no. 2–3, pp. 97–263, 2007.

[10] R. E. Blahut, "Hypothesis testing and information theory," *IEEE Transaction on Information Theory*, vol. 20, no. 4, pp. 405–417, jul 1974.

[11] A. Ben-Tal, M. Teboulle, and A. Charnes, "The role of duality in optimization problems involving entropy functionals with applications to information theory," *Journal of Optimization Theory and Applications*, vol. 58, no. 2, pp. 209–223, aug 1988.

[12] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Cambridge University Press (CUP), 2011.

[13] M. V. Burnashev and A. S. Holevo, "On the reliability function for a quantum communication channel," *Problems of information transmission*, vol. 34, no. 2, pp. 97–107, 1998.

[14] A. Holevo, "Reliability function of general classical-quantum channel," *IEEE Transaction on Information Theory*, vol. 46, no. 6, pp. 2256–2261, 2000.

[15] H.-C. Cheng and M.-H. Hsieh, "Concavity of the auxiliary function for classical-quantum channels," *IEEE Transactions on Information Theory*, vol. 62, no. 10, pp. 5960 – 5965, 2016.

[16] A. Winter, "Coding theorems of quantum information theory," *PhD Thesis, Universität Bielefeld*, 1999.

[17] M. Dalai, "Lower bounds on the probability of error for classical and classical-quantum channels," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8027–8056, dec 2013.

[18] M. Dalai and A. Winter, "Constant compositions in the sphere packing bound for classical-quantum channels," in *2014 IEEE International Symposium on Information Theory.* Institute of Electrical & Electronics Engineers (IEEE), jun 2014.

[19] D. Petz, "Quasi-entropies for finite quantum systems," *Reports on Mathematical Physics*, vol. 23, no. 1, pp. 57–65, feb 1986.

[20] S. Golden, "Lower bounds for the Helmholtz function," *Physical Review*, vol. 137, no. 4B, pp. B1127–B1128, feb 1965.

[21] C. J. Thompson, "Inequality with applications in statistical mechanics," *Journal of Mathematical Physics*, vol. 6, no. 11, p. 1812, 1965.

[22] Y. Altuğ and A. B. Wagner, "A refinement of the random coding bound," in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton).* Institute of Electrical and Electronics Engineers (IEEE), oct 2012.

[23] J. Scarlett, A. Martinez, and A. Guillén i F'abregas, "Mismatched decoding: Error exponents, second-order rates and saddlepoint approximations," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2647–2666, may 2014.

[24] J. Scarlett, "Reliable communication under mismatched decoding," *PhD Thesis (University of Cambridge)*, 2014.

[25] J. Honda, "Exact asymptotics for the random coding error probability," `arXiv:1506.03355 [cs.IT]`.

[26] R. R. Bahadur and R. R. Rao, "On deviations of the sample mean," *The Annals of Mathematical Statistics*, vol. 31, no. 4, pp. 1015–1027, dec 1960.

[27] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications.* Springer, 1998.

[28] Y. Altuğ and A. B. Wagner, "Moderate deviations in channel coding," *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4417–4426, aug 2014.

[29] H.-C. Cheng and M.-H. Hsieh, "Moderate deviation analysis for classical-quantum channels and quantum hypothesis testing," `arXiv:1701.03195 [quant-ph]`.

[30] C. T. Chubb, V. Y. F. Tan, and M. Tomamichel, "Moderate deviation analysis for classical communication over quantum channels," `arXiv:1701.03114 [quant-ph]`.

[31] J. K. Omura, "A lower bounding method for channel and source coding probabilities," *Information and Control*, vol. 27, no. 2, pp. 148–177, feb 1975.

[32] Y. Altuğ and A. B. Wagner, "Refinement of the sphere-packing bound: Asymmetric channels," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1592–1614, mar 2014.

[33] N. Elkayam and M. Feder, "Sphere packing bound for constant composition," 2016, (in preparation).

[34] U. Augustin, "Error estimates for low rate codes," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 14, no. 1, pp. 61–88, 1969.

[35] ——, "Noisy channels," 1978, habilitation thesis, Universitat Erlangen.

[36] B. Nakiboğlu, "Augustin's method - part I: The renyi center," `arXiv:1608.02424 [cs.IT]`.

[37] ——, "Augustin's method - part II: The sphere packing bound," `arXiv:1611.06924 [cs.IT]`.

[38] H. Umegaki, "Conditional expectation in an operator algebra. IV. entropy and information," *Kodai Mathematical Seminar Reports*, vol. 14, no. 2, pp. 59–85, 1962.

[39] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Communications in Mathematical Physics*, vol. 143, no. 1, pp. 99–114, dec 1991.

[40] M. Tomamichel and M. Hayashi, "A hierarchy of information quantities for finite block length analysis of quantum tasks," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7693–7710, nov 2013.

[41] K. Li, "Second-order asymptotics for quantum hypothesis testing," *The Annals of Statistics*, vol. 42, no. 1, pp. 171–189, feb 2014.

[42] M. Tomamichel and V. Y. F. Tan, "Second-order asymptotics for the classical capacity of image-additive quantum channels," *Communications in Mathematical Physics*, vol. 338, no. 1, pp. 103–137, may 2015.

[43] T. Ogawa and H. Nagaoka, “Strong converse and Stein’s lemma in quantum hypothesis testing,” *IEEE Transaction on Information Theory*, vol. 46, no. 7, pp. 2428–2433, 2000.

[44] M. Mosonyi and T. Ogawa, “Strong converse exponent for classical-quantum channel coding,” `arXiv:1409.3562 [quant-ph]`.

[45] E. H. Lieb, “Convex trace functions and the Wigner-Yanase-Dyson conjecture,” *Advances in Mathematics*, vol. 11, no. 3, pp. 267–288, dec 1973.

[46] F. Hiai, “Concavity of certain matrix trace and norm functions. II,” *Linear Algebra and its Applications*, vol. 496, pp. 193–220, may 2016.

[47] B. Schumacher and M. D. Westmoreland, “Sending classical information via noisy quantum channels,” *Physical Review A*, vol. 56, no. 1, pp. 131–138, jul 1997.

[48] A. Holevo, “The capacity of the quantum channel with general signal states,” *IEEE Transaction on Information Theory*, vol. 44, no. 1, pp. 269–273, 1998.

[49] M. Hayashi and M. Tomamichel, “Correlation detection and an operational interpretation of the rényi mutual information,” *Journal of Mathematical Physics*, vol. 57, no. 10, p. 102201, oct 2016.

[50] M. M. Wilde, A. Winter, and D. Yang, “Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy,” *Communications in Mathematical Physics*, vol. 331, no. 2, pp. 593–622, jul 2014.

[51] N. Sharma and N. A. Warsi, “Fundamental bound on the reliability of quantum information transmission,” *Physical Review Letters*, vol. 110, no. 8, feb 2013.

[52] M. Hayashi, *Quantum Information: An Introduction.* Springer, 2006.

[53] H. Nagaoka and M. Hayashi, “An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses,” *IEEE Transactions on Information Theory*, vol. 53, no. 2, pp. 534–549, feb 2007.

[54] K. M. R. Audenaert, M. Nussbaum, A. Szkoła, and F. Verstraete, “Asymptotic error rates in quantum hypothesis testing,” *Communications in Mathematical Physics*, vol. 279, no. 1, pp. 251–283, feb 2008.

[55] M. Nussbaum and A. Szkoła, “The Chernoff lower bound for symmetric quantum hypothesis testing,” *Annals of Statistics*, vol. 37, no. 2, pp. 1040–1057, apr 2009.

[56] H. Nagaoka, “The converse part of the theorem for quantum Hoeffding bound,” `arXiv:quant-ph/0611289`.

[57] R. E. Blahut, *Principles and practice of information theory.* Addison-Wesley, 1987.

[58] V. Y. F. Tan, “Asymptotic estimates in information theory with non-vanishing error probabilities,” *Foundations and Trends® in Communications and Information Theory*, vol. 10, no. 4, pp. 1–184, 2014.

[59] Y. Altuğ and A. B. Wagner, “The third-order term in the normal approximation for singular channels,” in *2014 IEEE International Symposium on Information Theory.* Institute of Electrical and Electronics Engineers (IEEE), jun 2014.

[60] Y. Polyanskiy, H. V. Poor, and S. Verdu, “Channel coding rate in the finite blocklength regime,” *IEEE Trans. Inform. Theory*, vol. 56, no. 5, pp. 2307–2359, may 2010.

[61] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, “Discriminating states: The quantum Chernoff bound,” *Physical Review Letters*, vol. 98, p. 160501, apr 2007.

[62] M. Tomamichel, *Quantum Information Processing with Finite Resources.* Springer International Publishing, 2016.

[63] C. Berge, *Topological Spaces.* Oliver & Boyd, 1963.

[64] B. Pshenichnyi, *Necessary Conditions for an Extremum Pshenichnyi.* CRC Press, 1971.

[65] R. T. Rockafellar, “Minimax theorems and conjugate saddle-functions.” *Mathematica Scandinavica*, vol. 14, p. 151, jun 1964.

[66] ——, *Convex Analysis.* Walter de Gruyter GmbH, jan 1970.

[67] S. Weis, “Quantum convex support,” *Linear Algebra and its Applications*, vol. 435, no. 12, pp. 3168–3188, dec 2011.

[68] F. Hiai and D. Petz, *Introduction to Matrix Analysis and Applications.* Springer International Publishing, 2014.

[69] H.-C. Cheng and M.-H. Hsieh, "New characterizations of matrix Φ-entropies, Poincaré and Sobolev inequalities and an upper bound to Holevo quantity," `arXiv:1506.06801 [quant-ph]`.

[70] H.-C. Cheng, M.-H. Hsieh, and Tomamichel, "Exponential decay of matrix Φ-entropies on Markov semigroups with applications to dynamical evolutions of quantum ensembles," `arXiv:1511.02627 [quant-ph]`.

[71] H.-C. Cheng and M.-H. Hsieh, "Characterizations of matrix and operator-valued Φ-entropies, and operator EfronStein inequalities," *Proceedings of the Royal Society of London A*, p. 20150563, 2016.

[72] R. M. Dudley, *Real Analysis and Probability.* Cambridge University Press (CUP), 2002.

[73] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of Convex Analysis.* Springer Nature, 2001.

[74] X. Wang, W. Xie, and R. Duan, "Semidefinite programming strong converse bounds for classical capacity," `arXiv:1610.06381 [quant-ph]`.

Verifiable fault-tolerance in measurement-based quantum computation

Keisuke Fujii[1] [2] [*] Masahito Hayashi[3] [4] [†]

[1] *Photon Science Center, Graduate School of Engineering, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-8656, Japan*

[2] *JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan*

[3] *Graduate School of Mathematics, Nagoya University, Nagoya, 464-8602, Japan*

[4] *Centre for Quantum Technologies, National University of Singapore, 117543, Singapore*

Abstract. Quantum systems, in general, cannot be simulated efficiently by a classical computer, and hence is useful for solving certain mathematical problems and simulating quantum many-body systems. This also implies, unfortunately, that verification of the output of the quantum systems is not so trivial, since predicting the output is exponentially hard. Here we propose a framework for verification of the output of fault-tolerant quantum computation in the measurement-based model. Contrast to existing analyses on fault-tolerance, we do not assume any noise model on the resource state, but an arbitrary resource state is tested by using only single-qubit measurements to verify whether the output of measurement-based quantum computation on it is correct or not. Verifiability is equipped by a constant time repetition of the original meausrement-based quantum computation in appropriate meausrement bases. Since full characterization of quantum noise is exponentially hard for large-scale quantum computing systems, our framework provides an efficient way of practical verification of experimental quantum error correction. (The long version of the present work is located on arXiv [1].)

Keywords: verification, fault-tolerant quantum computation, measurement-based quantum computation, blind quantum computation

1 Introduction

Quantum computation provides a new paradigm of information processing offering both fast and secure information processing, which could not be realized in classical computation [2]. Recently, a lot of experimental efforts have been paid to realize quantum computation [3, 4, 5]. There, fault-tolerant quantum computation with quantum error correction [2, 6] is inevitable to obtain quantum advantage using noisy quantum devices.

Due to the recent rapid progresses on experimental quantum error correction techniques [7, 8, 9, 10], there is an increasing demand on an efficient way of a performance analysis of fault-tolerant quantum computation. There are three categories for this purpose, characterization, validation and verification of quantum systems (QCVV) [11]. In the majority of existing performance analyses of fault-tolerant quantum computation, a specific noise model, such as independent and identical Pauli error operation and some specific correlation models, is assumed apriori [12, 13, 14, 15, 16, 17, 19, 18, 20]. By characterizing the elementary quantum operations experimentally, these could serve as validation of quantum computing devices [21]. However, in actual experiments, more general noise might occur including general trace preserving completely positive (TP-CP) maps with various correlation between qubits [22, 23]. Since full tomographic approch does not work efficiently, we need a novel scheme for the third category, verification, to guarantee correctness of the output of a quantum computer without assuming the underlying noise model. Unfortunately, existing fault-tolerant quantum computations have not equipped such an efficient verification scheme yet.

2 Verifiable fault-tolerance

The aim of this work is to develop fault-tolerant quantum computation being equipped with a verification scheme without assuming the underlying noise model. As requirements of verifiable fault-tolerance, we define the following two concepts. One is *detectability* which means that if the error of a quantum computer is not correctable, such a faulty output of the quantum computation is detected with high probability. In this stage, any assumption on the underlying noise model should not be made. The other is *acceptability* which means that an appropriately constructed quantum computer can pass the verification with high probability. In other words, under a realistic noise model, the test accepts the quantum computation with high probability. Both properties are important to characterize performance of test in statistical hypothesis testing [24].

3 Our main contribution

In this work, we develop verifiable fault-tolerance in measurement-based quantum computation (MBQC) [25, 26], which satisfies both detectability and acceptability (see Ref [1] for the detail). We take a rather different approach to fault-tolerance than conventional one. We do not assume any noise model underlying, but define a correctable set of errors on a resource state of MBQC and test whether the error on a given resource state belongs to such a set or not. To this end, we employ the stabilizer test proposed in Ref. [27], where an efficient verification of MBQC can be carried out by testing the graph state. However, this method is not fault-tolerant lacking acceptability; any small amount of noise on the

[*] fujii@qi.t.u-tokyo.ac.jp

[†] masahito@math.nagoya-u.ac.jp

graph state causes rejection regardless whether or not it is correctable. Although the paper [28] extended the stabilizer test to the self-testing for the measurement basis, it still has the same problem. Therefore, we crucially extend the stabilizer test [27] for a noisy situation, so that we can decide whether the given resource states belong to a set of fault-tolerant resource states or not (See *"Test for veirification of fault-tolerance"* in Ref [1]). In Theorem 1 of Ref. [1], we show under the condition of a successful pass of the test, that the accuracy of fault-tolerant MBQC is guaranteed to be arbitrarily high (i.e., contraposition of detectability). Our verification scheme works quite efficiently by simply repeating fault-tolerant MBQC without verification for a constant time in appropriate measurement bases. Therefore, we do not need any special resource state nor entangling operation for verification. The total overhead is only factored by a constant to the original fault-tolerant MBQC. In order to demonstrate acceptability, we consider a concrete example, and explicitly define a set of correctable errors on the resource state for topologically protected MBQC [14, 16, 6] (see *"Verifiable fault-tolerance for topological ly protected MBQC.—"* and Appendix A and B in Ref. [1] for the detail). Under a realistic noise model, we calculate a lower bound of the acceptance probability concretely and show that it can be made close to one (see *"Acceptance probability under a typical error model.—"* and Appendix C in Ref. [1]).

4 Verifiable blind quantum computation

We also address an application of the proposed verification scheme in a different context, blind quantum computation [29, 30, 33, 32, 34, 31, 35]. A promising application of the proposed framework is verification of measurement-only blind quantum computation [33]. Suppose a quantum server generates two-colorable graph states and sends them to a client who execute universal quantum computation by only single-qubit measurements, where client employ the proposed verification. First, our protocol is a one-way quantum communication from Bob to Alice, and therefore, the blindness is guaranteed by the no-signaling principle as in the protocol of Ref. [33], which contrasts to verifiable blind quantum computation [30, 35] of BFK (Broadbent-Fitzsimons-Kashefi) type [29]. According to detectability (Theorem 1 of Ref. [1]), under the condition of acceptance, the accuracy of the output is guaranteed. Contrast to the earlier verifiable blind quantum computation [30, 27], by virtue of acceptability, the proposed verification scheme can accept the delegated quantum computation even under quantum server's deviation or quantum channel noise as long as they are correctable. In this way, we can verify the quantum server is honest enough to obtain a correct output by only using single-qubit measurements. While fault-tolerance of verifiable blind quantum computation has been an open problem in the field [36], the proposed verifiable fault-tolerance in the measurement-based model combined with measurement-only blind quantum computation [33] resolved it successfully.

References

[1] K. Fujii and M. Hayashi, arXiv:1610.05216.

[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).

[3] T. D. Ladd *et al.,*, Nature **464**, 45 (2010).

[4] E. Gibney, Nature **516**, 24 (2014).

[5] S. Benjamin and J. Kelly, Nature Mat. **14**, 561 (2015).

[6] K. Fujii, *Quantum Computation with Topological Codes -From Qubit to Topological Fault- Tolerance-*. SpringerBriefs in Mathematical Physics vol. **8** (Springer-Verlag 2015).

[7] R. Barends *et al.*, Nature **508**, 500 (2014).

[8] J. Kelly *et al.*, Nature **519**, 66 (2015).

[9] J. M. Chow *et al.*, Nature Communications **5**, 4015 (2014).

[10] M. Takita *et al.*, arXiv:1605.01351.

[11] J. M. Gambetta, J. M. Chow and M. Steffen, npj Quant. Info. **3**, 2 (2017).

[12] A. M. Steane, Nature (London) **399**, 124 (1999).

[13] E. Knill, Nature (London) **434**, 39 (2005).

[14] R. Raussendorf, J. Harrington, and K. Goyal, Ann. of Phys. **321**, 2242 (2006).

[15] R. Raussendorf, J. Harrington, and K. Goyal, New J. Phys. **9**, 199 (2007).

[16] R. Raussendorf and J. Harrington, Phys. Rev. Lett. **98**, 190504 (2007).

[17] A. G. Fowler, A. M. Stephens, and P Groszkowski, Phys. Rev. A **80**, 052312 (2009).

[18] K. Fujii and K. Yamamoto, Phys. Rev. A **81**, 042324 (2010).

[19] K. Fujii and K. Yamamoto, Phys. Rev. A **82**, 060301(R) (2010).

[20] D. S. Wang, A. G. Fowler, and L. C. L. Hollenberg, Phys. Rev. A **83**, 020302(R) (2011).

[21] R. Kueng, D. M. Long, A. C. Doherty, and S. T. Flammia, Phys. Rev. Lett. **117**, 170502 (2016).

[22] J. Wallman, C. Granade, R. Harper, and S. T. Flammia New J. Phys. **17**, 113020 (2015).

[23] H. Ball, T. M. Stace, S. T. Flammia, and M. J. Biercuk Phys. Rev. A **93**, 022303 (2016).

[24] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses.* Springer Texts in Statistics, Springer (2008).

[25] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).

[26] R. Raussendorf, D. E. Browne, and H. J. Briegel, Phys. Rev. A **68**, 022312 (2003).

[27] M. Hayashi and T. Morimae, Phys. Rev. Lett., **115**, 220502 (2015).

[28] M. Hayashi and M. Hajdušek, arXiv:1603.02195.

[29] A. Broadbent, J. Fitzsimons, and E. Kashefi, in *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science* (IEEE Computer Society, Los Alamitos, USA, 2009), p. 517.

[30] J. F. Fitzsimons and E. Kashefi, arXiv:1203.5217.

[31] S. Barz *et al.*, Science **335**, 303 (2012).

[32] T. Morimae and K. Fujii, Nat. Commun. **3**, 1036 (2012).

[33] T. Morimae and K. Fujii, Blind quantum computation for Alice who does only measurements. Phys. Rev. A **87**, 050301(R) (2013).

[34] T. Morimae and K. Fujii, Phys. Rev. Lett. **111**, 020502 (2013).

[35] Y. Takeuchi, K. Fujii, T. Morimae, and N. Imoto, arXiv:1607.01568.

[36] D. Aharonov, M. Ben-Or, E. Eban, and U. Mahadev, arXiv:1704.04487.

Bayesian Quantum Noise Spectroscopy

Chris Ferrie[1], Chris Granade[2 3], Gerardo Paz-Silva[4 5], and Howard Wiseman[4 5]

[1] University of Technology Sydney, Centre for Quantum Software and Information, Ultimo NSW 2007, Australia

[2] Centre for Engineered Quantum Systems, University of Sydney, Sydney, NSW, Australia

[3] School of Physics, University of Sydney, Sydney, NSW, Australia

[4] Centre for Quantum Dynamics, Griffith University, Brisbane, Queensland 4111, Australia

[5] Centre for Quantum Computation and Communication Technology, Griffith University, Brisbane, Queensland 4111, Australia

July 24, 2017

This short abstract is based on Ref. [1]. As commonly understood, the noise spectroscopy problem is ill-posed. Ad-hoc solutions assume implicit structure which is often never determined. Thus it is unclear when the method will succeed or whether one should trust the solution obtained. Here we propose to treat the problem from the point of view of statistical estimation theory. We develop a Bayesian solution to the problem which allows one to easily incorporate assumptions which render the problem solvable. We compare several numerical techniques for noise spectroscopy and find the Bayesian approach to be superior in many respects.

The development of quantum technologies requires accurate characterisation not only for validation but also for control. *Quantum noise spectroscopy* protocols of varying generality have been developed and implemented in recent years [2–10] as a protocol to probe some aspects of the characterisation problem. Their objective is to characterize the actual noise affecting a quantum system of interest, regardless of its source, in terms of its correlations, or more specifically the set of power poly-spectra [11]. The key point is that the information these protocols output should be enough to enable its use, in tandem with optimal control techniques, to design control routines tailored to suppress the actual noise affecting the quantum system of interest [12, 13]. Operationally, spectroscopy protocols measure the response of a quantum system, in terms of expectation values of observables, in a known initial state, to the noise affecting it and user-determined control routines. The main difficulty is that noise correlations influence the dynamics of the quantum system in a highly non-linear way. Thus, inferring these correlations in detail from the response of the quantum system is generally an ill-posed problem, unless constraints are imposed or, equivalently, if a priori information on the noise is assumed. Even when standard assumptions such as Gaussian noise or a dephasing coupling are satisfied, the problem remains non-linear and inverting it carries along a set of non-trivial complications that in turn constraint the type of noise that can be characterized. For example, in Refs. [5–9] a control induced frequency comb approach is used in order to overcome the non-linear character of the problem but it comes at the cost of being only effective when the noise correlations are *smooth* functions in frequency space.

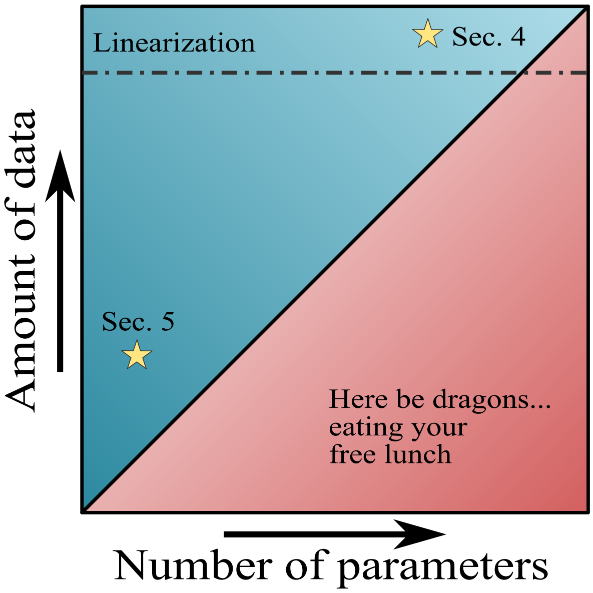


Figure 1: Estimation demands data. At some point (the exact location of which depends are far too many factors to quantify), the distribution of data becomes well-approximated by a Gaussian, allowing an effective linearization of the problem. This greatly simplifies the calculations required to solve the estimation problem. When this is not the case, the problem demands more resources and more clever numerical algorithms to approximate the solution. In any case, the more parameters one has in their model, the more data is require to learn anything. (The section numbers refer to the companion paper [1].)

We propose that many of these problems can be alleviated, or at least properly quantified, using a statistically principled approach. Within the statistical phrasing of the problem we provide a Bayesian solution [14], complete with a numerical implementation. We show the problem can be solved analytically with no numerical approximations in the limit of large of amounts of experimental data. At the other extreme—the small-data limit—a numerically stable Monte Carlo algorithm [15] approximates the full Bayesian solution. Our two approaches provide a robust solution to the software side of the noise spectroscopy problem. These two regimes are schematically depicted in Figure 1. For brevity, only the former will be described in this abstract.

We briefly describe the essence of the spectroscopy protocol. Consult the companion technical paper for full details of the physical model [1]. First, a +1 eigenstate of σ_x is prepared on the system at time $t = 0$, in such way that the expectation value of the observable σ_x at the final time $t = T$, given $H_{\mathrm{ctrl}}(t)$, is determined by

$$\langle \sigma_x \rangle = e^{-\int_0^\infty \frac{d\omega}{2\pi} F(\omega) S(\omega)}, \tag{1}$$

where $F(\omega)$ is known as the *filter function*. Different choices of $H_{\mathrm{ctrl}}(t)$ result in different filters $F(\omega)$, and different experimentally accessible values of $\langle\sigma_x\rangle$. In principle, it should be possible to choose a sufficiently large set of different control sequences in such way that the integral in the exponent can be deconvolved, and information about $S(\omega)$ can be inferred. Different approaches to this problem, under different simplifying assumptions, have been proposed and even experimentally implemented [7, 8, 10, 16, 17].

In the physical description we often make reference to observations as being the average values of observables. By contrast, in real experiments, observations are made by acquiring single bits of data at a time through projective measurements of single quantum systems. These two views of experimental observations agree only in the limit that very large numbers of projective measurements are made on identical copies of the system. Reasoning about noise spectroscopy in the presence of experimental constraints is thus, at its core, a statistical problem not suited to the "data-fitting" paradigm we are more used to. To make this precise we first extract the core mathematical elements of the problem. Mathematically, we are interested in

$$\chi(S; F_j) = \frac{1}{2\pi} \int_0^\infty S(\omega) F_j(\omega) d\omega, \tag{2}$$

where F_j indexes many different control sequences which result in different filter functions. In our simulations and algorithms we take the simple approach of numerically integrating this as

$$\chi(S; F_j) \approx \frac{1}{4\pi} \sum_k F_j(\omega_k) S(\omega_k)(\omega_k - \omega_{k-1}). \tag{3}$$

Recall that we do not have direct access to χ as it is only exposed experimentally through the statistical model in (1). Moreover, expectation values of observables also cannot be measured directly and will always come with fluctuations due to finite sample sizes. Thus, we prefer to work from the bottom up, considering the precise distribution of each bit of data. To this end, let r be a the binary random variable with distribution

$$\Pr(r = 1|S; F_j) = \frac{1}{2}\left(1 + e^{-\chi(S;F_j)}\right), \tag{4}$$

such that the expectation value in (1) obeys

$$\langle \sigma_x \rangle = \Pr(r = 1|S; F_j) - \Pr(r = -1|S; F_j).$$

This is the most fundamental statistical model and we should process data at this level whenever possible. But wait, what does it mean to *process data*? This is where Bayes come in.

The notation $\Pr(A|B)$ is read "the probability of A being true given B is known to be true". So, $\Pr(r = 1|S; F_j)$ is the probability of observing $r = 1$ given the filter F_j is used and the spectrum is S. Ah, but that seems a bit awkward, doesn't it? Isn't the spectrum the thing we don't know? To rectify this, we *invert* the probability using Bayes' rule:

$$\Pr(S|r; F_j) = \frac{\Pr(r|S; F_j)\Pr(S|F_j)}{\Pr(r|F_j)}. \tag{5}$$

Some terminology: $\Pr(r|S; F_j)$ is called the *likelihood function* and in physics it is always given by the physical model; $\Pr(r|F_j)$ is called the *evidence* and it is usually ignored as it can be determined by normalization; $\Pr(S|F_j)$ is called the *prior* and encodes the information we have about the spectrum before the data is take; and finally, $\Pr(S|r; F_j)$ is called the *posterior*, which is the information we have about the spectrum *after* the experiment—exactly what we want to know!

In general, performing this inversion is both analytically and computationally intractable. There are two general approaches to solving this problem. Either we make analytical approximations or we employ clever numerical integration techniques. Here we demonstrate both. But, the problem and solutions are also not decoupled from how much can be assumed known about the spectrum—the *dimension* of model—and the amount of data available, such that the domain of applicability of each solution is restricted in subtle ways. This is shown pictorially in Figure 1.

In the large data we effectively linearize the model and use what is known as *Gaussian process* (GP) regression [18]. this $S(\omega) \sim \mathcal{GP}(\mu(\omega), k(\omega, \omega'))$, where μ is the *mean function* and k is the covariance function, or *kernel*. In standard notation,

$$\mu(\omega) = \mathbb{E}_S[S(\omega)] \text{ and} \tag{6}$$

$$k(\omega, \omega') = \mathbb{E}_S[(S(\omega) - \mu(\omega))(S(\omega') - \mu(\omega'))]. \tag{7}$$

In principle we can choose any functions μ and k as our mean and kernel functions. However, there are natural choices and ones that have been found to perform well in a broad range of problems. The most common kernel is the so-called squared exponential

$$k(\omega, \omega') = \kappa e^{-\frac{(\omega-\omega')^2}{\delta}}, \tag{8}$$

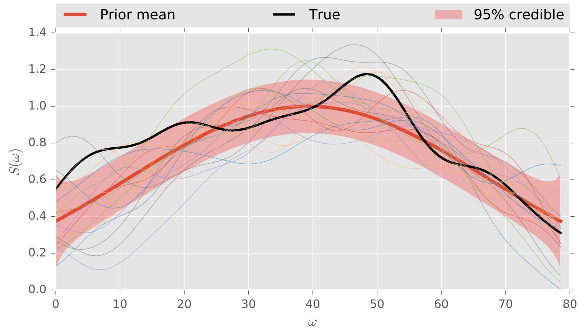


Figure 2: A visualization of a Gaussian Process. Here the mean function μ is taken to be a Gaussian function and we use the squared exponential kernel in (8) with parameters $\kappa = 0.02$ and $\delta = 100$. In red, the mean and 95% credible band is plotted. The other curves are samples from this GP. One of them, in solid black, we take to be the true spectrum.

where δ is a hyper-parameter which controls the correlation in S for nearby ω and κ controls the overall prior uncertainty. We have plotted a visualization of the GP we will use in Figure 2.

If we begin with a GP prior and the distribution of data is also Gaussian, then the posterior is Gaussian and we can derive an analytic expression for its mean function and kernel. Denote $\boldsymbol{G}$ as the matrix with entries $G_{kj} = F_j(\omega_k)(\omega_k - \omega_{k-1})/4\pi$. Then, Bayesian updating amounts to updating the covariance and mean as follows [18]:

$$\boldsymbol{k} \mapsto \boldsymbol{k}' = \boldsymbol{G}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{G} + \boldsymbol{k}^{-1}, \tag{9}$$

$$\boldsymbol{\mu} \mapsto \boldsymbol{k}'^{-1}\left(\boldsymbol{\chi}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{G} + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{k}^{-1}\right), \tag{10}$$

where the bold font simply means the vector of values defined by the level of discretization.

In describing one of our numerical results, we compare the GP estimator with a naive estimator. As a point of reference, we can also treat the prior mean function as an estimator and calculate its loss. The result of 400 trials is shown in Figure 3. As expected, the posterior loss is lower than the prior loss, indicating that the algorithm is learning. The naive loss does a respectable job as well, but is convincingly beaten by the GP estimator—especially given the fact that GP estimator comes with all the added benefits of the Bayesian methodology discussed above.

In this work, we formulated the noise spectroscopy problem in the language of statistical estimation theory. This allows us to provide a robust and principled solution to the problem using Bayesian analysis.

References

[1] C. Ferrie, C. Granade, G. A. Paz-Silva, and H. M. Wiseman, "Bayesian quantum noise spectroscopy," (2017), arXiv:1707.05088 .

[2] F. Yan, J. Bylander, S. Gustavsson, F. Yoshihara, K. Harrabi, D. G. Cory, T. P. Orlando, Y. Nakamura, J.-S. Tsai, and W. D. Oliver, Physical Review B **85**, 174521 (2012).

[3] F. Yan, S. Gustavsson, J. Bylander, X. Jin, F. Yoshihara, D. G. Cory, Y. Nakamura, T. P. Orlando, and W. D. Oliver, Nature Communications **4**, 2337 (2013).

[4] O. E. Dial, M. D. Shulman, S. P. Harvey, H. Bluhm, V. Umansky, and A. Yacoby, Physical Review Letters **110**, 146804 (2013).

[5] T. Yuge, S. Sasaki, and Y. Hirayama, Physical Review Letters **107**, 170504 (2011).

[6] K. C. Young and K. B. Whaley, Physical Review A **86**, 012314 (2012).

[7] G. A. Álvarez and D. Suter, Physical Review Letters **107**, 230501 (2011).

[8] L. M. Norris, G. A. Paz-Silva, and L. Viola, Phys. Rev. Lett. **116**, 150503 (2016).

[9] G. A. Paz-Silva, L. M. Norris, and L. Viola, Physical Review A **95**, 022121 (2017), arXiv:1609.01792 .

[10] J. T. Muhonen, J. P. Dehollain, A. Laucht, F. E. Hudson, R. Kalra, T. Sekiguchi, K. M. Itoh, D. N. Jamieson, J. C. McCallum, A. S. Dzurak, and A. Morello, Nat Nano **9**, 986 (2014).

[11] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002).

[12] M. D. Shulman, S. P. Harvey, J. M. Nichol, S. D. Bartlett, A. C. Doherty, V. Umansky, and A. Yacoby, Nature Communications **5**, 5156 (2014).

[13] Y. Wang, M. Um, J. Zhang, S. An, M. Lyu, J. N. Zhang, L.-M. Duan, D. Yum, and K. Kim, arXiv:1701.04195 (2017).

[14] G. L. Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation* (Springer New York, 1988).

[15] A. Doucet and A. M. Johansen, *A Tutorial on Particle Filtering and Smoothing: Fifteen Years Later* (2011).

[16] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. G. Cory, Y. Nakamura, J.-S. Tsai, and W. D. Oliver, Nature Physics **7**, 565 (2011).

[17] Y. Wang, M. Um, J. Zhang, S. An, and M. Lyu, arXiv , 1701.04195.

[18] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)* (The MIT Press, 2005).

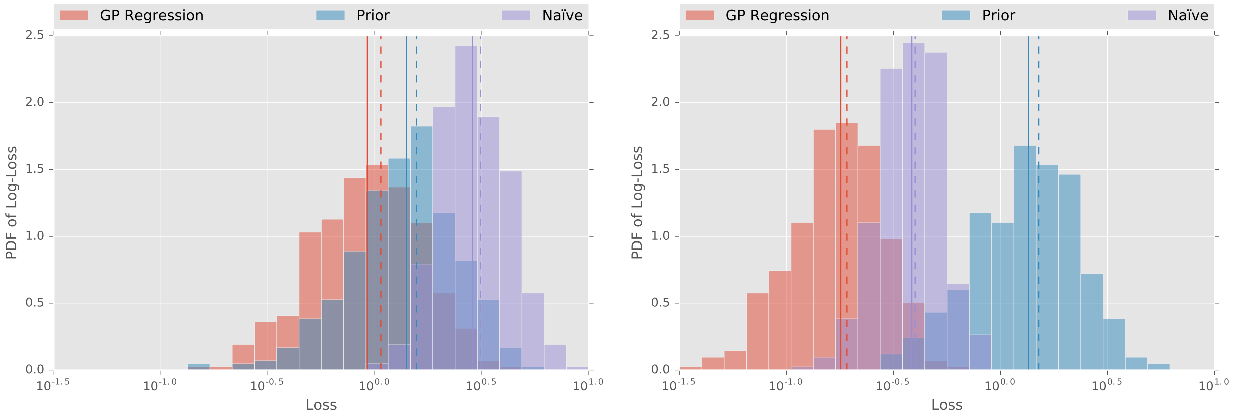


Figure 3: The performance of the Gaussian process estimator and naive estimator in relation to the prior loss. Plotted is a normalized histogram of the log-loss over 400 trials. The simulated experiment is that of $N = 100$ (Left) and $N = 1000$ (Right) single-shot repetitions of each of the 25 control sequences described in the text. Both the median loss (solid line) and the mean loss / Bayes risk (dashed line) are shown to guide the eye. The prior for the Gaussian process estimator is taken to be that shown in Figure 2, and the true spectra are sampled from the prior.

Moderate Deviations for Classical-Quantum Channels

Hao-Chung Cheng[1 2 *] Min-Hsiu Hsieh[2 †]

[1] *Graduate Institute Communication Engineering, National Taiwan University, Taiwan (R.O.C.)*
[2] *Centre for Quantum Computation & Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology Sydney, NSW 2007, Australia*

Abstract. We show that the reliable communication through a classical-quantum channel is possible when the transmission rate approaches the channel capacity sufficiently slowly. This scenario exists between the non-vanishing error probability regime, where the rate tends to capacity with a fixed error, and the small error probability regime, where the error vanishes given a rate below capacity. The proof employs a sharp concentration bound in strong large deviation theory, and the asymptotic expansions of the error-exponent functions. Our complete paper can be found in arXiv:1701.03195 [quant-ph].

1 Introduction

The interplay between the transmission rate, blocklength and error probability is one of the core problems in Shannon theory. Based on different ranges of the error probability, its study roughly falls into the following three categories: (i) *large error probability* or *non-vanishing error probability* regime; (ii) *medium error probability* regime; and (iii) *small probability error* regime. In the non-vanishing error probability regime, the target is to find the largest rate of a code given a blocklength n up to an error ϵ. Strassen [2] applied the *central limit theorem (CLT)* to show that the maximum message size of an n-blocklength code through a discrete memoryless channel (DMC) W yields an asymptotic expansion to the order $\sqrt{n}$, and hence this is called *second-order analysis* [3]:

$$\log M^*(W^n, \epsilon) = nC + \sqrt{nV}\,\Phi^{-1}(\epsilon) + O(\log n), \tag{1}$$

where the quantities C and V denote the capacity and the dispersion [3] of the channel, and Φ is the cumulative distribution function of a standard normal random variable. Equivalently, the second-order result in Eq. (1) can be formulated for estimating the optimal decoding error with blocklength n and rate $C - A/\sqrt{n}$ for any constant A:

$$\lim_{n\to+\infty} \epsilon^*\left(n, C - A/\sqrt{n}\right) = \Phi\left(\frac{A}{\sqrt{V}}\right). \tag{2}$$

In the *small error probability* regime, Shannon [4] introduced a *reliability function* $E(R)$ as the optimal exponent of the exponential decreases of the error for any rate R below C:

$$\epsilon^*(n, R) = \mathrm{e}^{-nE(R)+o(n)}. \tag{3}$$

This seminal work entails the *error exponent analysis* of a broad class of channels [5, 6, 8]. The exponential decay of the error in Eq. (3) is the consequence of the *large deviation principle (LDP)* [10]. Hence, the errors in Eqs. (2) and (3), respectively, fall into the CLT regime and LDP regime.

Altuğ and Wagner [11, 12] pioneered the study of the medium error probability regime, and investigated the asymptotic behaviour of the optimal decoding error when the coding rate converges to capacity slowly. Specifically, they studied the conditions under which the error is asymptotically equal to[1]

$$\epsilon^*(n, C - a_n) \sim \Phi\left(\frac{\sqrt{n}a_n}{\sqrt{V}}\right) \sim \mathrm{e}^{\frac{-na_n^2}{2V}}, \tag{4}$$

where the sequence $(a_n)_{n\in\mathbb{N}}$ satisfies

$$\text{(i)} \lim_{n\to+\infty} a_n = 0, \quad \text{and} \quad \text{(ii)} \lim_{n\to+\infty} na_n^2 = +\infty. \tag{5}$$

A DMC with errors satisfying Eq. (4) possesses a *moderate deviation property (MDP)* [10]. These three approaches—(i), (ii), and (iii)—all have theoretical significance and practical value; however, this paper focuses on the medium error probability regime, which is rarely explored in the quantum scenario.

Our main contribution is, for any classical-quantum (c-q) channel with a non-zero dispersion $V > 0$,

$$\lim_{n\to+\infty} \frac{\log \epsilon^*(n, C - a_n)}{na_n^2} = -\frac{1}{2V}, \tag{6}$$

where $(a_n)_n$ is an arbitrary sequence satisfies Eq. (5). The result in Eq. (6) shows that reliable communication for a c-q channel is possible when the transmission rate approaches capacity at the scale of $\Theta(n^{-t})$, $t \in (0, \frac{1}{2})$. Our proof employs techniques from error exponent analysis. For the achievability part, we start from Hayashi's upper bound of the average error for c-q channels [13] followed by an asymptotic expansion of the error exponent. For the converse, we exploit a refined sphere-packing bound [7, 8]. We remark that Altuğ and Wagner's converse proof [12, Theorem 2.2] cannot be directly generalized to c-q channels because their sphere-packing bound is of a weaker form [9, 8] and hence naively following their converse approach will result in a gap between

[*] F99942118@ntu.edu.tw
[†] Min-Hsiu.Hsieh@uts.edu.au

[1] We denote $f_n \sim g_n$ if and only if $\lim_{n\to+\infty} \frac{f_n}{g_n} = 1$.

the achievability and converse results. Different from the approaches in this work, we remark that a recent and independent paper [1] proceeds from another way of the non-vanishing error regime and also accomplish a MDP result in Eq. (6).

2 Notation and Main Result

2.1 Notation

Throughout this paper, we consider a finite-dimensional Hilbert space $\mathcal{H}$. Denote by $\mathcal{X}$ a finite input alphabet, and let $\mathscr{P}(\mathcal{X})$ be the set of probability distributions on $\mathcal{X}$. A c-q channel $\mathscr{W}$ maps elements of $\mathcal{X}$ to the density operators in $\mathcal{S}(\mathcal{H})$, i.e. $\mathscr{W} : x \mapsto W_x$. Let $\mathcal{M}$ be a finite alphabetical set with size $M = |\mathcal{M}|$. An (n-block) *encoder* is a map $f_n : \mathcal{M} \to \mathcal{X}^n$ that encodes each message $m \in \mathcal{M}$ to a codeword $\mathbf{x}^n(m) := x_1(m)\dots x_n(m) \in \mathcal{X}^n$. The codeword $\mathbf{x}^n(m)$ is then mapped to a state $W^{\otimes n}_{\mathbf{x}^n(m)} = W_{x_1(m)} \otimes \cdots \otimes W_{x_n(m)} \in \mathcal{S}(\mathcal{H}^{\otimes n})$. The *decoder* is described by a positive operator-valued measurement (POVM) $\Pi_n = \{\Pi_{n,1}, \dots, \Pi_{n,M}\}$ on $\mathcal{H}^{\otimes n}$, where $\Pi_{n,i} \geq 0$ and $\sum_{i=1}^{M} \Pi_{n,i} = \mathbb{1}_{\mathcal{H}}$. The pair $(f_n, \Pi_n) =: \mathcal{C}_n$ is called a *code* with *rate* $R = \frac{1}{n}\log|\mathcal{M}|$. The error probability of sending a message m with the code $\mathcal{C}_n$ is $\epsilon_m(\mathcal{C}_n) := 1 - \operatorname{Tr}\left(\Pi_{n,m} W_{\mathbf{x}^n(m)}\right)$. We use $\bar{\epsilon}(\mathcal{C}_n) = \frac{1}{M}\sum_{m\in\mathcal{M}} \epsilon_m(\mathcal{C}_n)$ to denote the *average* error probability. Denote the relative entropy and relative entropy variance by $D(\rho\|\sigma) := \operatorname{Tr}\left[\rho\left(\log\rho - \log\sigma\right)\right]$ and $V(\rho\|\sigma) := \operatorname{Tr}\left[\rho\left(\log\rho - \log\sigma\right)^2\right] - D(\rho\|\sigma)^2$, respectively. We define the mutual information by $I(P, \mathscr{W}) := D\left(P \circ \mathscr{W} \| P \otimes P\mathscr{W}\right)$, where $P\circ\mathscr{W} := \sum_{x\in\mathcal{X}} P(x)|x\rangle\langle x| \otimes W_x$. Hence, the (classical) *information capacity* of the channel is $C := \max_{P\in\mathscr{P}(\mathcal{X})} I(P, \mathscr{W})$. The information variance is defined by$V(P, \mathscr{W}) := V\left(\mathscr{W}\|P\mathscr{W}|P\right)$, where $P\mathscr{W} := \sum_x P(x)W_x$. Further, we define $V := \min_{P\in\mathscr{P}(\mathcal{X}):\, I(P,\mathscr{W})=C} V(P, \mathscr{W})$.

2.2 Main Results

Theorem 1 (Achievability). *For any c-q channel with $V > 0$ and any sequence $(a_n)_{n\geq 1}$ satisfying Eq. (5), there exists a sequence of codes $\{\mathcal{C}_n\}_{n\geq 1}$ with rates $R_n = C - a_n$ so that*

$$\limsup_{n\to+\infty} \frac{1}{na_n^2} \log \bar{\epsilon}\left(\mathcal{C}_n\right) \leq -\frac{1}{2V}. \tag{7}$$

Theorem 2 (Converse). *For any c-q channel with $V > 0$, any sequence $\{a_n\}_{n\geq 1}$ satisfying Eq. (5), and any sequence of codes $\{\mathcal{C}_n\}_{n\geq 1}$ with rates $R_n = C - a_n$, it holds that*

$$\liminf_{n\to+\infty} \frac{1}{na_n^2} \log \bar{\epsilon}\left(\mathcal{C}_n\right) \geq -\frac{1}{2V}. \tag{8}$$

The proof can be found in Section 3 of the main text.

3 Auxiliary functions and their properties.

In order to prove Theorems 1 and 2, we need the following crucial properties of the auxiliary functions, Propositions 3, 4, and 5. The proof can be found in Appendix A of the main text.

The auxiliary function of a classical-quantum channel is defined as [14, 15]

$$E_0(s, P) := -\log \operatorname{Tr}\left[\left(\sum_{x\in\mathcal{X}} P(x) W_x^{1/(1+s)}\right)^{1+s}\right].$$

In this paper, we will require three variants of the above auxiliary function: $\forall s \geq 0$ and $\sigma \in \mathcal{S}(\mathcal{H})$,

$$\widetilde{E}_0(s, P, \sigma) := sD_{1-s}\left(P \circ \mathscr{W} \| P \otimes \sigma\right) \tag{9}$$

$$E_{\mathrm{h}}(s, P, \sigma) := sD_{\frac{1}{1+s}}\left(\mathscr{W}\|\sigma|P\right), \tag{10}$$

$$\widetilde{E}_{\mathrm{h}}(s, P, \sigma) := sD^{\flat}_{\frac{1}{1+s}}\left(\mathscr{W}\|\sigma|P\right), \tag{11}$$

where D_α is the (Petz's) quantum Rényi divergence and $D^{\flat}_\alpha(\rho\|\sigma) := \frac{1}{\alpha}\log\operatorname{Tr}[\exp\{\alpha\log\rho + (1-\alpha)\log\sigma\}]$ is the log-Euclidean Rényi divergence; $D_\alpha\left(\mathscr{W}\|\sigma|P\right) := \sum_{x\in\mathcal{X}} P(x) D_\alpha\left(W_x\|\sigma\right)$.

Proposition 3 (Properties of $\widetilde{E}_0(s, P, \sigma)$).

(a) $\widetilde{E}_0(s, P, \sigma)$ and its partial derivatives $\partial\widetilde{E}_0(s, P, \sigma)/\partial s$, $\partial^2\widetilde{E}_0(s, P, \sigma)/\partial s^2$, $\partial^3\widetilde{E}_0(s, P, \sigma)/\partial s^3$ are all continuous in $(s, P) \in [0, +\infty) \times \mathscr{P}(\mathcal{X})$.

(b) For every $P \in \mathscr{P}(\mathcal{X})$, the function $\widetilde{E}_0(s, P, \sigma)$ is concave in $s \in [0, +\infty)$.

(c) For every $P \in \mathscr{P}(\mathcal{X})$, $\left.\frac{\partial\widetilde{E}_0(s,P,\sigma)}{\partial s}\right|_{s=0} = D\left(P \circ \mathscr{W}\|P \otimes \sigma\right)$.

(d) For every $P \in \mathscr{P}(\mathcal{X})$, $\lim_{s\to+\infty} \frac{\partial\widetilde{E}_0(s,P,\sigma)}{\partial s} \leq \frac{\partial\widetilde{E}_0(s,P,\sigma)}{\partial s} \leq D\left(P \circ \mathscr{W}\|P \otimes \sigma\right)$, $\forall s \in [0, +\infty)$.

(e) For every $P \in \mathscr{P}(\mathcal{X})$, $\left.\frac{\partial^2\widetilde{E}_0(s,P,\sigma)}{\partial s^2}\right|_{s=0} = -V\left(P \circ \mathscr{W}\|P \otimes \sigma\right)$.

Proposition 4 (Properties of $E_{\mathrm{h}}(s, P, \sigma)$).

(a) $E_{\mathrm{h}}(s, P, \sigma)$ and its partial derivatives $\partial E_{\mathrm{h}}(s, P, \sigma)/\partial s$, $\partial^2 E_{\mathrm{h}}(s, P, \sigma)/\partial s^2$, $\partial^3 E_{\mathrm{h}}(s, P, \sigma)/\partial s^3$ are continuous for $(s, P) \in [0, +\infty) \times \mathscr{P}(\mathcal{X})$.

(b) For every $P \in \mathscr{P}(\mathcal{X})$, the function $E_{\mathrm{h}}(s, P, \sigma)$ is concave in s for all $s \in [0, +\infty)$.

(c) For every $P \in \mathscr{P}(\mathcal{X})$, $\left.\frac{\partial E_{\mathrm{h}}(s,P,\sigma)}{\partial s}\right|_{s=0} = D\left(\mathscr{W}\|\sigma|P\right)$.

(d) For every $P \in \mathscr{P}(\mathcal{X})$, $\lim_{s\to+\infty} \frac{\partial E_{\mathrm{h}}(s,P,\sigma)}{\partial s} \leq \frac{\partial E_{\mathrm{h}}(s,P,\sigma)}{\partial s} \leq D\left(\mathscr{W}\|\sigma|P\right)$, $\forall s \in [0, +\infty)$.

(e) For every $P \in \mathscr{P}(\mathcal{X})$, $\left.\frac{\partial^2 E_{\mathrm{h}}(s,P,\sigma)}{\partial s^2}\right|_{s=0} = -V\left(\mathscr{W}\|\sigma|P\right)$, where $V\left(\mathscr{W}\|\sigma|P\right) := \sum_{x\in\mathcal{X}} P(x) V\left(W_x\|\sigma\right)$.

Proposition 5 (Properties of $\widetilde{E}_{\mathrm{h}}(s,P,\sigma)$)**.**

(a) $\widetilde{E}_{\mathrm{h}}(s,P,\sigma)$ *and its partial derivatives* $\partial \widetilde{E}_{\mathrm{h}}(s,P,\sigma)/\partial s$, $\partial^2 \widetilde{E}_{\mathrm{h}}(s,P,\sigma)/\partial s^2$, $\partial^3 \widetilde{E}_{\mathrm{h}}(s,P,\sigma)/\partial s^3$ *are all continuous for* $(s,P)\in[0,+\infty)\times\mathscr{P}(\mathcal{X})$.

(b) For every $P\in\mathscr{P}(\mathcal{X})$*, the function* $\widetilde{E}_{\mathrm{h}}(s,P,\sigma)$ *is concave in* s *for all* $s\in[0,+\infty)$.

(c) For every $P\in\mathscr{P}(\mathcal{X})$, $\left.\frac{\partial \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s}\right|_{s=0} = D\left(\mathscr{W}\|\sigma|P\right)$.

(d) For every $P\in\mathscr{P}(\mathcal{X})$, $\lim_{s\to+\infty}\frac{\partial \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s} \le \frac{\partial \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s} \le D\left(\mathscr{W}\|\sigma|P\right)$, $\forall s\in[0,+\infty)$.

(e) For every $P\in\mathscr{P}(\mathcal{X})$, $\left.\frac{\partial^2 \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s^2}\right|_{s=0} = -\widetilde{V}\left(\mathscr{W}\|\sigma|P\right)$, *where* $\widetilde{V}\left(\mathscr{W}\|\sigma|P\right) := \sum_{x\in\mathcal{X}} P(x)\widetilde{V}\left(W_x\|\sigma\right)$ *and* $\widetilde{V}\left(\rho\|\sigma\right) := \int_0^1 \mathrm{d}t\,\mathrm{Tr}\left[\rho^{1-t}(\log\rho-\log\sigma)\rho^t(\log\rho-\log\sigma)\right] - D(\rho\|\sigma)^2$.

4 Discussions

We consider a scenario that involves the interplay between three parameters—optimal error probability, the transmission rate, and the coding blocklength. Our result shows that the optimal error of a c-q channel with positive channel dispersion tends to zero as the rate approaches channel capacity slower than $\frac{1}{\sqrt{n}}$. Our proof strategy is based on a strong large deviation inequality [8] and the asymptotic behaviour expansion the error exponent function. It is interesting that a recent and independent work [1], which proceeds from the other extreme of the non-vanishing error probability regime, also accomplishes the MDP result.

References

[1] C. T. Chubb, V. Y. F. Tan, and M. Tomamichel, "Moderate deviation analysis for classical communication over quantum channels," arXiv:1701.03114 [quant-ph].

[2] V. Strassen, "Asymptotische abschätzungen in Shannon's informationstheorie," *Trans. of the Third Prague Conference on Inform. Theory*, pp. 689–723, 1962.

[3] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inform. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[4] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell System Technical Journal*, vol. 38, no. 3, pp. 611–656, may 1959.

[5] C. Shannon, R. Gallager, and E. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. I," *Information and Control*, vol. 10, no. 1, pp. 65–103, Jan 1967.

[6] R. Gallager, *Information Theory and Reliable Communication.* Wiley, 1968.

[7] H.-C. Cheng, M.-H. Hsieh and M. Tomamichel, "Sphere-Packing Bound for Symmetric Classical-Quantum Channels," arXiv:1701.02957 [quant-ph].

[8] H.-C. Cheng, M.-H. Hsieh, and M. Tomamichel "Quantum Sphere-Packing Bounds with Polynomial Prefactors," arXiv:1704.05703 [quant-ph].

[9] M. Dalai and A. Winter, "Constant compositions in the sphere packing bound for classical-quantum channels," in *2014 IEEE ISIT*. Jun 2014.

[10] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications.* Springer, 1998.

[11] Y. Altuğ and A. B. Wagner, "Moderate deviation analysis of channel coding: Discrete memoryless case," in *2010 IEEE ISIT*, Jun 2010.

[12] ——, "Moderate deviations in channel coding," *IEEE Trans. Inform. Theory*, vol. 60, no. 8, pp. 4417–4426, 2014.

[13] M. Hayashi, "Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding," *Physical Review A*, vol. 76, no. 6, 2007.

[14] A. Holevo, "Reliability function of general classical-quantum channel," *IEEE Transaction on Information Theory*, vol. 46, no. 6, pp. 2256–2261, 2000.

[15] H.-C. Cheng and M.-H. Hsieh, "Concavity of the auxiliary function for classical-quantum channels," *IEEE Transactions on Information Theory*, vol. 62, no. 10, pp. 5960 – 5965, 2016.

[16] A. Winter, "Coding theorems of quantum information theory," 1999, (PhD Thesis, Universität Bielefeld).

[17] H.-C. Cheng and M.-H. Hsieh, "Moderate deviation analysis for classical-quantum channels and quantum hypothesis testing," arXiv:1701.03195 [quant-ph].

[18] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Cambridge University Press (CUP), 2011.

[19] M. Tomamichel and V. Y. F. Tan, "Second-order asymptotics for the classical capacity of image-additive quantum channels," *Commun. Math. Phys.*, vol. 338, no. 1, pp. 103–137, May 2015.

[20] K. M. R. Audenaert, M. Nussbaum, A. Szkoła, and F. Verstraete, "Asymptotic error rates in quantum hypothesis testing," *Commun. Math. Phys.*, vol. 279, no. 1, pp. 251–283, Feb 2008.

Moderate Deviation Analysis for Classical-Quantum Channels and Quantum Hypothesis Testing

Hao-Chung Cheng[1,2] and Min-Hsiu Hsieh[1]

[1] *Centre for Quantum Software and Information (UTS:Q|SI⟩),*
Faculty of Engineering and Information Technology, University of Technology Sydney, Australia
[2] *Graduate Institute Communication Engineering, National Taiwan University, Taiwan (R.O.C.)*

Abstract. In this work, we study the tradeoffs between the error probabilities of classical-quantum channels and the blocklength n when the transmission rates approach the channel capacity at a rate slower than $1/\sqrt{n}$, a research topic known as moderate deviation analysis. We show that the optimal error probability vanishes under this rate convergence. Our main technical contributions are a tight quantum sphere-packing bound, obtained via Chaganty and Sethuraman's concentration inequality in strong large deviation theory, and asymptotic expansions of error-exponent functions. Moderate deviation analysis for quantum hypothesis testing is also established. The converse directly follows from our channel coding result, while the achievability relies on a martingale inequality.

1. Introduction

Investigating the interplay between the transmission rate, blocklength and error probability is one of the core problems in information theory. Based on different ranges of the error probability, the analysis of communication performance roughly falls into the following three categories: (i) *large error probability* or *non-vanishing error probability* regime; (ii) *medium error probability* regime; and (iii) *small error probability* regime. In the non-vanishing error probability regime, the largest transmission rate, given a coding length n and an error probability no more than ϵ, is one of the main research focuses. Strassen [1] first demonstrated that the maximum size of an n-blocklength code through a discrete memoryless channel (DMC) $\mathcal{W}$, denoted by $M^*(\mathcal{W}^n, \epsilon)$, yields an asymptotic expansion to the order $\sqrt{n}$, and hence this is called *second-order analysis*:

$$\log M^*(W^n, \epsilon) = nC + \sqrt{nV}\, \Phi^{-1}(\epsilon) + O(\log n), \tag{1.1}$$

where the quantities C and V denote the capacity [2] and the dispersion [3] of the channel, and Φ is the cumulative distribution function of a standard normal random variable. Equivalently, Eq. (1.1) yields the following relationship between the optimal decoding error with blocklength n and rate $C - A/\sqrt{n}$ for any constant A:

$$\lim_{n\to+\infty} \epsilon^*\left(n, C - A/\sqrt{n}\right) = \Phi\left(\frac{A}{\sqrt{V}}\right). \tag{1.2}$$

Strassen's result relied on the *Gaussian approximation* or the *central limit theorem (CLT)*. His work was latter refined by Hayashi [4], Polyanskiy *et al.* [3], and extended to quantum channels [5, 6, 7, 8]. The results for higher-order asymptotics are referred to Refs. [9, 10, 11].

In the *small error probability* regime, Shannon [12] introduced the *reliability function* $E(R)$ as the optimal error exponent:

$$\lim_{n\to+\infty} -\frac{1}{n} \log \epsilon^*\left(n, R\right) = E(R), \tag{1.3}$$

E-mail address: F99942118@ntu.edu.tw, Min-Hsiu.Hsieh@uts.edu.au.

for rate R below the channel capacity[1] C. This seminal work entails the *error exponent analysis* of a broad class of channels [14, 13, 15, 16, 17, 18]. The exponential decay of the error probability in Eq. (1.3) is a consequence of the *large deviation principle (LDP)* [19]. In summary, the errors in Eqs. (1.2) and (1.3), respectively, fall into the CLT regime and large-deviation regime.

Altuğ and Wagner [20, 21] pioneered the study of the medium error probability regime, and investigated the asymptotic behaviour of the optimal decoding error when the coding rate converges to capacity sufficiently slowly. Specifically, they studied under which conditions the error is asymptotically equal to[2]

$$\epsilon^*(n, C - a_n) \sim \Phi\left(\frac{\sqrt{n}a_n}{\sqrt{v}}\right) \sim \mathrm{e}^{\frac{-na_n^2}{2v}}, \tag{1.4}$$

where the sequence $(a_n)_{n\in\mathbb{N}}$ satisfies

$$\begin{aligned} &\text{(i)} \lim_{n\to+\infty} a_n = 0; \\ &\text{(ii)} \lim_{n\to+\infty} a_n\sqrt{n} = +\infty. \end{aligned} \tag{1.5}$$

Evidently, the transmission rate in Eq. (1.4) approaches capacity slower than $1/\sqrt{n}$. A DMC with errors satisfying Eq. (1.4) possesses a *moderate deviation property (MDP)* [19, Section 3.7]. The constant v in Eq. (1.4) equals the channel dispersion V when both the limit in Eq. (1.2) and MDP hold [22, Theorem 1]. We refer the interested readers to Refs. [22, 24, 21] for further results in classical channel coding. These three approaches—(i), (ii), and (iii)—all have theoretical significance and practical value, and this paper will focus on the medium error probability regime, which is rarely explored in the quantum scenario.

Our main contribution is, for any classical-quantum (c-q) channel with a non-zero dispersion $V > 0$,

$$\lim_{n\to+\infty} \frac{\log \epsilon^*(n, C - a_n)}{na_n^2} = -\frac{1}{2V}, \tag{1.6}$$

where $(a_n)_{n\in\mathbb{N}}$ is any sequence satisfying Eq. (1.5). The result in Eq. (1.6) shows that reliable communication over a c-q channel is possible when the transmission rate approaches capacity at the scale slower than $1/\sqrt{n}$. Our proof employs techniques from the error exponent analysis (the LDP regime). For the achievability part, we start from Hayashi's upper bound of the average error for c-q channels [27] followed by an asymptotic expansion of the error-exponent function. For the converse, we employ a sharp converse bound based on a strong large deviation inequality (Proposition 7). This bound is more general than the previous result in Ref. [18, Proposition 14], since it allows the transmission rates to depend on the blocklength instead of being fixed. We remark that Altuğ and Wagner's converse proof [21, Theorem 2.2] is not sufficient for proving Eq. (1.6) because their sphere-packing bound is of a weaker form in general c-q channels [18, Theorem 6] (see also [29]). Thus, naively following their converse approach will result in a gap between the achievability and converse results (see Remark 3.1).

As a special case of c-q channel coding, we obtain the moderate deviations for binary quantum hypothesis testing (see Theorems 9 and 10):

$$\lim_{n\to+\infty} \frac{1}{na_n^2} \log \widehat{\alpha}_{\exp\{-n[D(\rho\|\sigma)-a_n]\}}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) = -\frac{1}{2V(\rho\|\sigma)}, \tag{1.7}$$

where $\widehat{\alpha}_\mu$ denotes the smallest type-I error when the type-II error does not exceed μ; $D(\rho\|\sigma)$ and $V(\rho\|\sigma)$ denote the relative entropy and relative variance of ρ and σ, respectively. The converse part directly follows from the channel coding, and we provide two proofs for the achievability part. The first one comes from Audeneart *et al.*'s error exponent analysis [30], while the second one employs a martingale inequality [24]. We remark that the moderate deviation analysis for classical hypothesis testing was studied by Sason [24], and by Watanabe and Hayashi [25]. Moreover, a recent work by Rouzé and Datta [26] formulated the quantum hypothesis problem into a martingale, which is similar to our approach for proving the achievability.

[1] To the best of our knowledge, the reliability function $E(R)$ is only known in the high rate regime, i.e. at rates above a *critical rate* (see e.g. [13, p. 160]).

[2] We denote $f_n \sim g_n$ if and only if $\lim_{n\to+\infty} \frac{f_n}{g_n} = 1$.

Unlike our proof techniques relying on error exponent analysis (the LDP regime), a recent and independent paper [31] obtained the same result, but proceeds from the second-order analysis (the CLT regime). Their achievability proof follows from the one-shot capacity by Wang and Renner [32]; while the converse part generalizes Polyanskiy and Verdú's result [22] (which in turn relies on Strassen's Gaussian approximation [1]) and a powerful inequality in probability [33] to the quantum scenario. We summarize the error behaviors in these three regimes in Table 1.

This paper is organized as follows. We introduce notation and preliminaries in Section 2. Section 3 contains our main result—the moderate deviation analysis for c-q channel coding. In Section 4, we present the moderate deviations for quantum hypothesis testing. Lastly, we conclude this paper in Section 5.

Error Regimes	Concentration Phenomena	Hypothesis Testing	Channel Coding
Large Error	CLT: $\Pr\left(S_n \geq \sqrt{n}x\right) \to 1 - \Phi\left(\frac{x}{\sqrt{v}}\right)$	$\widehat{\alpha}_{\exp\left\{-n\left[D-\frac{A}{\sqrt{n}}\right]\right\}} \to \Phi\left(\frac{A}{\sqrt{V}}\right)$	$\epsilon^*\left(n, C - \frac{A}{\sqrt{n}}\right) \to \Phi\left(\frac{A}{\sqrt{V}}\right)$
Medium Error	MDP: $\Pr\left(S_n \geq na_nx\right) = \mathrm{e}^{-\frac{na_n^2}{2v}x + o(na_n^2)}$	$\widehat{\alpha}_{\exp\{-n[D-a_n]\}} = \mathrm{e}^{-\frac{na_n^2}{2V} + o(na_n^2)}$	$\epsilon^*(n, C - a_n) = \mathrm{e}^{-\frac{na_n^2}{2V} + o(na_n^2)}$
Small Error	LDP: $\Pr\left(S_n \geq nx\right) = \mathrm{e}^{-n\Lambda^*(x) + o(n)}$	$\widehat{\alpha}_{\exp\{-nr\}} = \mathrm{e}^{-n\phi(r) + o(n)}$	$\epsilon^*(n, R) = \mathrm{e}^{-nE(R) + o(n)}$

Table 1. This table compares the asymptotic error behaviors of quantum hypothesis testing and classical-quantum channel coding in three error probability regimes: (i) large error (central limit theorem), (ii) medium error (moderate deviation principle), and (iii) small error (large deviation principle). The quantity S_n denotes the sum of n independent and identically distributed random variables with zero mean and variance v. The exponent Λ^* is the Legendre-Fenchel transform of the normalized cumulant generating function of S_n [19]. The error $\widehat{\alpha}_{\exp\{-nr\}}$ is defined as the minimum type-I error with the type-II error smaller than $\exp\{-nr\}$. The quantities D and V in the hypothesis testing column denote the quantum relative entropy and the relative entropy variance, respectively. The optimal error probability with blocklength n and rate R is denoted by $\epsilon^*(n, R)$. The quantities C and V in the channel coding column indicate the channel capacity and the channel dispersion, respectively. The sequence $(a_n)_{n\in\mathbb{N}}$ satisfies Eq. (1.5). The quantity $E(R)$ is the reliability function of the channel.

2. Preliminaries and Notation

We first introduce necessary notation. Throughout this paper, we consider a Hilbert space $\mathcal{H}$ with finite dimension d. The set of density operators (i.e. positive semi-definite operators with unit trace) and non-singular density operators on $\mathcal{H}$ are defined by $\mathcal{S}(\mathcal{H})$ and $\mathcal{S}_{>0}(\mathcal{H})$, respectively. The identity operator on $\mathcal{H}$ is denoted by $\mathbb{1}_{\mathcal{H}}$, or simply $\mathbb{1}$ if there is no possibility of confusion. We use $\operatorname{Tr}[\,\cdot\,]$ as the trace function. Let $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{R}_{\geq 0}$ denote the set of integers, real numbers, and non-negative real numbers, respectively. Define $[n] := \{1, 2, \ldots, n\}$ for $n \in \mathbb{N}$.

The power of a positive semi-definite operator A is defined as: $A^p = \sum_{i:a_i\neq 0} a_i^p P_i$, where $(a_i)_i$ and $(P_i)_i$ are the eigenvalues and eigenprojections of $A = \sum_i a_i P_i$. We use $\mathtt{supp}(A)$ to denote support of the operator A. We write $A \ll B$ if $\mathtt{supp}(A) \subset \mathtt{supp}(B)$.

2.1. Quantum Hypothesis Testing and Channel Coding.

Consider a binary hypothesis testing problem whose null and alternative hypotheses are $\rho \in \mathcal{S}(\mathcal{H})$ and $\sigma \in \mathcal{S}(\mathcal{H})$, respectively. The *type-I error* and *type-II error* of the hypothesis testing, for an operator $0 \leq Q \leq \mathbb{1}$, are defined as follows:

$$\alpha\left(Q;\rho\right) := \operatorname{Tr}\left[(\mathbb{1} - Q)\rho\right], \tag{2.1}$$

$$\beta\left(Q;\sigma\right) := \operatorname{Tr}\left[Q\sigma\right]. \tag{2.2}$$

There is a trade-off relation between these two errors. Thus we can define the minimum type-I error when the type-II error is below $\mu \in (0, 1)$ as

$$\widehat{\alpha}_{\mu}\left(\rho\|\sigma\right) := \min_{0\leq Q\leq\mathbb{1}} \left\{\alpha\left(Q;\rho\right) : \beta\left(Q;\sigma\right) \leq \mu\right\}. \tag{2.3}$$

Denote by $\mathcal{X}$ a finite input alphabet, and let $\mathcal{P}(\mathcal{X})$ be the set of probability distributions on $\mathcal{X}$. For a sequence $\mathbf{x}^n \in \mathcal{X}^n$, we denote by

$$P_{\mathbf{x}^n}(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\left\{x = x_i\right\}, \tag{2.4}$$

where x_i is the i-th element of $\mathbf{x}^n$.

A c-q channel $\mathcal{W}$ maps elements of $\mathcal{X}$ to the density operators in $\mathcal{S}(\mathcal{H})$, i.e. $\mathcal{W} : x \mapsto W_x$. We denote the image of the channel $\mathcal{W}$ by

$$\mathsf{im}\left(\mathcal{W}\right) := \left\{\rho \in \mathcal{S}(\mathcal{H}) \,|\, \exists\, x \in \mathcal{X} : \rho = W_x\right\}, \tag{2.5}$$

and its closure by $\overline{\mathsf{im}(\mathcal{W})}$. Without loss of generality, we assume that $\mathsf{im}\left(\mathcal{W}\right)$ has full support on the Hilbert space $\mathcal{H}$ throughout this paper.

Let $\mathcal{M}$ be a finite alphabetical set with size $M = |\mathcal{M}|$. An (n-block) *encoder* is a map $f_n : \mathcal{M} \to \mathcal{X}^n$ that encodes each message $m \in \mathcal{M}$ to a codeword $\mathbf{x}^n(m) := x_1(m) \dots x_n(m) \in \mathcal{X}^n$. The c-q channel then produces an output state $W^{\otimes n}_{\mathbf{x}^n(m)}$ with the input codeword $\mathbf{x}^n(m)$, where

$$W^{\otimes n}_{\mathbf{x}^n(m)} = W_{x_1(m)} \otimes \cdots \otimes W_{x_n(m)} \in \mathcal{S}(\mathcal{H}^{\otimes n}). \tag{2.6}$$

The *decoder* is described by a positive operator-valued measurement (POVM) $\Pi_n = \{\Pi_{n,1}, \dots, \Pi_{n,M}\}$ on $\mathcal{H}^{\otimes n}$, where $\Pi_{n,i} \geq 0$ and $\sum_{i=1}^{M} \Pi_{n,i} = \mathbb{1}$. The pair $(f_n, \Pi_n) =: \mathcal{C}_n$ is called a *code* with *rate* $R = \frac{1}{n}\log|\mathcal{M}|$. The error probability of sending a message m with the code $\mathcal{C}_n$ is $\epsilon_m(\mathcal{W}, \mathcal{C}_n) := 1 - \mathrm{Tr}\left(\Pi_{n,m} W_{\mathbf{x}^n(m)}\right)$. We use $\epsilon_{\max}(\mathcal{W}, \mathcal{C}_n) = \max_{m\in\mathcal{M}} \epsilon_m(\mathcal{W}, \mathcal{C}_n)$ and $\bar{\epsilon}(\mathcal{W}, \mathcal{C}_n) = \frac{1}{M}\sum_{m\in\mathcal{M}} \epsilon_m(W, \mathcal{C}_n)$ to denote the *maximal* error probability and the *average* error probability, respectively. Denote by $\epsilon^*\left(n, R\right)$ the smallest average error probability among all codes $\mathcal{C}_n$ with message size $|\mathcal{M}| = \exp\{nR\}$.

2.2. **Information Quantities.** For any $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, we define the quantum relative entropy, (Petz's) quantum Rényi divergence [43], and the log-Euclidean Rényi divergence [48, 18], respectively, as follows:

$$D(\rho\|\sigma) := \mathrm{Tr}\left[\rho\left(\log\rho - \log\sigma\right)\right], \tag{2.7}$$

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\mathrm{Tr}[\rho^\alpha\sigma^{1-\alpha}], \tag{2.8}$$

$$D^\flat_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\mathrm{Tr}\left[\mathrm{e}^{\alpha\log\rho + (1-\alpha)\log\sigma}\right]. \tag{2.9}$$

We define two types of the quantum relative entropy variances [5, 6] by

$$V(\rho\|\sigma) := \mathrm{Tr}\left[\rho\left(\log\rho - \log\sigma\right)^2\right] - D(\rho\|\sigma)^2 \tag{2.10}$$

$$\widetilde{V}(\rho\|\sigma) := \int_0^1 \mathrm{d}t\,\mathrm{Tr}\left[\rho^{1-t}(\log\rho - \log\sigma)\rho^t(\log\rho - \log\sigma)\right] - D(\rho\|\sigma)^2. \tag{2.11}$$

It is well-known that both quantities are non-negative, and

$$V(\rho\|\sigma) > 0 \quad \text{implies} \quad D(\rho\|\sigma) > 0. \tag{2.12}$$

We define the *conditional quantum relative entropy* of two channels $\bar{\mathcal{W}}, \mathcal{W}$ and $P \in \mathcal{P}(\mathcal{X})$ to be

$$D\left(\bar{\mathcal{W}}\|\mathcal{W}|P\right) := \sum_{x\in\mathcal{X}} P(x) D\left(\bar{W}_x\|W_x\right). \tag{2.13}$$

Similarly, we define the following conditional entropic quantities for $\sigma \in \mathcal{S}(\mathcal{H})$ and $P \in \mathcal{P}(\mathcal{X})$:

$$D\left(\mathcal{W}\|\sigma|P\right) := \sum_{x\in\mathcal{X}} P(x) D\left(W_x\|\sigma\right), \tag{2.14}$$

$$D_\alpha\left(\mathcal{W}\|\sigma|P\right) := \sum_{x\in\mathcal{X}} P(x) D_\alpha\left(W_x\|\sigma\right), \tag{2.15}$$

$$V\left(\mathcal{W}\|\sigma|P\right) := \sum_{x\in\mathcal{X}} P(x) V\left(W_x\|\sigma\right), \tag{2.16}$$

$$\widetilde{V}\left(\mathcal{W}\|\sigma|P\right) := \sum_{x\in\mathcal{X}} P(x) \widetilde{V}\left(W_x\|\sigma\right). \tag{2.17}$$

The *mutual information* of the channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ with a prior distribution $P \in \mathcal{P}(\mathcal{X})$ is defined by

$$I(P, \mathcal{W}) := D\left(P \circ \mathcal{W} \| P \otimes P\mathcal{W}\right) = D\left(\mathcal{W}\|P\mathcal{W}|P\right), \tag{2.18}$$

where $P \circ \mathcal{W} := \sum_{x\in\mathcal{X}} P(x)|x\rangle\langle x| \otimes W_x$ and $P\mathcal{W} := \sum_{x\in\mathcal{X}} P(x) W_x$. Hence, the (classical) *information capacity* of the channel $\mathcal{W}$ is

$$C_\mathcal{W} := \max_{P\in\mathcal{P}(\mathcal{X})} I(P, \mathcal{W}). \tag{2.19}$$

The *conditional information variance* and the *unconditional information variance* of $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ with a prior distribution $P \in \mathcal{P}(\mathcal{X})$ are defined, respectively, by

$$\begin{aligned} V(P, \mathcal{W}) &:= V\left(\mathcal{W}\|P\mathcal{W}|P\right), \\ U(P, \mathcal{W}) &:= V\left(P \circ \mathcal{W}\|P \otimes P\mathcal{W}\right). \end{aligned} \tag{2.20}$$

It is known that (see e.g. [3, Lemma 62]) that $V(P^\star, \mathcal{W}) = U(P^\star, \mathcal{W})$ for every capacity-achieving distribution $P^\star \in \mathcal{P}(\mathcal{X})$, i.e. $I(P^\star, \mathcal{W}) = C_\mathcal{W}$. Similarly, we also define the unconditional information variance in terms of $\widetilde{V}(\rho\|\sigma)$:

$$\widetilde{V}(P, \mathcal{W}) := \widetilde{V}\left(\mathcal{W}\|P\mathcal{W}|P\right). \tag{2.21}$$

The *minimal peripheral information variance* and its variant are defined by

$$V_\mathcal{W} := \min_{P\in\mathcal{P}(\mathcal{X}) : I(P,\mathcal{W}) = C_\mathcal{W}} V(P, \mathcal{W}), \tag{2.22}$$

$$\widetilde{V}_\mathcal{W} := \min_{P\in\mathcal{P}(\mathcal{X}) : I(P,\mathcal{W}) = C_\mathcal{W}} \widetilde{V}(P, \mathcal{W}). \tag{2.23}$$

Furthermore, one can verify that

$$V_\mathcal{W} > 0 \quad \text{implies} \quad C_\mathcal{W} > 0. \tag{2.24}$$

2.2.1. *Auxiliary functions and their properties.* The auxiliary function of a classical-quantum channel is defined as [35, 36, 37, 38, 39]

$$E_0(s, P) := -\log \operatorname{Tr}\left[\left(\sum_{x\in\mathcal{X}} P(x) W_x^{1/(1+s)}\right)^{1+s}\right].$$

In this paper, we will require three variants of the above auxiliary function: $\forall s \geq 0$ and $\sigma \in \mathcal{S}(\mathcal{H})$,

$$\widetilde{E}_0(s, P, \sigma) := s D_{1-s}\left(P \circ \mathcal{W}\|P \otimes \sigma\right) \tag{2.25}$$

$$E_\mathrm{h}(s, P, \sigma) := s D_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right), \tag{2.26}$$

$$\widetilde{E}_\mathrm{h}(s, P, \sigma) := s D^\flat_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right), \tag{2.27}$$

where D_α and $D_\alpha^\flat$ are the (Petz's) quantum Rényi divergence and the log-Euclidean Rényi divergence, respectively.

The function $\widetilde{E}_0(s, P, \sigma)$ will play a major role in the achievability part of our main result (see Theorem 4 in Section 3). This quantity yields an upper bound to the average error probability (see [27, Eq. (9)]):

$$\bar{\epsilon}(\mathcal{W}, \mathcal{C}_n) \leq 4 \exp \left\{ -n \left[\max_{0 \leq s \leq 1} \max_{P \in \mathcal{P}(\mathcal{X})} \left\{ -sR + \widetilde{E}_0(s, P, P\mathcal{W}) \right\} \right] \right\}. \tag{2.28}$$

Properties of E_h and $\widetilde{E}_\mathrm{h}$ will be crucial in the analysis of the converse part of our main result. The following proposition summarizes properties of $\widetilde{E}_0(s, P, \sigma)$. We provide the proof in Appendix A.1.

Proposition 1 (Properties of $\widetilde{E}_0(s, P, \sigma)$)**.** *Consider a classical-quantum channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$, a distribution $P \in \mathcal{P}(\mathcal{X})$, and a state $\sigma \in \mathcal{S}(\mathcal{H})$ with $W_x \ll \sigma$ for all $x \in \mathsf{supp}(P)$. Then $\widetilde{E}_0(s, P, \sigma)$ defined in Eq.* (2.25) *enjoys the following properties.*

(a) $\widetilde{E}_0(s, P, \sigma)$ and its partial derivatives $\partial \widetilde{E}_0(s, P, \sigma)/\partial s$, $\partial^2 \widetilde{E}_0(s, P, \sigma)/\partial s^2$, $\partial^3 \widetilde{E}_0(s, P, \sigma)/\partial s^3$ are all continuous in $(s, P) \in \mathbb{R}_{\geq 0} \times \mathcal{P}(\mathcal{X})$.

(b) For every $P \in \mathcal{P}(\mathcal{X})$, the function $\widetilde{E}_0(s, P, \sigma)$ is concave in $s \in \mathbb{R}_{\geq 0}$.

(c) For every $P \in \mathcal{P}(\mathcal{X})$,

$$\left. \frac{\partial \widetilde{E}_0(s, P, \sigma)}{\partial s} \right|_{s=0} = D\left(P \circ \mathcal{W} \| P \otimes \sigma\right). \tag{2.29}$$

(d) For every $P \in \mathcal{P}(\mathcal{X})$,

$$\lim_{s \to +\infty} \frac{\partial \widetilde{E}_0(s, P, \sigma)}{\partial s} \leq \frac{\partial \widetilde{E}_0(s, P, \sigma)}{\partial s} \leq D\left(P \circ \mathcal{W} \| P \otimes \sigma\right), \ \forall s \in \mathbb{R}_{\geq 0}. \tag{2.30}$$

(e) For every $P \in \mathcal{P}(\mathcal{X})$,

$$\left. \frac{\partial^2 \widetilde{E}_0(s, P, \sigma)}{\partial s^2} \right|_{s=0} = -V\left(P \circ \mathcal{W} \| P \otimes \sigma\right). \tag{2.31}$$

Properties of $E_\mathrm{h}(s, P, \sigma)$ are collected in the following proposition. The proof can be found in Appendix A.2.

Proposition 2 (Properties of $E_\mathrm{h}(s, P, \sigma)$)**.** *Consider a classical-quantum channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$, a distribution $P \in \mathcal{P}(\mathcal{X})$, and a state $\sigma \in \mathcal{S}(\mathcal{H})$ with $W_x \ll \sigma$ for all $x \in \mathsf{supp}(P)$. Then $E_\mathrm{h}(s, P, \sigma)$ defined in Eq.* (2.26) *enjoys the following properties.*

(a) $E_\mathrm{h}(s, P, \sigma)$ and its partial derivatives $\partial E_\mathrm{h}(s, P, \sigma)/\partial s$, $\partial^2 E_\mathrm{h}(s, P, \sigma)/\partial s^2$, $\partial^3 E_\mathrm{h}(s, P, \sigma)/\partial s^3$ are continuous for $(s, P) \in \mathbb{R}_{\geq 0} \times \mathcal{P}(\mathcal{X})$.

(b) For every $P \in \mathcal{P}(\mathcal{X})$, the function $E_\mathrm{h}(s, P, \sigma)$ is concave in s for all $s \in \mathbb{R}_{\geq 0}$.

(c) For every $P \in \mathcal{P}(\mathcal{X})$,

$$\left. \frac{\partial E_\mathrm{h}(s, P, \sigma)}{\partial s} \right|_{s=0} = D\left(\mathcal{W} \| \sigma | P\right). \tag{2.32}$$

(d) For every $P \in \mathcal{P}(\mathcal{X})$,

$$\lim_{s \to +\infty} \frac{\partial E_\mathrm{h}(s, P, \sigma)}{\partial s} \leq \frac{\partial E_\mathrm{h}(s, P, \sigma)}{\partial s} \leq D\left(\mathcal{W} \| \sigma | P\right), \ \forall s \in \mathbb{R}_{\geq 0}. \tag{2.33}$$

(e) For every $P \in \mathcal{P}(\mathcal{X})$,

$$\left. \frac{\partial^2 E_\mathrm{h}(s, P, \sigma)}{\partial s^2} \right|_{s=0} = -V\left(\mathcal{W} \| \sigma | P\right). \tag{2.34}$$

Proposition 3 below lists the properties of $\widetilde{E}_\mathrm{h}$, and the proof is provided in Appendix A.3.

Proposition 3 (Properties of $\widetilde{E}_\mathrm{h}(s, P, \sigma)$)**.** *Consider a classical-quantum channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$, a distribution $P \in \mathcal{P}(\mathcal{X})$, and a state $\sigma \in \mathcal{S}(\mathcal{H})$ with $W_x \ll \sigma$ for all $x \in \mathsf{supp}(P)$. Then $\widetilde{E}_\mathrm{h}(s, P, \sigma)$ defined in Eq.* (2.27) *enjoys the following properties.*

(a) $\widetilde{E}_{\mathrm{h}}(s,P,\sigma)$ *and its partial derivatives* $\partial\widetilde{E}_{\mathrm{h}}(s,P,\sigma)/\partial s$, $\partial^2\widetilde{E}_{\mathrm{h}}(s,P,\sigma)/\partial s^2$, $\partial^3\widetilde{E}_{\mathrm{h}}(s,P,\sigma)/\partial s^3$ *are all continuous for* $(s,P)\in\mathbb{R}_{\geq 0}\times\mathcal{P}(\mathcal{X})$.

(b) For every $P\in\mathcal{P}(\mathcal{X})$, *the function* $\widetilde{E}_{\mathrm{h}}(s,P,\sigma)$ *is concave in* s *for all* $s\in\mathbb{R}_{\geq 0}$.

(c) For every $P\in\mathcal{P}(\mathcal{X})$,

$$\left.\frac{\partial\widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s}\right|_{s=0}=D\left(\mathcal{W}\|\sigma|P\right). \tag{2.35}$$

(d) For every $P\in\mathcal{P}(\mathcal{X})$,

$$\lim_{s\to+\infty}\frac{\partial\widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s}\leq\frac{\partial\widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s}\leq D\left(\mathcal{W}\|\sigma|P\right),\ \forall s\in\mathbb{R}_{\geq 0}. \tag{2.36}$$

(e) For every $P\in\mathcal{P}(\mathcal{X})$,

$$\left.\frac{\partial^2\widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s^2}\right|_{s=0}=-\widetilde{V}\left(\mathcal{W}\|\sigma|P\right). \tag{2.37}$$

2.2.2. *Error Exponents.* Auxiliary functions allow us to concisely define sphere-packing exponent functions of a classical-quantum channel. We will use notation similar to Refs. [40, 28, 18]. Define

$$\widetilde{E}_{\mathrm{sp}}(R,P,\sigma):=\min_{\bar{\mathcal{W}}:\mathcal{X}\to\mathcal{S}_\circ}\left\{D\left(\bar{\mathcal{W}}\|\mathcal{W}|P\right):D\left(\bar{\mathcal{W}}\|\sigma|P\right)\leq R\right\} \tag{2.38}$$

$$=\sup_{s\geq 0}\left\{\widetilde{E}_{\mathrm{h}}(s,P)-sR\right\}, \tag{2.39}$$

$$E_{\mathrm{sp}}^{(2)}(R,P,\sigma):=\sup_{s\geq 0}\left\{E_{\mathrm{h}}\left(s,P\right)-sR\right\}, \tag{2.40}$$

for all $R>0$, $P\in\mathcal{P}(\mathcal{X})$, and $\sigma\in\mathcal{S}_{>0}(\mathcal{H})$. The equality in Eq. (2.39) follows from [18, Theorem 6]. From the definitions in Eqs. (2.38) and (D.9), it is not hard to see that [30]

$$\widetilde{E}_{\mathrm{sp}}(R,P,\sigma)=0,\quad\forall R\geq D\left(\mathcal{W}\|\sigma|P\right). \tag{2.41}$$

and

$$E_{\mathrm{sp}}^{(2)}(R,P,\sigma)=\begin{cases}+\infty, & R<D_0\left(\mathcal{W}\|\sigma|P\right),\\ 0, & R\geq D\left(\mathcal{W}\|\sigma|P\right).\end{cases} \tag{2.42}$$

3. Moderate Deviations for Classical-Quantum Channels

This section presents our main results—the error performance of classical-quantum channels satisfies the moderate deviation property, Eq. (1.4). The achievability part is stated in Theorem 4, and its proof is given in Section 3.1. Our proof strategy employs Hayashi's bound [27] and the properties of the modified auxiliary function (Proposition 1). Theorem 5 contains the converse part, and is proved in Section 3.2. The proof involves a weak sphere-packing bound (Proposition 6), a sharp converse lower bound (Proposition 7), and an approximation of the error-exponent function around capacity (Proposition 8).

Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of real numbers satisfying

$$\begin{aligned}&\text{(i) } a_n\to 0,\quad\text{as}\quad n\to+\infty,\\ &\text{(ii) } a_n\sqrt{n}\to+\infty,\quad\text{as}\quad n\to+\infty.\end{aligned} \tag{3.1}$$

Theorem 4 (Achievability)**.** *For any* $\mathcal{W}:\mathcal{X}\to\mathcal{S}(\mathcal{H})$ *with* $V_{\mathcal{W}}>0$ *and any sequence* $(a_n)_{n\geq 1}$ *satisfying Eq. (3.1), there exists a sequence of codes* $\{\mathcal{C}_n\}_{n\geq 1}$ *with rates* $R_n=C_{\mathcal{W}}-a_n$ *so that*

$$\limsup_{n\to+\infty}\frac{1}{na_n^2}\log\bar{\epsilon}\left(\mathcal{W},\mathcal{C}_n\right)\leq-\frac{1}{2V_{\mathcal{W}}}. \tag{3.2}$$

The proof is given in Section 3.1.

Theorem 5 (Converse). *For any $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ with $V_\mathcal{W} > 0$, any sequence $\{a_n\}_{n\geq 1}$ satisfying Eq. (3.1), and any sequence of codes $\{\mathcal{C}_n\}_{n\geq 1}$ with rates $R_n = C_\mathcal{W} - a_n$, it holds that*

$$\liminf_{n\to+\infty} \frac{1}{na_n^2} \log \bar{\epsilon}\left(\mathcal{W}, \mathcal{C}_n\right) \geq -\frac{1}{2V_\mathcal{W}}. \tag{3.3}$$

The proof is given in Section 3.2.

Remark 3.1. Altuğ and Wagner [21] proved Theorem 5 for discrete classical channels by a weak sphere-packing bound with the expression of $\widetilde{E}_{\text{sp}}$. Although such a weak sphere-packing bound indeed holds for c-q channels (see Proposition 6 and Remark B.1 in Appendix B), Proposition 8 in Section 3.2 shows that it will lead to

$$\limsup_{n\to+\infty} \frac{1}{na_n^2} \log \bar{\epsilon}\left(\mathcal{W}, \mathcal{C}_n\right) \leq -\frac{1}{2\widetilde{V}_\mathcal{W}}, \tag{3.4}$$

where $\widetilde{V}_\mathcal{W}$ is defined in Eq. (2.23). Since $\widetilde{V}(\rho\|\sigma) \leq V(\rho\|\sigma)$ [42, Theorem 1.2], it holds that $\widetilde{V}_\mathcal{W} \leq V_\mathcal{W}$ and the equality happens if and only if the channel reduces to classical. Hence, Altuğ and Wagner's method yields a weaker result in quantum regime; namely, a gap between the achievability and the converse. In Section 3.2, we will employ a sharp converse bound from strong large deviation theory to achieve our result, Theorem 5.

3.1. **Proof of Achievability: Theorem 4.** Let $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ satisfy $V_\mathcal{W} > 0$. Let $\{a_n\}_{n\geq 1}$ be any sequence of real numbers satisfying Eq. (3.1). Since $V_\mathcal{W} > 0$, Eq. (2.24) shows that $C_\mathcal{W} > 0$. Hence, we have $C_\mathcal{W} - a_n > 0$, for all sufficiently large n. Fix such an integer n onwards, Hayashi's upper bound, Eq. (2.28), implies that there exists a code $\mathcal{C}_n$ with $R_n = C_\mathcal{W} - a_n$ so that

$$\bar{\epsilon}(\mathcal{W}, \mathcal{C}_n) \leq 4 \exp\left(-n\left[\max_{0\leq s\leq 1}\left\{\widetilde{E}_0(s, P, P\mathcal{W}) - sR_n\right\}\right]\right), \tag{3.5}$$

for all $P \in \mathcal{P}(\mathcal{X})$. In the following, we denote by $\widetilde{E}_0(s, P) := \widetilde{E}_0(s, P, P\mathcal{W})$ for notational convenience. Simple algebra yields

$$\frac{1}{na_n^2} \log \bar{\epsilon}(W, \mathcal{C}_n) \leq \frac{\log 4}{na_n^2} - \frac{1}{a_n^2} \max_{0\leq s\leq 1}\left\{\widetilde{E}_0(s, P) - sR_n\right\}, \tag{3.6}$$

for all sufficiently large n and any $P \in \mathcal{P}(\mathcal{X})$.

Let $\widetilde{\mathcal{P}}(\mathcal{X})$ be the set of distributions that achieve the minimum in Eq. (2.22), and let $\widetilde{P} \in \widetilde{\mathcal{P}}(\mathcal{X})$. Note that Ref. [9, Lemma 3] implies that $\widetilde{\mathcal{P}}(\mathcal{X})$ is compact. Applying Taylor's theorem to $\widetilde{E}_0(s, \widetilde{P})$ at $s = 0$ together with Proposition 1 gives

$$\widetilde{E}_0\left(s, \widetilde{P}\right) = sC_\mathcal{W} - \frac{s^2}{2}V_\mathcal{W} + \frac{s^3}{6}\left.\frac{\partial^3 \widetilde{E}_0\left(s, \widetilde{P}\right)}{\partial s^3}\right|_{s=\bar{s}}, \tag{3.7}$$

for some $\bar{s} \in [0, s]$. Let $s_n = a_n/V_\mathcal{W}$. Then $s_n \leq 1$ for all sufficiently large n by the assumption in Eq. (3.1) and $V_\mathcal{W} > 0$. For all $s_n \leq 1$, Eq. (3.7) yields

$$\max_{0\leq s\leq 1}\left\{\widetilde{E}_0\left(s, \widetilde{P}\right) - sR_n\right\} \geq \widetilde{E}_0\left(s_n, \widetilde{P}\right) - s_nR_n \tag{3.8}$$

$$= \frac{a_n}{V_\mathcal{W}}\left(C_\mathcal{W} - R_n\right) - \frac{a_n^2}{2V_\mathcal{W}} + \frac{a_n^3}{6V_\mathcal{W}^3}\left.\frac{\partial^3 \widetilde{E}_0\left(s, \widetilde{P}\right)}{\partial s^3}\right|_{s=\bar{s}_n} \tag{3.9}$$

$$= \frac{a_n^2}{2V_\mathcal{W}} + \frac{a_n^3}{6V_\mathcal{W}^3}\left.\frac{\partial^3 \widetilde{E}_0\left(s, \widetilde{P}\right)}{\partial s^3}\right|_{s=\bar{s}_n}, \tag{3.10}$$

where $\bar{s}_n \in [0, s_n]$ and Eq. (3.10) holds since $R_n = C_\mathcal{W} - a_n$.

Define

$$\Upsilon = \max_{(s,P)\in[0,1]\times\widetilde{\mathcal{P}}(\mathcal{X})} \left| \frac{\partial^3 \widetilde{E}_0\left(s,P\right)}{\partial s^3} \right|, \tag{3.11}$$

which is finite due to the compact set $[0,1]\times\widetilde{\mathcal{P}}(\mathcal{X})$ and item (a) in Proposition 1. Therefore, Eq. (3.10) implies that

$$\max_{0\le s\le 1}\left\{\widetilde{E}_0\left(s,\widetilde{P}\right)-sR_n\right\} \ge \frac{a_n^2}{2V_{\mathcal{W}}}+\frac{a_n^3}{6V_{\mathcal{W}}^3}\left.\frac{\partial^3\widetilde{E}_0\left(s,\widetilde{P}\right)}{\partial s^3}\right|_{s=\bar{s}_n} \tag{3.12}$$

$$\ge \frac{a_n^2}{2V_{\mathcal{W}}}-\frac{a_n^3}{6V_{\mathcal{W}}^3}\left|\left.\frac{\partial^3\widetilde{E}_0\left(s,\widetilde{P}\right)}{\partial s^3}\right|_{s=\bar{s}_n}\right| \tag{3.13}$$

$$\ge \frac{a_n^2}{2V_{\mathcal{W}}}-\frac{a_n^3}{6V_{\mathcal{W}}^3}\Upsilon, \tag{3.14}$$

for all sufficiently large n.

Substituting Eq. (3.14) into Eq. (3.6) gives

$$\frac{1}{na_n^2}\log\bar{\epsilon}(\mathcal{W},\mathcal{C}_n)\le\frac{\log 4}{na_n^2}-\frac{1}{2V_{\mathcal{W}}}\left(1-\Upsilon\frac{a_n}{3V_{\mathcal{W}}^2}\right). \tag{3.15}$$

Recall Eq. (3.1) and let $n\to+\infty$, which completes the proof:

$$\limsup_{n\to+\infty}\frac{1}{na_n^2}\log\bar{\epsilon}(\mathcal{W},\mathcal{C}_n)\le-\frac{1}{2V_{\mathcal{W}}}. \tag{3.16}$$

□

3.2. **Proof of Converse: Theorem 5.** Our strategy consists of the following steps. First, we claim that it suffices to prove Eq. (3.3) for the maximal error probability of any code $\mathcal{C}_n$, i.e. $\epsilon_{\max}(\mathcal{W},\mathcal{C}_n)$. Recall the standard expurgation method (see e.g. [41, p. 96], [50, Theorem 20], [15, p. 395]): by removing half codewords with highest error probability to arrive at $\bar{\epsilon}\left(\mathcal{W},\mathcal{C}_n\right)\ge\frac{1}{2}\epsilon_{\max}\left(\mathcal{W},\mathcal{C}_n'\right)$ with $|\mathcal{C}_n'|=\lceil|\mathcal{C}_n|/2\rceil\ge\frac{1}{2}\exp\{nR_n\}=\exp\{n(R_n-\frac{1}{n}\log 2)\}$. Since the induced rate back-off is only $\frac{1}{n}\log 2=o(a_n)$, one might define another sequence $a_n':=a_n-\frac{1}{n}\log 2$ satisfying Eq. (3.1). Hence, without of loss generality, we only need to prove the converse part for $\epsilon_{\max}$.

Second, we employ the method of Ref. [18, Lemma 16] to relate the error probability $\epsilon_{\max}$ to the minimum type-I error:

$$\frac{\log\epsilon_{\max}(\mathcal{W},\mathcal{C}_n)}{na_n^2}\ge\max_{\sigma^n\in\mathcal{S}(\mathcal{H}^{\otimes n})}\min_{\mathbf{x}^n\in\mathcal{X}^n}\frac{\log\widehat{\alpha}_{\exp\{-nR_n\}}(W_{\mathbf{x}^n}^{\otimes n}\|\sigma^n)}{na_n^2} \tag{3.17}$$

$$\ge\min_{\mathbf{x}^n\in\mathcal{X}^n}\frac{\log\widehat{\alpha}_{\exp\{-nR_n\}}(W_{\mathbf{x}^n}^{\otimes n}\|(P^\star\mathcal{W})^{\otimes n})}{na_n^2}, \tag{3.18}$$

where $P^\star\in\mathcal{P}(\mathcal{X})$ is an arbitrary capacity-achieving distribution, i.e. $I(P^\star,\mathcal{W})=C_{\mathcal{W}}$.

Third, we divide the set of codewords into two groups. Fix an arbitrary $\eta\in(0,\frac{1}{2})$. Let $A:=\max_{\rho\in\mathcal{S}_\circ}V(\rho\|P^\star\mathcal{W})$ and let $\xi=\sqrt{2A/\eta}$. Define:

$$\Omega_{\text{good}}:=\{\mathbf{x}^n\in\mathcal{X}^n:D(\mathcal{W}\|P^\star\mathcal{W}|P_{\mathbf{x}^n})>R_n\}; \tag{3.19}$$

$$\Omega_{\text{bad}}:=\mathcal{X}^n\backslash\Omega_{\text{good}}. \tag{3.20}$$

For the codes in Ω_{bad}, we employ a weak converse bound in Proposition 6, and apply a sharp converse bound, Proposition 7, for Ω_{good}. Furthermore, we can assume $a_n>0$ for all sufficiently large $n\in\mathbb{N}$ owing to the assumption $\lim_{n\to+\infty}a_n\sqrt{n}=+\infty$. Subsequently, we will consider such n onwards.

Proof of Theorem 5. We start the proof with the case Ω_{bad}, and further consider two different cases:

$$\Omega_{\text{bad}}^{(1)} := \left\{ \mathbf{x}^n \in \mathcal{X}^n : D(\mathcal{W} \| P^\star \mathcal{W} | P_{\mathbf{x}^n}) \leq R_n - \frac{2\xi}{\sqrt{n}} \right\}; \tag{3.21}$$

$$\Omega_{\text{bad}}^{(2)} := \left\{ \mathbf{x}^n \in \mathcal{X}^n : R_n - \frac{2\xi}{\sqrt{n}} < D(\mathcal{W} \| P^\star \mathcal{W} | P_{\mathbf{x}^n}) \leq R_n \right\}. \tag{3.22}$$

We apply the following weak converse bound with $\sigma = P^\star \mathcal{W}$, whose proof is provided in Appendix B to further lower bound the right-hand side of Eq. (3.18).

Proposition 6 (A Weak Converse Bound)**.** *Consider a classical-quantum channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ with $\mathcal{S}_\circ := \overline{\mathsf{im}(\mathcal{W})}$, an arbitrary rate $R \geq 0$, and $\sigma \in \mathcal{S}_{>0}(\mathcal{H})$. For any $\eta \in (0, \frac{1}{2})$, let $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$,*

$$\mathrm{e}^{-\xi\sqrt{n}} \leq \frac{\eta}{2}, \tag{3.23}$$

where $\xi = \sqrt{2A/\eta}$ and $A := \max_{\rho \in \mathcal{S}_\circ} V(\rho \| \sigma)$. Then, it holds that for all $n \geq N_0$,

$$\widehat{\alpha}_{\exp\{-nR\}} \left(W_{\mathbf{x}^n}^{\otimes n} \| \sigma^{\otimes n} \right) \geq f(\eta) \exp \left\{ -n \left[\frac{\widetilde{E}_{\text{sp}} \left(R - \frac{2\xi}{\sqrt{n}}, P_{\mathbf{x}^n}, \sigma \right)}{1 - \eta} \right] \right\}, \tag{3.24}$$

where $f(\eta) = \exp\left\{ -\frac{h(1-\eta)}{1-\eta} \right\}$ and $h(p) := -p \log p - (1-p) \log(1-p)$ is the binary entropy function.

Let η and ξ be defined as above, and let N_1 be an integer satisfying Eq. (3.23). Then Eq. (3.24) gives, for all $n \geq N_1$,

$$\frac{\log \widehat{\alpha}_{\exp\{-nR_n\}}(W_{\mathbf{x}^n}^{\otimes n} \| (P^\star \mathcal{W})^{\otimes n})}{n a_n^2} \geq -\frac{\widetilde{E}_{\text{sp}} \left(R_n - \frac{2\xi}{\sqrt{n}}, P_{\mathbf{x}^n}, P^\star \mathcal{W} \right)}{a_n^2 (1 - \eta)} + \frac{\log f(\eta)}{n a_n^2}. \tag{3.25}$$

Further, Eq. (2.41) implies that for all $\mathbf{x}^n \in \Omega_{\text{bad}}^{(1)}$,

$$\widetilde{E}_{\text{sp}} \left(R_n - \frac{2\xi}{\sqrt{n}}, P_{\mathbf{x}^n}, P^\star \mathcal{W} \right) = 0. \tag{3.26}$$

Hence, we have for all $\mathbf{x}^n \in \Omega_{\text{bad}}^{(1)}$,

$$\frac{\log \widehat{\alpha}_{\exp\{-nR_n\}}(W_{\mathbf{x}^n}^{\otimes n} \| (P^\star \mathcal{W})^{\otimes n})}{n a_n^2} \geq \frac{\log f(\eta)}{n a_n^2} \tag{3.27}$$

$$\geq -\frac{1}{2V_{\mathcal{W}}} + \frac{\log f(\eta)}{n a_n^2}, \tag{3.28}$$

where the last inequality follows from $V_{\mathcal{W}} > 0$. Since $f(\eta) < +\infty$, taking the infimum limit of $n \to +\infty$ and using Eq. (3.1) give, for all $\mathbf{x}^n \in \Omega_{\text{bad}}^{(1)}$,

$$\liminf_{n \to +\infty} \frac{\log \widehat{\alpha}_{\exp\{-nR_n\}} \left(W_{\mathbf{x}^n}^{\otimes n} \| (P^\star \mathcal{W})^{\otimes n} \right)}{n a_n^2} \geq -\frac{1}{2V_{\mathcal{W}}}. \tag{3.29}$$

Next, we move on to $\mathbf{x}^n \in \Omega_{\text{bad}}^{(2)}$. In this case, $\widetilde{E}_{\text{sp}}$ in Eq. (3.25) is not equal to zero for any finite n, we employ Eq. (3.47) in Proposition 8 below with $\delta_n = a_n + 2\xi/\sqrt{n}$ and $b_n = a_n$ to arrive at

$$\liminf_{n \to +\infty} \frac{\log \widehat{\alpha}_{\exp\{-nR_n\}} \left(W_{\mathbf{x}^n}^{\otimes n} \| (P^\star \mathcal{W})^{\otimes n} \right)}{n a_n^2} \geq -\lim_{n \to +\infty} \frac{4\xi^2}{n \left(a_n + \frac{2\xi}{\sqrt{n}} \right)^2} \cdot \frac{1}{2\widetilde{V}_{\mathcal{W}}(1 - \eta)} \tag{3.30}$$

$$= 0 \tag{3.31}$$

$$\geq -\frac{1}{2V_{\mathcal{W}}}, \tag{3.32}$$

where the equality follows since $\lim_{n \to +\infty} n a_n^2 = +\infty$.

In the last case of $\mathbf{x}^n \in \Omega_{\text{good}}$, we employ a tighter bound, Proposition 7, to lower bound the right-hand side of Eq. (3.18). The proof is delayed to Appendix C.

Proposition 7 (A Sharp Converse Bound)**.** *Consider a classical-quantum channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ and a state $\sigma \in \mathcal{S}(\mathcal{H})$. Suppose the sequence $\mathbf{x}^n \in \mathcal{X}^n$ satisfies*

$$\nu \leq V\left(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}\right) < +\infty \tag{3.33}$$

for some $\nu > 0$, and suppose the sequence of rates $(R_n)_{n\in\mathbb{N}}$ satisfies[3] $D_0(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}) < R_n < D(\mathcal{W}\|\sigma|P_{\mathbf{x}^n})$. *Then, there exists an $N_0 \in \mathbb{N}$ such that, for all $n \geq N_0$,*

$$\widehat{\alpha}_{\exp\{-nR_n\}}(W_{\mathbf{x}^n}^{\otimes n}\|\sigma^{\otimes n}) \geq \frac{A}{s_n^\star\sqrt{n}} \exp\left\{-nE_{\text{sp}}^{(2)}\left(R_n - c_n, P_{\mathbf{x}^n}, \sigma\right)\right\}, \tag{3.34}$$

where $c_n = \frac{K\log n}{n}$ and $A, K > 0$ are finite constants independent of the sequence $\mathbf{x}^n$, and

$$s_n^\star := \arg\max_{s\geq 0}\left\{E_{\text{h}}(s, P_{\mathbf{x}^n}, \sigma) - sR_n\right\}. \tag{3.35}$$

Before applying Proposition 7, we verify that the condition, Eq. (3.33), is satisfied. Define

$$v(\delta) := \min_{P\in\mathcal{P}(\mathcal{X})}\left\{V\left(\mathcal{W}\|P^\star\mathcal{W}|P\right) : D(\mathcal{W}\|P^\star\mathcal{W}|P) \geq C_{\mathcal{W}} - \delta\right\}. \tag{3.36}$$

Note that the map $\delta \mapsto v(\delta)$ is monotone decreasing and continuous at 0 from above, i.e. $\lim_{\delta\downarrow 0} v(\delta) = v(0) = V_{\mathcal{W}}$ [7, Lemma 22]. For any $\kappa \in (0, 1)$, we can choose a sufficiently small $\gamma > 0$ independent of the sequence $\mathbf{x}^n$ such that $v(\gamma) \geq (1-\kappa)V_{\mathcal{W}} =: \nu > 0$. Further, let $N_2 \in \mathbb{N}$ such that $a_n \leq \gamma$ for all $n \geq N_2$. Then, one finds, for all $\mathbf{x}^n \in \Omega_{\text{good}}$ and $n \geq N_2$,

$$V\left(\mathcal{W}\|P^\star\mathcal{W}|P_{\mathbf{x}^n}\right) \geq v(\gamma) \geq \nu > 0. \tag{3.37}$$

Moreover, since $V_{\mathcal{W}} > 0$ implies that $C_{\mathcal{W}} = \max_{P\in\mathcal{P}(\mathcal{X})} D(\mathcal{W}\|P^\star\mathcal{W}|P) > \max_{P\in\mathcal{P}} D_0(\mathcal{W}\|P^\star\mathcal{W}|P)$, one can choose a sufficiently large n, say $N_3 \in \mathbb{N}$, such that $R_n > D_0(\mathcal{W}\|P^\star\mathcal{W}|P_{\mathbf{x}^n})$ for all $n \geq N_3$. Now, we have for all $\mathbf{x}^n \in \Omega_{\text{good}}$ and $n \geq \max_{\{} N_2, N_3\}$ that

$$\max_{P\in\mathcal{P}(\mathcal{X})} D_0(\mathcal{W}\|P^\star\mathcal{W}|P) < R_n < D(\mathcal{W}\|P^\star\mathcal{W}|P_{\mathbf{x}^n}); \tag{3.38}$$

$$0 < \nu \leq V(\mathcal{W}\|P^\star\mathcal{W}|P_{\mathbf{x}^n}). \tag{3.39}$$

Together with Eqs. (3.18) and (3.37) and letting $\sigma = P^\star\mathcal{W}$, Proposition 7 yields, for all $\mathbf{x}^n \in \Omega_{\text{good}}$ and all sufficiently large n, say $n \geq N_4 \in \mathbb{N}$,

$$\frac{\log\widehat{\alpha}_{\exp\{-nR_n\}}\left(W_{\mathbf{x}^n}^{\otimes n}\|(P^\star\mathcal{W})^{\otimes n}\right)}{na_n^2} \geq -\frac{E_{\text{sp}}^{(2)}\left(R_n - c_n, P_{\mathbf{x}^n}, P^\star\mathcal{W}\right)}{a_n^2} - \frac{\log s_n^\star\sqrt{n}}{na_n^2} + \frac{\log A}{na_n^2}. \tag{3.40}$$

Recall Eq. (3.48) in Proposition 8 below with $b_n = 0$ and $\delta_n = a_n + c_n$ that $\limsup_{n\to+\infty}\frac{s_n^\star}{a_n+c_n} \leq \frac{1}{V_{\mathcal{W}}}$. Hence, one can fix an arbitrary $\zeta > 0$ and there exists an $N_5 \in \mathbb{N}$ such that $\frac{s_n^\star\sqrt{n}}{(a_n+c_n)\sqrt{n}} \leq \frac{1}{V_{\mathcal{W}}} + \zeta$ for all $n \geq N_5$. This then leads to for all sufficiently large $n \geq \max\{N_2, N_3, N_4, N_5\}$ and all $\mathbf{x}^n \in \Omega_{\text{good}}$,

$$\frac{\log\widehat{\alpha}_{\exp\{-nR_n\}}\left(W_{\mathbf{x}^n}^{\otimes n}\|(P^\star\mathcal{W})^{\otimes n}\right)}{na_n^2} \geq -\frac{E_{\text{sp}}^{(2)}\left(R_n - c_n, P_{\mathbf{x}^n}, P^\star\mathcal{W}\right)}{a_n^2} - \frac{\log(a_n+c_n)\sqrt{n}}{na_n^2} + \frac{\log\frac{A}{\frac{1}{V_{\mathcal{W}}}+\zeta}}{na_n^2}. \tag{3.41}$$

Taking $n \to +\infty$, the second and the third terms on the right-hand side of Eq. (3.41) vanish since $c_n = K\frac{\log n}{n} = o(a_n)$ and the assumption $\lim_{n\to+\infty} a_n\sqrt{n} = +\infty$.

[3] Note that $D_0(\mathcal{W}\|\sigma|P) = D(\mathcal{W}\|\sigma|P)$ implies $W_x = \sigma$ for all $x \in \mathsf{supp}(P)$ [46, Collorary 4.1]. This further gives $V(\mathcal{W}\|\sigma|P) = 0$. However, the assumption in Eq. (3.33) ensures that $\liminf_{n\in\mathbb{N}} D(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}) - D_0(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}) > 0$. Hence, the intervals $[D_0(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}), D(\mathcal{W}\|\sigma|P_{\mathbf{x}^n})]$ for all $\mathbf{x}^n$ satisfying Eq. (3.33) are not measure zero.

Next, we apply Eq. (3.46) in Proposition 8 again to bound the error-exponent function $E_{\text{sp}}^{(2)}$ in Eq. (3.40): for all $\mathbf{x}^n \in \Omega^{(3)}$

$$\begin{aligned}
\liminf_{n\to+\infty} \frac{\log \widehat{\alpha}_{\exp\{-nR_n\}}\left(W_{\mathbf{x}^n}^{\otimes n} \| (P^\star \mathcal{W})^{\otimes n}\right)}{n a_n^2} &\geq -\limsup_{n\to+\infty} \frac{E_{\text{sp}}^{(2)}\left(C_{\mathcal{W}} - \delta_n, P_{\mathbf{x}^n}, P^\star\mathcal{W}\right)}{a_n^2} && (3.42)\\
&= -\limsup_{n\to+\infty} \frac{E_{\text{sp}}^{(2)}\left(C_{\mathcal{W}} - \delta_n, P_{\mathbf{x}^n}, P^\star\mathcal{W}\right)}{\delta_n^2} && (3.43)\\
&\geq -\frac{1}{2V_{\mathcal{W}}}. && (3.44)
\end{aligned}$$

Finally, combining Eqs. (3.18), (3.29), (3.32) and (3.44) concludes the desired Eq. (3.3).

Proposition 8 (Error Exponent around Capacity)**.** *Let $(b_n)_{n\in\mathbb{N}}$ be a sequence of real numbers with $\lim_{n\to+\infty} b_n = 0$ and let $(\delta_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers with $\lim_{n\to+\infty}\delta_n = 0$. Suppose the sequence of distributions $(P_n)_{n\in\mathbb{N}}$ satisfies*

$$C_{\mathcal{W}} - \delta_n < D(\mathcal{W}\|P^\star\mathcal{W}|P_n) \leq C_{\mathcal{W}} - b_n. \tag{3.45}$$

The following hold:

$$\limsup_{n\to+\infty} \frac{E_{\text{sp}}^{(2)}\left(C_{\mathcal{W}} - \delta_n, P_n, P^\star\mathcal{W}\right)}{\delta_n^2} \leq \limsup_{n\to+\infty} \frac{(\delta_n - b_n)^2}{2V_{\mathcal{W}}\delta_n^2}; \tag{3.46}$$

$$\limsup_{n\to+\infty} \frac{\widetilde{E}_{\text{sp}}\left(C_{\mathcal{W}} - \delta_n, P_n, P^\star\mathcal{W}\right)}{\delta_n^2} \leq \limsup_{n\to+\infty} \frac{(\delta_n - b_n)^2}{2\widetilde{V}_{\mathcal{W}}\delta_n^2}; \tag{3.47}$$

$$\limsup_{n\to+\infty} \frac{s_n^\star}{\delta_n} \leq \frac{1}{V_{\mathcal{W}}}, \tag{3.48}$$

where

$$s_n^\star := \operatorname*{arg\,max}_{s\geq 0} \left\{ E_{\text{h}}(s, P_n, P^\star\mathcal{W}) - s\left(C_{\mathcal{W}} - \delta_n\right)\right\}. \tag{3.49}$$

The proof of Proposition 8 is provided in Appendix D.

□

4. Moderate Deviations for Quantum Hypothesis Testing

In this section, we show that a special case of channel coding yields the moderate deviation result for quantum hypothesis testing. The achievability part is given in Theorem 9. In Section 4.1, we provide two proofs. The first proof follows the idea of asymptotic expansions in Theorem 4; however, we will employ Audenaet *et al.*'s quantum Hoeffding bound [30], instead of Hayashi's inequality [27]. The second proof relies on a martingale inequality [24]. The converse part and its proof are provided in Theorem 10 and Section 4.2, respectively.

Theorem 9 (Achievability)**.** *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ be the density operators with finite relative variance $V := V\left(\rho\|\sigma\right) > 0$. For any sequence of real numbers $(a_n)_{n\in\mathbb{N}}$ satisfying Eq. (3.1), there exists a sequence $r_n := D\left(\rho\|\sigma\right) - a_n$ such that*

$$\limsup_{n\to+\infty} \frac{1}{n a_n^2} \log \widehat{\alpha}_{\exp\{-nr_n\}}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) \leq -\frac{1}{2V}. \tag{4.1}$$

Theorem 10 (Converse)**.** *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ be the density operators with non-zero and finite relative variance $V := V\left(\rho\|\sigma\right) > 0$. For any sequence of real numbers $\{a_n\}_{n\in\mathbb{N}}$ satisfying Eq. (3.1), there exists a sequence $r_n := D\left(\rho\|\sigma\right) - a_n$ such that*

$$\liminf_{n\to+\infty} \frac{1}{n a_n^2} \log \widehat{\alpha}_{\exp\{-nr_n\}}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) \geq -\frac{1}{2V}. \tag{4.2}$$

4.1. **Proof of Achievability: Theorem 9.** In this section, we present two proofs for Theorem 9. The first one relies on the quantum Hoeffding bound [30] and the Taylor's expansion of the exponent function E_{h}.

The first proof of Theorem 9. Recall the following achievability of the quantum Hoeffding bound:

Lemma 11 (Theorem 5, Section 5.5 of [30])**.** *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$. For any $r \geq 0$ and any $n \in \mathbb{N}$, we have*

$$\widehat{\alpha}_{\exp\{-nr\}}\left(\rho^{\otimes n} \| \sigma^{\otimes n}\right) \leq \exp\left\{-n\left[\sup_{0<\alpha\leq 1}\left\{\frac{\alpha-1}{\alpha}\left(r - D_\alpha\left(\rho\|\sigma\right)\right)\right\}\right]\right\}. \tag{4.3}$$

Since $D(\rho\|\sigma) > 0$ (due to Eq. (2.12)), we have

$$r_n := D(\rho\|\sigma) - a_n > 0 \tag{4.4}$$

for all sufficiently large n. Choose such n onwards, then Eq. (4.3) implies that:

$$\begin{aligned} \frac{1}{na_n^2}\log\widehat{\alpha}_{\exp\{-nr_n\}}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) &\leq -\frac{1}{a_n^2}\sup_{0<\alpha\leq 1}\left\{\frac{\alpha-1}{\alpha}\left(r_n - D_\alpha\left(\rho\|\sigma\right)\right)\right\} &(4.5)\\ &= -\frac{1}{a_n^2}\sup_{s\geq 0}\left\{E_{\mathrm{h}}(s) - sr_n\right\}, &(4.6)\end{aligned}$$

where we substitute $s = \frac{1-\alpha}{\alpha}$ and let

$$E_{\mathrm{h}}(s) := sD_{\frac{1}{1+s}}\left(\rho\|\sigma\right). \tag{4.7}$$

Taylor's theorem followed by simple calculation yields

$$E_{\mathrm{h}}(s) = sD(\rho\|\sigma) - \frac{s^2}{2}V + \frac{s^3}{6}\frac{\partial^3 E_{\mathrm{h}}(s)}{\partial s^3}\bigg|_{s=\bar{s}} \tag{4.8}$$

for some $\bar{s} \in [0, s]$ and all $s \geq 0$. The above equation is also a simple consequence of items (c) and (e) in Proposition 2. Now let $s_n = a_n/V$, for all $n \in \mathbb{N}$. Then for all sufficiently large n and for some $\bar{s}_n \in [0, s_n]$, Eq. (4.8) yields

$$\begin{aligned}\sup_{s\geq 0}\left\{E_{\mathrm{h}}(s) - sr_n\right\} &\geq E_h(s_n) - s_n r_n &(4.9)\\ &= \frac{a_n}{V}\left(D(\rho\|\sigma) - r_n\right) - \frac{a_n^2}{2V} + \frac{a_n^3}{6V^3}\frac{\partial^3 E_{\mathrm{h}}(s)}{\partial s^3}\bigg|_{s=\bar{s}_n} &(4.10)\\ &= \frac{a_n^2}{2V} + \frac{a_n^3}{6V^3}\frac{\partial^3 E_{\mathrm{h}}(s)}{\partial s^3}\bigg|_{s=\bar{s}_n}, &(4.11)\end{aligned}$$

where we substitute $r_n = D(\rho\|\sigma) - a_n$ in Eq. (4.11).

Define

$$\Upsilon := \max_{s\in[0,1]}\left|\frac{\partial^3 E_{\mathrm{h}}(s)}{\partial s^3}\right|, \tag{4.12}$$

which is finite. Therefore, Eq. (4.11) leads to

$$\begin{aligned}\sup_{s\geq 0}\left\{E_{\mathrm{h}}(s) - sr_n\right\} &\geq \frac{a_n^2}{2V} + \frac{a_n^3}{6V^3}\frac{\partial^3 E_{\mathrm{h}}(s)}{\partial s^3}\bigg|_{s=\bar{s}_n} &(4.13)\\ &\geq \frac{a_n^2}{2V} - \frac{a_n^3}{6V^3}\Upsilon &(4.14)\end{aligned}$$

for all sufficiently large n. Substituting Eq. (4.14) into Eq. (4.6) yields

$$\frac{1}{na_n^2}\log\widehat{\alpha}_{\exp\{-nr_n\}}\left(\rho\|\sigma\right) \leq -\frac{1}{2V}\left(1 - \Upsilon\frac{a_n}{3V^2}\right), \tag{4.15}$$

which implies the desired achievability part:

$$\limsup_{n\to+\infty}\frac{1}{na_n^2}\log\widehat{\alpha}_{\exp\{-nr_n\}}\left(\rho\|\sigma\right) \leq -\frac{1}{2V}. \tag{4.16}$$

□

In the following, we give an alternative proof of Theorem 9 by employing a martingale inequality [24].

The second proof of Theorem 9. We follow the idea in Ref. [6] to write the eigendecomposition of $\rho^{\otimes n}$ and $\sigma^{\otimes n}$, respectively, as

$$\rho^{\otimes n} = \sum_{x^n} \lambda^n(x^n) |f^n_{x^n}\rangle\langle f^n_{x^n}|; \quad \sigma^{\otimes n} = \sum_{y^n} \gamma^n(y^n) |g^n_{y^n}\rangle\langle g^n_{y^n}|, \tag{4.17}$$

where $x^n := x_1 x_2 \dots x_n$; $y^n := y_1 y_2 \dots y_n$; $\lambda^n(x^n) = \prod_{i=1}^n \lambda(x_i)$; $\mu^n(y^n) = \prod_{i=1}^n \mu(y_i)$; $|f^n_{x^n}\rangle = |f_{x_1}\rangle \otimes |f_{x_2}\rangle \otimes \cdots \otimes |f_{x_n}\rangle$; and $|g^n_{y^n}\rangle = |g_{y_1}\rangle \otimes |g_{y_2}\rangle \otimes \cdots \otimes |g_{y_n}\rangle$. Further, we define a pair of random variables (X, Y) via the Nussbaum-Szkoła mapping [58], i.e. $P_{X,Y}(x,y) = \lambda(x)|\gamma_{xy}|^2$, where $\gamma_{xy} := \langle g_y | f_x\rangle \in \mathbb{C}$. It is well-known that

$$D(\rho\|\sigma) = D(\lambda(X)\|\mu(Y)) = \mathbb{E}_{(X,Y)}\left[\log \frac{\lambda(X)}{\mu(Y)}\right], \tag{4.18}$$

$$V(\rho\|\sigma) = V(\lambda(X)\|\mu(Y)) = \mathrm{Var}_{(X,Y)}\left[\log \frac{\lambda(X)}{\mu(Y)}\right]. \tag{4.19}$$

Let $T_n := \exp\{nr_n\}$. For every sequence x^n, we define a sub-normalized vector:

$$|\xi^n_{x_n}\rangle := \sum_{y^n : \lambda^n(x^n)/\mu^n(y^n) \geq T_n} \gamma^n_{x^n y^n} |g^n_{y^n}\rangle \tag{4.20}$$

with $\gamma^n_{x^n y^n} = \prod_{i=1}^n \gamma_{x_i y_i}$ and $\sum_x |\gamma_{xy}|^2 = \sum_y |\gamma_{xy}|^2 = 1$. Applying the Gram-Schmidt orthonormalization process on $\{|\xi^n_{x^n}\rangle\}_{x^n}$ to obtain an orthonormal vectors

$$|\hat{\xi}^n_{x^n}\rangle = \sum_{y^n : \lambda^n(x^n)/\mu^n(y^n) \geq T_n} t^n_{x^n y^n} |g^n_{y^n}\rangle \tag{4.21}$$

for some $t^n_{x^n y^n} \in \mathbb{C}$ and

$$\sum_{y^n : \lambda^n(x^n)/\mu^n(y^n) \geq T^n} |t^n_{x^n y^n}|^2 = 1. \tag{4.22}$$

We define a test of the hypotheses by

$$Q_n := \sum_{x^n} |\hat{\xi}^n_{x^n}\rangle\langle \hat{\xi}^n_{x^n}|. \tag{4.23}$$

Then, it suffices to show $\beta\left(Q_n; \sigma^{\otimes n}\right) \leq \exp\{-nr_n\}$ and

$$\lim_{n\to+\infty} \frac{1}{na_n^2} \log \alpha\left(Q_n; \rho^{\otimes n}\right) \leq -\frac{1}{2V} \tag{4.24}$$

to complete the proof. The former follows Eqs. (4.17), (4.21), and (4.22):

$$\begin{aligned} \beta\left(Q_n; \sigma^{\otimes n}\right) &= \sum_{x^n} \mathrm{Tr}\left[\sigma^{\otimes n} |\hat{\xi}^n_{x^n}\rangle\langle \hat{\xi}^n_{x^n}|\right] \\ &= \sum_{x^n} \sum_{y^n : \lambda^n(x^n)/\mu^n(y^n) \geq T^n} |t^n_{x^n y^n}|^2 \mu^n(y^n) \\ &\leq \sum_{x^n} \frac{\lambda^n(x^n)}{T_n} = \frac{1}{T_n} = \exp\{-nr_n\}. \end{aligned} \tag{4.25}$$

Likewise, since $\frac{|\xi_{x^n}\rangle\langle\xi_{x^n}|}{|\langle\xi_{x^n}|\xi_{x^n}\rangle|^2} \leq Q_n$, one can verify that

$$\alpha\left(Q_n;\rho^{\otimes n}\right) \leq 1 - \sum_{x^n} \lambda^n(x^n)\langle\xi^n_{x^n}|\xi^n_{x^n}\rangle \tag{4.26}$$

$$= \Pr\left\{\frac{\lambda^n(X^n)}{\mu^n(Y^n)} < T_n\right\} \tag{4.27}$$

$$= \Pr\left\{\log\frac{\lambda^n(X^n)}{\mu^n(Y^n)} < nr_n\right\}. \tag{4.28}$$

Next, we adopt Sason's approach [24] to construct a martingale sequence $\{U_k, \mathfrak{M}_k\}_{k=0}^n$, where $\mathfrak{M}_k$ denotes the sigma-algebra formed by $(X_l, Y_l)_{l=1}^k$; $\mathfrak{M}_0 \subseteq \mathfrak{M}_1 \subseteq \ldots \subseteq \mathfrak{M}_n$ is the filtration; and

$$U_k := \mathbb{E}_{(X^n,Y^n)}\left[\log\frac{\lambda^n(X^n)}{\mu^n(Y^n)}\middle|\mathfrak{M}_k\right] \tag{4.29}$$

$$= \sum_{i=1}^k \log\frac{\lambda(X_i)}{\mu(Y_i)} + \sum_{i=k+1}^n \mathbb{E}_{X^n}\left[\log\frac{\lambda(X_i)}{\mu(Y_i)}\right] \tag{4.30}$$

$$= \sum_{i=1}^k \log\frac{\lambda(X_i)}{\mu(Y_i)} + (n-k)D(\lambda(X)\|\mu(Y)). \tag{4.31}$$

In particular, we have

$$U_0 = nD\left(\lambda(X)\|\mu(Y)\right);\, U_n = \log\frac{\lambda(X^n)}{\mu(Y^n)} = \sum_{i=1}^n \log\frac{\lambda(X_i)}{\mu(Y_i)}.$$

Hence, it can be verified that:

$$\begin{aligned}
&U_k - U_{k-1} = \log\frac{\lambda(X_k)}{\mu(Y_k)} - D(\lambda(X)\|\mu(Y));\\
&\mathbb{E}_{X^n}\left[U_k - U_{k-1}|\,\mathfrak{M}_{k-1}\right] = 0;\\
&\mathbb{E}_{X^n}\left[(U_k - U_{k-1})^2\middle|\mathfrak{M}_{k-1}\right] = V\left(\lambda(X)\|\mu(Y)\right) = V.
\end{aligned}$$

Let

$$b := \max_{(x,y):x=y}\left|\log\frac{\lambda(x)}{\mu(y)} - D(\lambda(X)\|\mu(Y))\right|, \tag{4.32}$$

which is a finite number due to the assumption of the finite-dimensional Hilbert space. Then, we have $|U_k - U_{k-1}| \leq b$ almost surely for every $k \in [n]$. Equipped with the notation above, Eq. (4.28) can be expressed as:

$$\alpha\left(Q_n;\rho^{\otimes n}\right) = \Pr\left\{U_n - U_0 \leq -na_n\right\}. \tag{4.33}$$

In the following, we borrow the idea from Sason [24] to employ a martingale inequality to upper bound Eq. (4.33).

Theorem 12 (Refined Azuma's Inequality [24, Theorem 2])**.** *Let* $(X_k)_{k=1}^n$ *be a martingale with respect to the filtration* $(\mathfrak{M}_k)_{k=0}^n$ *such that the following requirements are satisfied almost surely:* (i) $\mathbb{E}\left[X_k|\mathfrak{M}_{k-1}\right] = 0$*;* (ii) $\mathbb{E}\left[X_k^2|\mathfrak{M}_{k-1}\right] \leq v$*;* (iii) $\|X_k\|_\infty \leq b_k$*. For any* $x \geq 0$,

$$\Pr\left\{\sum_{k=1}^n X_k \geq xn\right\} = \Pr\left\{\sum_{k=1}^n X_k \leq -xn\right\}$$

$$\leq 2\exp\left\{-nh\left(\frac{bx+v}{b^2+v}\middle\|\frac{v}{b^2+v}\right)\right\}, \tag{4.34}$$

where $h(p\|q) := p\log\frac{p}{q} + (1-p)\log\frac{1-p}{1-q}$.

Apply Theorem 12 to Eq. (4.33) with $x = a_n$, $X_k = U_k - U_{k-1}$ for ever $k \in [n]$:

$$\alpha\left(Q_n;\rho^{\otimes n}\right) \leq 2\exp\left\{-nh\left(\frac{ba_n+V}{b^2+V}\middle\|\frac{V}{b^2+V}\right)\right\}. \tag{4.35}$$

By using a scalar inequality [24, Lemma 1]:

$$(1+u)\log(1+u) \geq u + \frac{u^2}{2} - \frac{u^3}{6}, \quad u \geq 0, \tag{4.36}$$

and the definition of $h(\cdot\|\cdot)$ in Theorem 12, Eq. (4.35) leads to

$$\alpha\left(Q_n;\rho^{\otimes n}\right) \leq 2\exp\left\{-n\left[\frac{a_n^2}{2V}\left(1-\frac{a_n b}{3V(1+V/b^2)}\right)\right]\right\}. \tag{4.37}$$

Finally, recall that $\lim_{n\to+\infty} a_n = 0$ in Eq. (3.1), then

$$\limsup_{n\to+\infty}\frac{1}{na_n^2}\log\alpha_n\left(\eta_n\right) \leq -\frac{1}{2V}.$$

□

4.2. **Proof of Converse: Theorem 10.** The converse part is a direct consequence of the sharp converse Hoeffding bound, Theorem 7.

Let $\mathcal{X} = \{x\}$ and $W_x = \rho$. We apply Theorem 7 with $r = r_n$ to obtain

$$\widehat{\alpha}_{\exp\{-nr_n\}}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) \geq \frac{A}{s_n^\star\sqrt{n}}\exp\left\{-n\left[\sup_{0<\alpha\leq 1}\frac{\alpha-1}{\alpha}\left(r_n - c_n - D_\alpha\left(\rho\|\sigma\right)\right)\right]\right\}, \tag{4.38}$$

for sufficiently large $n \in \mathbb{N}$ and some constant $A > 0$. Here

$$s_n^\star := \arg\max_{s\geq 0}\left\{sD_{\frac{1}{1+s}}(\rho\|\sigma) - sr_n\right\}. \tag{4.39}$$

Now let

$$\delta_n := a_n + c_n, \quad \forall n \in \mathbb{N}, \tag{4.40}$$

and invoke Proposition 8 with $W_x = \rho$, $P(x) = 1$, and substitute $P^\star\mathcal{W}$ with σ to obtain

$$\limsup_{n\to+\infty}\frac{\sup_{s\geq 0}\left\{-s\left(D\left(\rho\|\sigma\right)-\delta_n\right)+sD_{\frac{1}{1+s}}\left(\rho\|\sigma\right)\right\}}{\delta_n^2} \leq \frac{1}{2V}. \tag{4.41}$$

Moreover, Eq. (3.48) in Proposition 8 gives that $\lim_{n\to+\infty}\frac{s_n^\star}{\delta_n} = 1/V$. Combining Eqs. (4.38) and (4.41) concludes our claim:

$$\liminf_{n\to+\infty}\frac{\log\widehat{\alpha}_{\exp\{-nr_n\}}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right)}{n\delta_n^2} \geq -\frac{1}{2V}. \tag{4.42}$$

5. Conclusion

A practical question in quantum information theory is that—is it possible for a reliable communication through a c-q channel when the transmission rate approaches capacity in blocklength? In this paper, we propose a moderate deviation analysis for c-q channel and thus give an affirmative answer. Moreover, we also establish the moderate deviations for quantum hypothesis testing.

Our proof strategy is based on a strong large deviation theory [28, 18] and the study of the asymptotic behaviour of the error exponent function. As a result, we successfully bridge the connection between small error regime and the medium error regime. On the other hand, the recent work from the authors [31] also obtains the moderate deviation result via the techniques in the non-vanishing error regime. It is remarkable that both methods from different regimes arrive at the same place, and hence both this work along with Ref. [31] illuminate the whole picture of the three regimes in quantum information theory. □

Acknowledgements

MH is supported by an ARC Future Fellowship under Grant FT140100574. We would like to thank Vincent Tan for introducing us to Altuğ and Wagner's work of moderate deviation analysis. We also thank Christopher Chubb and Marco Tomamichel for the insightful discussions and the useful comments.

Appendix A. Properties of Auxiliary Functions

This section contains proofs of Propositions 1 and 2. Most results follow from properties of Petz quantum Rényi divergence [43] (see also [44, 45, 46]).

A.1. Proof of Proposition 1.

Proposition 1 (Properties of $\widetilde{E}_0(s,P,\sigma)$)**.** *For any classical-quantum channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$, the modified auxiliary function $\widetilde{E}_0(s,P,\sigma)$ admits the following properties.*

- *(a) $\widetilde{E}_0(s,P,\sigma)$ and its partial derivatives $\partial\widetilde{E}_0(s,P,\sigma)/\partial s$, $\partial^2\widetilde{E}_0(s,P,\sigma)/\partial s^2$, $\partial^3\widetilde{E}_0(s,P,\sigma)/\partial s^3$ are all continuous in $(s,P) \in \mathbb{R}_{\geq 0} \times \mathcal{P}(\mathcal{X})$.*
- *(b) For every $P \in \mathcal{P}(\mathcal{X})$, the function $\widetilde{E}_0(s,P,\sigma)$ is concave in $s \in \mathbb{R}_{\geq 0}$.*
- *(c) For every $P \in \mathcal{P}(\mathcal{X})$,*

$$\left.\frac{\partial\widetilde{E}_0(s,P,\sigma)}{\partial s}\right|_{s=0} = D(P \circ \mathcal{W} \| P \otimes \sigma) \tag{A.1}$$

- *(d) For every $P \in \mathcal{P}(\mathcal{X})$,*

$$\lim_{s\to+\infty} \frac{\partial\widetilde{E}_0(s,P)}{\partial s} \leq \frac{\partial\widetilde{E}_0(s,P)}{\partial s} \leq D(P \circ \mathcal{W} \| P \otimes \sigma), \ \forall s \in \mathbb{R}_{\geq 0}. \tag{A.2}$$

- *(e) For every $P \in \mathcal{P}(\mathcal{X})$,*

$$\left.\frac{\partial^2\widetilde{E}_0(s,P)}{\partial s^2}\right|_{s=0} = -V(P \circ \mathcal{W} \| P \otimes \sigma). \tag{A.3}$$

Proof of Proposition 1.

(1-(a)) The continuity can be proved by the standard approach of functional calculus (see e.g. [44, Lemma III.1] and [45, Section 4.2]). Let $\widetilde{F}(s) := \sum_{x\in\mathcal{X}} P(x) \operatorname{Tr}\left[W_x^{1-s}(\sigma)^s\right]$. Direct calculation shows that

$$\frac{\partial\widetilde{E}_0(s,P,\sigma)}{\partial s} = -\frac{\widetilde{F}'(s)}{\widetilde{F}(s)}, \tag{A.4}$$

$$\frac{\partial^2\widetilde{E}_0(s,P,\sigma)}{\partial s^2} = -\frac{\widetilde{F}''(s)}{\widetilde{F}(s)} + \left(\frac{\partial\widetilde{E}_0(s,P,\sigma)}{\partial s}\right)^2, \tag{A.5}$$

$$\frac{\partial^3\widetilde{E}_0(s,P,\sigma)}{\partial s^3} = -\frac{\widetilde{F}'''(s,P)}{\widetilde{F}(s,P)} + 3\frac{\partial\widetilde{E}_0(s,P,\sigma)}{\partial s}\frac{\partial^2\widetilde{E}_0(s,P,\sigma)}{\partial s^2} - \left(\frac{\partial\widetilde{E}_0(s,P,\sigma)}{\partial s}\right)^3, \tag{A.6}$$

and

$$\widetilde{F}'(s) = \sum_{x\in\mathcal{X}} P(x)\,\mathrm{Tr}\left[-W_x^{1-s}\log(W_x)(\sigma)^s + W_x^{1-s}(\sigma)^s\log(\sigma)\right], \tag{A.7}$$

$$\begin{aligned}\widetilde{F}''(s) = \sum_{x\in\mathcal{X}} P(x)\,\mathrm{Tr}\big[&W_x^{1-s}\log^2(W_x)(\sigma)^s - W_x^{1-s}\log(W_x)(\sigma)^s\log(\sigma)\\ &-W_x^{1-s}\log(W_x)(\sigma)^s\log(\sigma) + W_x^{1-s}(\sigma)^s\log^2(\sigma)\big],\end{aligned} \tag{A.8}$$

$$\begin{aligned}\widetilde{F}'''(s) = \sum_{x\in\mathcal{X}} P(x)\,\mathrm{Tr}\big[&-W_x^{1-s}\log^3(W_x)(\sigma)^s + W_x^{1-s}\log^2(W_x)(\sigma)^s\log(\sigma)\\ &+2W_x^{1-s}\log^2(W_x)(\sigma)^s\log(\sigma) - 2W_x^{1-s}\log(W_x)(\sigma)^s\log^2(\sigma)\\ &-W_x^{1-s}\log(W_x)(\sigma)^s\log^2(\sigma) + W_x^{1-s}(\sigma)^s\log^3(\sigma)\big].\end{aligned} \tag{A.9}$$

Since the matrix power function is continuous (with respect to the strong topology; see e.g. [47, Theorem 1.19]), we conclude the continuity of the partial derivatives Eqs. (A.4)-(A.6) in item (a).

(1-(b)) The claim follows from the concavity of the map $s \mapsto sD_{1-s}(\,\cdot\,\|\,\cdot\,)$ (see e.g. [48, Lemma III.11]).

(1-(c)) The results can be derived from evaluating Eqs. (A.4), (A.5), (A.7), and (A.8) at $s = 0$. We provide an alternative proof here. One can verify

$$\begin{aligned}\left.\frac{\partial \widetilde{E}_0(s,P,\sigma)}{\partial s}\right|_{s=0} &= D_{1-s}\left(P\circ\mathcal{W}\|P\otimes\sigma\right) - sD'_{1-s}\left(P\circ\mathcal{W}\|P\otimes\sigma\right)\big|_{s=0} &\text{(A.10)}\\ &= D_{1-s}\left(P\circ\mathcal{W}\|P\otimes\sigma\right)\big|_{s=0} &\text{(A.11)}\\ &= D(P\circ\mathcal{W}\|P\otimes\sigma). &\text{(A.12)}\end{aligned}$$

(1-(d)) The concavity of the map $s\mapsto \widetilde{E}(s,P,\sigma)$ in item (b) ensures that $\partial\widetilde{E}(s,P,\sigma)/\partial s$ is non-increasing in s. Along with Eq. (A.12), we conclude Eq. (2.30).

(1-(e)) Following from item (c), one obtain

$$\begin{aligned}\left.\frac{\partial^2 \widetilde{E}_0(s,P,\sigma)}{\partial s^2}\right|_{s=0} &= -2D'_{1-s}\left(P\circ\mathcal{W}\|P\otimes\sigma\right) + sD''_{1-s}\left(P\circ\mathcal{W}\|P\otimes\sigma\right)\big|_{s=0} &\text{(A.13)}\\ &= -2D'_{1-s}\left(P\circ\mathcal{W}\|P\otimes\sigma\right)\big|_{s=0} &\text{(A.14)}\\ &= -V(P\circ\mathcal{W}\|P\otimes\sigma), &\text{(A.15)}\end{aligned}$$

where the last equality (A.15) follows from the fact $D'_{1/1+s}(\cdot\|\cdot)|_{s=0} = V(\cdot\|\cdot)/2$ [45, Theorem 2].

□

A.2. **Proof of Proposition 2.**

Proposition 2 (Properties of $E_{\mathrm{h}}(s,P,\sigma)$)**.** *Consider a classical-quantum channel $\mathcal{W}:\mathcal{X}\to\mathcal{S}(\mathcal{H})$, a distribution $P\in\mathcal{P}(\mathcal{X})$, and a state $\sigma\in\mathcal{S}(\mathcal{H})$ with $W_x \ll \sigma$ for all $x\in\mathsf{supp}(P)$. Then $E_{\mathrm{h}}(s,P,\sigma)$ defined in Eq. (2.26) enjoys the following properties.*

(a) The partial derivatives $\partial E_{\mathrm{h}}(s,P,\sigma)/\partial s$, $\partial^2 E_{\mathrm{h}}(s,P,\sigma)/\partial s^2$, $\partial^3 E_{\mathrm{h}}(s,P,\sigma)/\partial s^3$, and $E_{\mathrm{h}}(s,P)$ are all continuous for $(s,P)\in\mathbb{R}_{\geq 0}\times\mathcal{P}(\mathcal{X})$.

(b) For every $P\in\mathcal{P}(\mathcal{X})$, the function $E_{\mathrm{h}}(s,P,\sigma)$ is concave in s for all $s\in\mathbb{R}_{\geq 0}$.

(c) For every $P\in\mathcal{P}(\mathcal{X})$,

$$\left.\frac{\partial E_{\mathrm{h}}(s,P,\sigma)}{\partial s}\right|_{s=0} = D\left(\mathcal{W}\|\sigma|P\right). \tag{A.16}$$

(d) For every $P\in\mathcal{P}(\mathcal{X})$,

$$\lim_{s\to+\infty}\frac{\partial E_{\mathrm{h}}(s,P,\sigma)}{\partial s} \leq \frac{\partial E_{\mathrm{h}}(s,P,\sigma)}{\partial s} \leq D\left(\mathcal{W}\|\sigma|P\right),\ \forall s\in\mathbb{R}_{\geq 0}. \tag{A.17}$$

(e) For every $P \in \mathcal{P}(\mathcal{X})$,

$$\left.\frac{\partial^2 E_{\mathrm{h}}(s,P,\sigma)}{\partial s^2}\right|_{s=0} = -V\left(\mathcal{W}\|\sigma|P\right). \tag{A.18}$$

Proof Proposition 2.

(2-(a)) Direct calculation yields that

$$\frac{\partial E_{\mathrm{h}}(s,P,\sigma)}{\partial s} = D_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right) - \frac{s}{(1+s)^2} D'_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right) \tag{A.19}$$

$$\frac{\partial^2 E_{\mathrm{h}}(s,P,\sigma)}{\partial s^2} = -\frac{2}{(1+s)^3} D'_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right) + \frac{s}{(1+s)^4} D''_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right) \tag{A.20}$$

$$\begin{aligned}\frac{\partial^3 E_{\mathrm{h}}(s,P,\sigma)}{\partial s^3} &= \frac{6}{(1+s)^4} D'_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right) + \frac{3-3s}{(1+s)^5} D''_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right) \\ &\quad - \frac{s}{(1+s)^6} D'''_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right).\end{aligned} \tag{A.21}$$

From Eqs. (A.19)-(A.21) and the fact that $D_{1/(1+s)}\left(\mathcal{W}\|\sigma|P\right)$, $D'_{1/(1+s)}\left(\mathcal{W}\|\sigma|P\right)$, $D''_{1/(1+s)}\left(\mathcal{W}\|\sigma|P\right)$, and $D'''_{1/(1+s)}\left(\mathcal{W}\|\sigma|P\right)$ are continuous for $(s,P) \in \mathbb{R}_{\geq 0} \times \mathcal{P}(\mathcal{X})$, we deduce the continuity property in item (a).

(2-(b)) The proof strategy follows closely with [48, Appendix B]. Let $\psi(\alpha) = \sum_{x\in\mathcal{X}} P(x) \log \operatorname{Tr}\left[W_x^{\alpha}\sigma^{1-\alpha}\right]$. Since $\alpha \mapsto \psi(\alpha)$ is convex for all $\alpha \in (0,1]$ [48, Lemma III.11], it can be written as the supremum of affine functions, i.e.

$$\psi(\alpha) = \sup_{i\in\mathcal{I}} \left\{c_i \alpha + d_i\right\} \tag{A.22}$$

for some index set $\mathcal{I}$. Hence,

$$-E_{\mathrm{h}}(s,P,\sigma) = (1+s)\psi\left(\frac{1}{1+s}\right) = \sup_{i\in\mathcal{I}} \left\{c_i + d_i(1+s)\right\}. \tag{A.23}$$

The right-hand side of Eq. (A.23), in turn, implies that the map $s \mapsto E_{\mathrm{h}}(s,P,\sigma)$ is convex for all $s \in \mathbb{R}_{\geq 0}$.

(2-(c)) From Eqs. (A.19) and (A.20), one finds

$$\left.\frac{\partial E_{\mathrm{h}}(s,P,\sigma)}{\partial s}\right|_{s=0} = D\left(\mathcal{W}\|\sigma|P\right). \tag{A.24}$$

(2-(d)) The concavity of the map $s \mapsto E_{\mathrm{h}}(s,P,\sigma)$ in item (b) ensures that $\partial E_{\mathrm{h}}(s,P,\sigma)/\partial s$ is non-increasing in s. Along with Eq. (A.24) in item (c), we conclude Eq. (2.33).

(2-(e)) Applying $D'_{1/1+s}(\cdot\|\cdot)|_{s=0} = V(\cdot\|\cdot)/2$ [45, Theorem 2], it holds that

$$\left.\frac{\partial^2 E_{\mathrm{h}}(s,P,\sigma)}{\partial s^2}\right|_{s=0} = -V\left(\mathcal{W}\|\sigma|P\right). \tag{A.25}$$

□

A.3. Proof of Proposition 3.

Proposition 3 (Properties of $\widetilde{E}_{\mathrm{h}}(s,P,\sigma)$)**.** *Consider a classical-quantum channel $\mathcal{W}: \mathcal{X} \to \mathcal{S}(\mathcal{H})$, a distribution $P \in \mathcal{P}(\mathcal{X})$, and a state $\sigma \in \mathcal{S}(\mathcal{H})$ with $W_x \ll \sigma$ for all $x \in$ $\mathtt{supp}(P)$. Then $\widetilde{E}_{\mathrm{h}}(s,P,\sigma)$ defined in Eq. (2.27) enjoys the following properties.*

(a) The partial derivatives $\partial \widetilde{E}_{\mathrm{h}}(s,P,\sigma)/\partial s$, $\partial^2 \widetilde{E}_{\mathrm{h}}(s,P,\sigma)/\partial s^2$, $\partial^3 \widetilde{E}_{\mathrm{h}}(s,P,\sigma)/\partial s^3$, and $\widetilde{E}_{\mathrm{h}}(s,P,\sigma)$ are all continuous for $(s,P) \in \mathbb{R}_{\geq 0} \times \mathcal{P}(\mathcal{X})$.

(b) For every $P \in \mathcal{P}(\mathcal{X})$, the function $\widetilde{E}_{\mathrm{h}}(s,P,\sigma)$ is concave in s for all $s \in \mathbb{R}_{\geq 0}$.

(c) For every $P \in \mathcal{P}(\mathcal{X})$,

$$\left.\frac{\partial \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s}\right|_{s=0} = D\left(\mathcal{W}\|\sigma|P\right). \tag{A.26}$$

(d) For every $P \in \mathcal{P}(\mathcal{X})$,

$$\lim_{s\to+\infty} \frac{\partial \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s} \leq \frac{\partial \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s} \leq D\left(\mathcal{W}\|\sigma|P\right), \ \forall s \in \mathbb{R}_{\geq 0}. \tag{A.27}$$

(e) For every $P \in \mathcal{P}(\mathcal{X})$,

$$\left.\frac{\partial^2 \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s^2}\right|_{s=0} = -\widetilde{V}\left(\mathcal{W}\|\sigma|P\right). \tag{A.28}$$

Proof of Proposition 3. This proof follows similarly from Proposition 2.

(3-(a)) Direct calculation yields that

$$\begin{aligned}\frac{\partial \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s} &= \widetilde{D}_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right) - \frac{s}{(1+s)^2}\widetilde{D}'_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right) && \text{(A.29)}\\ \frac{\partial^2 \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s^2} &= -\frac{2}{(1+s)^3}\widetilde{D}'_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right) + \frac{s}{(1+s)^4}\widetilde{D}''_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right) && \text{(A.30)}\\ \frac{\partial^3 \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s^3} &= \frac{6}{(1+s)^4}\widetilde{D}'_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right) + \frac{3-3s}{(1+s)^5}\widetilde{D}''_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right)\\ &\quad - \frac{s}{(1+s)^6}\widetilde{D}'''_{\frac{1}{1+s}}\left(\mathcal{W}\|\sigma|P\right). && \text{(A.31)}\end{aligned}$$

From Eqs. (A.29)-(A.31) and the fact that $\widetilde{D}_{1/(1+s)}\left(\mathcal{W}\|\sigma|P\right)$, $\widetilde{D}'_{1/(1+s)}\left(\mathcal{W}\|\sigma|P\right)$, $\widetilde{D}''_{1/(1+s)}\left(\mathcal{W}\|\sigma|P\right)$, and $D'''_{1/(1+s)}\left(\mathcal{W}\|\sigma|P\right)$ are continuous for $(s,P) \in \mathbb{R}_{\geq 0} \times \mathcal{P}(\mathcal{X})$, we deduce the continuity property in item (a).

(3-(b)) The proof strategy follows closely with [48, Appendix B]. Let

$$\tilde{\psi}(\alpha) = \sum_{x\in\mathcal{X}} P(x) \log \mathrm{Tr}\left[\mathrm{e}^{\alpha \log W_x + (1-\alpha)\log\sigma}\right]. \tag{A.32}$$

Since $\alpha \mapsto \tilde{\psi}(\alpha)$ is convex for all $\alpha \in (0,1]$ [48, Lemma III.11], it can be written as the supremum of affine functions, i.e.

$$\tilde{\psi}(\alpha) = \sup_{i\in\mathcal{I}} \{c_i\alpha + d_i\} \tag{A.33}$$

for some index set $\mathcal{I}$. Hence,

$$-\widetilde{E}_{\mathrm{h}}(s,P,\sigma) = (1+s)\tilde{\psi}\left(\frac{1}{1+s}\right) = \sup_{i\in\mathcal{I}}\{c_i + d_i(1+s)\}. \tag{A.34}$$

The right-hand side of Eq. (A.34), in turn, implies that the map $s \mapsto \widetilde{E}_{\mathrm{h}}(s,P,\sigma)$ is convex for all $s \in \mathbb{R}_{\geq 0}$.

(3-(c)) From Eqs. (A.29) and (A.30) and recalling [48, Lemma III.4], one finds

$$\left.\frac{\partial \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s}\right|_{s=0} = D\left(\mathcal{W}\|\sigma|P\right). \tag{A.35}$$

(3-(d)) The concavity of the map $s \mapsto E_{\mathrm{h}}(s,P)$ in item (b) ensures that $\partial E_{\mathrm{h}}(s,P)/\partial s$ is non-increasing in s. Along with Eq. (A.35) in item (c), we conclude Eq. (2.36).

(3-(e)) Following similar steps in [45, Proposition 4], it can be verifies that

$$\widetilde{D}'_{\alpha}(\rho\|\sigma)\Big|_{\alpha=1} = \lim_{\alpha\uparrow 1} \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2}\log f(\alpha) = \frac{f(1)f''(1)-(f'(1))^2}{2(f(1))^2}, \tag{A.36}$$

where $f(\alpha) := \operatorname{Tr}\left[\mathrm{e}^{\alpha\log\rho+(1-\alpha)\sigma}\right]$. Further, the Fréchet derivative of the exponential (see e.g. [49, Example X.4.2]) gives

$$f'(\alpha) = \operatorname{Tr}\left[\mathrm{e}^{\alpha\log\rho+(1-\alpha)\sigma}\left(\log\rho-\log\sigma\right)\right], \tag{A.37}$$

$$f''(\alpha) = \int_0^1 \mathrm{d}t\, \operatorname{Tr}\left[\mathrm{e}^{t(\alpha\log\rho+(1-\alpha)\sigma)}\left(\log\rho-\log\sigma\right)\mathrm{e}^{(1-t)(\alpha\log\rho+(1-\alpha)\sigma)}\left(\log\rho-\log\sigma\right)\right], \tag{A.38}$$

Therefore, Eq. (A.36) equals

$$\widetilde{D}'_{\alpha}(\rho\|\sigma)\Big|_{\alpha=1} = \frac{1}{2}\left(\int_0^1 \mathrm{d}t\, \operatorname{Tr}\left[\rho^{1-t}(\log\rho-\log\sigma)\rho^{t}(\log\rho-\log\sigma)\right] - D(\rho\|\sigma)^2\right) \tag{A.39}$$

$$= \frac{1}{2}\widetilde{V}(\rho\|\sigma). \tag{A.40}$$

Finally, combining with Eq. (A.30) yields

$$\left.\frac{\partial^2 \widetilde{E}_{\mathrm{h}}(s,P,\sigma)}{\partial s^2}\right|_{s=0} = -\widetilde{V}\left(\mathcal{W}\|\sigma|P\right). \tag{A.41}$$

□

Appendix B. A Weak Converse Bound: Proof of Proposition 6

Proposition 6 (Weak Converse Bound with Polynomial Prefactors)**.** *Consider a classical-quantum channel $\mathcal{W}:\mathcal{X}\to\mathcal{S}(\mathcal{H})$ with $\mathcal{S}_\circ := \overline{\mathsf{im}(\mathcal{W})}$, an arbitrary rate $R\geq 0$, and $\sigma\in\mathcal{S}_{>0}(\mathcal{H})$. For any $\eta\in(0,\frac{1}{2})$ and $c>0$, let $N_0\in\mathbb{N}$ such that for all $n\geq N_0$,*

$$c\cdot\mathrm{e}^{-\xi\sqrt{n}} \leq \frac{\eta}{2}, \tag{B.1}$$

where $\xi=\sqrt{2A/\eta}$ and $A:=\max_{\rho\in\mathcal{S}_\circ} V(\rho\|\sigma)$. Then, it holds that for all $n\geq N_0$,

$$\widehat{\alpha}_{c\exp\{-nR\}}\left(W_{\mathbf{x}^n}^{\otimes n}\|\sigma^{\otimes n}\right) \geq f(\eta)\exp\left\{-n\left[\frac{\widetilde{E}_{\mathrm{sp}}\left(R-\frac{2\xi}{\sqrt{n}},P_{\mathbf{x}^n},\sigma\right)}{1-\eta}\right]\right\}, \tag{B.2}$$

where $f(\eta)=\exp\left\{-\frac{h(1-\eta)}{1-\eta}\right\}$ and $h(p):=-p\log p-(1-p)\log(1-p)$ is the binary entropy function.

Remark B.1. Consider a constant composition code with common type $P_{\mathbf{x}^n}$ on a finite input alphabet $\mathcal{X}$. Recall the definition of the weak sphere-packing exponent [40, 18]:

$$\widetilde{E}_{\mathrm{sp}}(R,P_{\mathbf{x}^n}) := \min_{\bar{\mathcal{W}}:\mathcal{X}\to\mathcal{S}(\mathcal{H})}\left\{D\left(\bar{\mathcal{W}}\|\mathcal{W}|P_{\mathbf{x}^n}\right) : I(P_{\mathbf{x}^n},\bar{\mathcal{W}})\leq R\right\}. \tag{B.3}$$

Proposition 6, along with [18, Lemma 11], establishes a weak sphere-packing bound with polynomial prefactors, which generalizes Altuğ and Wagner's result [21, Lemma 3] to c-q channels: for any $\eta\in(0,\frac{1}{2})$ and for all sufficiently large n such that Eq. (B.1) holds, we have

$$\epsilon_{\max}(\mathcal{W},P_{\mathbf{x}^n}) \geq \max_{\sigma\in\mathcal{S}(\mathcal{H})}\widehat{\alpha}_{\exp\{-nR\}}\left(W_{\mathbf{x}^n}^{\otimes n}\|\sigma^{\otimes n}\right) \tag{B.4}$$

$$\geq \widehat{\alpha}_{\exp\{-nR\}}\left(W_{\mathbf{x}^n}^{\otimes n}\|(\sigma^\star)^{\otimes n}\right) \tag{B.5}$$

$$\geq f(\eta)\exp\left\{-n\left[\frac{\widetilde{E}_{\mathrm{sp}}\left(R-\frac{2\xi}{\sqrt{n}},P_{\mathbf{x}^n}\right)}{1-\eta}\right]\right\}, \tag{B.6}$$

where $\sigma^\star := P_{\mathbf{x}^n}\bar{\mathcal{W}}^\star$ and $\bar{\mathcal{W}}^\star$ is an arbitrary minimizer in Eq. (B.3). Moreover, Eq. (B.6) improves the prefactor of Winter's weak sphere-packing bound [40] from the order of subexponential to polynomial.

Proof of Proposition 6. Consider an arbitrary sequence $\mathbf{x}^n \in \mathcal{X}^n$ and a test Q_n on $\mathcal{H}^{\otimes n}$. For two c-q channels $\bar{\mathcal{W}}, \mathcal{W} : \mathcal{X} \to \mathcal{S}_\circ$, the data-processing inequality implies that

$$D\left(\bar{W}_{\mathbf{x}^n}^{\otimes n} \| W_{\mathbf{x}^n}^{\otimes n}\right) \geq \left[1 - \alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n})\right] \log \frac{1 - \alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n})}{1 - \alpha(Q_n; W_{\mathbf{x}^n}^{\otimes n})} + \alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n}) \log \frac{\alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n})}{\alpha(Q_n; W_{\mathbf{x}^n}^{\otimes n})} \tag{B.7}$$

$$= -h\left(\alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n})\right) - \alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n}) \log \alpha(Q_n; W_{\mathbf{x}^n}^{\otimes n})$$

$$- \left[1 - \alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n})\right] \log\left(1 - \alpha(Q_n; W_{\mathbf{x}^n}^{\otimes n})\right) \tag{B.8}$$

$$\geq -\alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n}) \log \alpha(Q_n; W_{\mathbf{x}^n}^{\otimes n}) - h\left(\alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n})\right), \tag{B.9}$$

where the last inequality (B.9) follows since the third term in (B.8) is non-negative. Continuing from Eq. (B.9), we have

$$\alpha(Q_n; W_{\mathbf{x}^n}^{\otimes n}) \geq \exp\left\{-\frac{D\left(\bar{W}_{\mathbf{x}^n}^{\otimes n} \middle\| W_{\mathbf{x}^n}^{\otimes n}\right) + h\left(\alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n})\right)}{\alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n})}\right\} \tag{B.10}$$

$$= \exp\left\{-\frac{n\, D\left(\bar{\mathcal{W}} \middle\| \mathcal{W} \middle| P_{\mathbf{x}^n}\right) + h\left(\alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n})\right)}{\alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n})}\right\}, \tag{B.11}$$

where Eq. (B.11) follows from the additivity of the relative entropy and the empirical distribution $P_{\mathbf{x}^n}$.

The next step is to replace $\alpha(Q_n; W_{\mathbf{x}^n}^{\otimes n})$ with a lower bound that does not depend on the dummy channel $\bar{W}$, provided that $\bar{W}$ satisfies certain conditions. This can be done using Proposition 13, Wolfowitz's strong converse bound. We delay its proof in Appendix B.1.

Proposition 13 (Wolfowitz's Strong Converse)**.** *Let $\mathcal{S}_\circ \subseteq \mathcal{S}(\mathcal{H})$ be closed and let $\bar{\mathcal{W}} : \mathcal{X} \to \mathcal{S}_\circ$ be an arbitrary classical-quantum channel. Consider the binary hypothesis testing:*

$$\mathsf{H}_0 : \bar{W}_{\mathbf{x}^n}^{\otimes n}, \tag{B.12}$$

$$\mathsf{H}_1 : \sigma^{\otimes n}, \tag{B.13}$$

where $\mathbf{x}^n \in \mathcal{X}^n$ and $\sigma \in \mathcal{S}_{>0}(\mathcal{H})$. For any test Q_n such that $\beta(Q_n; \sigma^{\otimes n}) \leq \mathrm{e}^{-nR}$ and $D\left(\bar{W}_{\mathbf{x}^n} \| \sigma | P_{\mathbf{x}^n}\right) \leq R - 2\kappa$, it holds that

$$\alpha\left(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n}\right) > 1 - \frac{A}{n\kappa^2} - \mathrm{e}^{-n\kappa}, \tag{B.14}$$

where $A := \max_{\rho \in \mathcal{S}_\circ} V\left(\rho \| \sigma\right)$.

Fix $0 < \eta < \frac{1}{2}$, and let $\xi^2 := \frac{2A}{\eta}$. Note that ξ^2 is finite because $A < +\infty$. For all $n \geq N_0$, we have

$$c \cdot \mathrm{e}^{-\xi\sqrt{n}} \leq \frac{\eta}{2} \tag{B.15}$$

by assumption in Proposition 6. Choose $\kappa = \xi/\sqrt{n}$. For any $\bar{\mathcal{W}} : \mathcal{X} \to \mathcal{S}_\circ$ with $D\left(\bar{\mathcal{W}} \| \sigma | P_{\mathbf{x}^n}\right) \leq R - \frac{2\xi}{\sqrt{n}}$ and any test Q_n such that $\beta(Q_n; \sigma^{\otimes n}) \leq \mathrm{e}^{-nR}$, Proposition 13 gives a lower bound to the type-I error:

$$\alpha(Q_n; \bar{W}_{\mathbf{x}^n}^{\otimes n}) \geq 1 - \frac{A}{n\kappa^2} - \mathrm{e}^{-n\kappa} \geq 1 - \eta. \tag{B.16}$$

Hence, combining Eqs. (B.11) and (B.16) yields that, for any $\beta(Q_n; \sigma^{\otimes n}) \leq c\mathrm{e}^{-nR}$,

$$\alpha(Q_n; W_{\mathbf{x}^n}^{\otimes n}) \geq \max_{\bar{\mathcal{W}} : D\left(\bar{\mathcal{W}} \| \sigma | P_{\mathbf{x}^n}\right) \leq R - \frac{2\xi}{\sqrt{n}}} \exp\left\{-\frac{n\, D\left(\bar{\mathcal{W}} \middle\| \mathcal{W} \middle| P_{\mathbf{x}^n}\right) + h\left(1 - \eta\right)}{1 - \eta}\right\}, \tag{B.17}$$

$$= \exp\left\{-\frac{h\left(1 - \eta\right)}{1 - \eta}\right\} \exp\left\{-\frac{n\, \widetilde{E}_{\mathrm{sp}}\left(R - \frac{2\xi}{\sqrt{n}}, P_{\mathbf{x}^n}, \sigma\right)}{1 - \eta}\right\}, \tag{B.18}$$

which concludes Proposition 6.

□

B.1. **Proof of Wolfowitz's Strong Converse: Proposition 13.** To prove our claim, we first introduce notation for generalized divergences. For any $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, and $\gamma > 0$, define the *hockey-stick divergence* by

$$\mathcal{D}_\gamma(\rho \| \sigma) := \operatorname{Tr}\left[(\rho - \gamma\sigma)_+\right], \tag{B.19}$$

where $A_+ := A\{A \geq 0\}$ denotes the self-adjoint matrix contributed only by its positive part. This divergence satisfies the data-processing inequality (DPI):

$$\operatorname{Tr}\left[(\rho - \gamma\varrho)_+\right] \geq \operatorname{Tr}\left[(\mathcal{N}(\rho) - \gamma\mathcal{N}(\varrho))_+\right], \tag{B.20}$$

for any completely positive and trace-preserving map $\mathcal{N} : \mathcal{S}(\mathcal{H}_{\text{in}}) \to \mathcal{S}(\mathcal{H}_{\text{out}})$ [51, Lemma 4]. Let

$$\rho_p := p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|, \quad \text{and} \quad \sigma_q := q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|, \tag{B.21}$$

for $0 \leq p, q \leq 1$ and some orthonormal basis $\{|0\rangle, |1\rangle\}$, and define

$$\mathrm{d}_\gamma\left(p \| q\right) := \mathcal{D}_\gamma\left(\rho_p \| \sigma_q\right). \tag{B.22}$$

Note that the quantity $\mathrm{d}_\gamma\left(p\|q\right)$ is independent of the choice of the basis $\{|0\rangle, |1\rangle\}$. Now we are ready to prove Proposition 13.

Proof of Proposition 13. Fix an arbitrary test Q_n on $\mathcal{H}^{\otimes n}$. For notational convenience, we shorthand $\rho^n = \bar{W}_{\mathbf{x}^n}^{\otimes n}$, $\tau^n = \sigma^{\otimes n}$, $\alpha = \alpha(Q_n; \rho^n)$ and $\beta = \beta = (Q_n; \tau^n)$. Further, we assume $\beta(Q_n; \tau^n) \leq \mathrm{e}^{-nR}$. From the definition of the classical divergence, Eqs. (B.19) and (B.22), and any $\gamma > 0$, we find

$$\begin{aligned} \mathrm{d}_\gamma(1-\alpha\|\beta) &= (1-\alpha-\gamma\beta)_+ + (\alpha - \gamma\left[1-\beta\right])_+ && \text{(B.23)}\\ &\geq 1-\alpha-\gamma\beta && \text{(B.24)}\\ &\geq 1-\alpha-\gamma\mathrm{e}^{-nR}. && \text{(B.25)} \end{aligned}$$

On the other hand, DPI and the measurement map $\operatorname{Tr}[Q_n(\cdot)]|0\rangle\langle 0| + (1-\operatorname{Tr}[Q_n(\cdot)])|1\rangle\langle 1|$ imply that

$$\mathcal{D}_\gamma\left(\rho^n \| \tau^n\right) \geq \mathrm{d}_\gamma\left(\operatorname{Tr}[Q_n\rho^n] \| \operatorname{Tr}[Q_n\tau^n]\right) = \mathrm{d}_\gamma(1-\alpha\|\beta). \tag{B.26}$$

Hence, Eqs. (B.25) and (B.26) lead to

$$\alpha \geq 1 - \mathcal{D}_\gamma\left(\rho^n \| \tau^n\right) - \gamma\mathrm{e}^{-nR}. \tag{B.27}$$

Since

$$\begin{aligned} \mathcal{D}_\gamma\left(\rho^n\|\tau^n\right) &= \operatorname{Tr}\left[\{\rho^n - \gamma\tau^n \geq 0\}\left(\rho^n - \gamma\tau^n\right)\right] && \text{(B.28)}\\ &\leq \operatorname{Tr}\left[\{\rho^n - \gamma\tau^n \geq 0\}\,\rho^n\right], && \text{(B.29)} \end{aligned}$$

continuing from Eq. (B.27) gives

$$\alpha \geq 1 - \operatorname{Tr}\left[\{\rho^n - \gamma\tau^n \geq 0\}\,\rho_n\right] - \gamma\mathrm{e}^{-nR}. \tag{B.30}$$

Next, invoking Lemma 14 below, for all $\log\gamma > D\left(\rho^n\|\tau^n\right)$, we have

$$\begin{aligned} \alpha &\geq 1 - \frac{V\left(\rho^n\|\tau^n\right)}{\left[\log\gamma - D\left(\rho^n\|\tau^n\right)\right]^2} - \gamma\mathrm{e}^{-nR} && \text{(B.31)}\\ &= 1 - \frac{V\left(\bar{\mathcal{W}}\|\sigma|P_{\mathbf{x}^n}\right)}{n\left[\frac{\log\gamma}{n} - D\left(\bar{\mathcal{W}}\|\sigma|P_{\mathbf{x}^n}\right)\right]^2} - \gamma\mathrm{e}^{-nR} && \text{(B.32)} \end{aligned}$$

Finally, recall $D\left(\bar{\mathcal{W}}\|\sigma|P_{\mathbf{x}^n}\right) \leq R - 2\kappa$ and $A := \max_{\rho\in\mathcal{S}_\circ} V\left(\rho\|\sigma\right)$ and choose $\log\gamma = nD\left(\bar{\mathcal{W}}\|\sigma|P_{\mathbf{x}^n}\right) + n\kappa$. Then, Eq. (B.32) yields, for any test Q_n and $\beta(Q_n; \sigma^{\otimes n}) \leq \mathrm{e}^{-nR}$,

$$\begin{aligned} \alpha\left(Q_n; \bar{\mathcal{W}}_{\mathbf{x}^n}^{\otimes n}\right) &\geq 1 - \frac{V\left(\bar{\mathcal{W}}\|\sigma|P_{\mathbf{x}^n}\right)}{n\kappa^2} - \mathrm{e}^{-n\kappa} && \text{(B.33)}\\ &\geq 1 - \frac{A}{n\kappa^2} - \mathrm{e}^{-n\kappa}, && \text{(B.34)} \end{aligned}$$

which concludes the proof.

Lemma 14 (Quantum Chebyshev's Inequality [51, Lemma 6]). *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ and assume $\log\gamma > D(\rho\|\sigma)$. Then*

$$\operatorname{Tr}\left[\rho\left\{\rho - \gamma\sigma \geq 0\right\}\right] \leq \frac{V(\rho\|\sigma)}{\left[\log\gamma - D(\rho\|\sigma)\right]^2}. \tag{B.35}$$

□

Appendix C. A Sharp Converse Bound from Strong Large Deviation

In this section, we provide the proof of Proposition 7. Our technique highly relies on a strong large deviation inequality.

C.1. **A Strong Large Deviation Inequality.** Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of independent, real-valued random variables with probability measures $(\mu_i)_{i=1}^n$. Let $Z_n := \sum_{i=1}^n X_i$ and let $\Lambda_n(t) := \log \mathbb{E}\left[\mathrm{e}^{tZ_n}\right]$. Define the Legendre-Fenchel transform of $\frac{1}{n}\Lambda_n(\cdot)$ by:

$$\Lambda_n^*(z) := \sup_{t\in\mathbb{R}} \left\{ zt - \frac{1}{n}\Lambda_n(t) \right\}, \quad \forall z \in \mathbb{R}. \tag{C.1}$$

Let $(T_n)_{n\in\mathbb{N}}$ be a bounded sequence of real numbers and $(t_n^\star)_{n\in\mathbb{N}}$ be a sequence satisfying for all $n \in \mathbb{N}$

$$t_n^\star \in (0,1); \tag{C.2}$$

$$T_n = \frac{1}{n}\Lambda_n'(t_n^\star); \tag{C.3}$$

$$\Lambda_n^*(T_n) = T_n t_n^\star - \frac{1}{n}\Lambda_n(t_n^\star). \tag{C.4}$$

With these definitions, we can now state the following sharp concentration inequality for $\frac{1}{n}Z_n$:

Theorem 15 (Chaganty-Sethuraman's Concentration Inequality [52, Theorem 3.3]). *For any $\eta \in (0,1)$, there exists an $N_0 \in \mathbb{N}$ such that, for all $n \geq N_0$,*

$$\Pr\left\{ \frac{1}{n}Z_n \geq T_n, \right\} \geq \frac{1-\eta}{t_n^\star\sqrt{2\pi n m_{2,n}}} \exp\{-n\Lambda_n^\star(T_n)\}, \tag{C.5}$$

where $m_{2,n} := \frac{1}{n}\sum_{i=1}^n \operatorname{Var}_{\tilde{\mu}_{n,i}}[X_i]$, and the measure $\tilde{\mu}_{n,i}$ is defined via

$$\frac{\mathrm{d}\tilde{\mu}_{n,i}}{\mathrm{d}\mu_i}(y) := \frac{\mathrm{e}^{yt_n^\star}}{\mathbb{E}\left[\mathrm{e}^{t_n^\star X_i}\right]}. \tag{C.6}$$

Remark C.1. Chaganty and Sethuraman in Ref. [52, Theorem 3.3] considered a more general sequence of random variables $\{Z_n\}_{n\in\mathbb{N}}$, which are not necessarily the sum of random variables. They proved Theorem 15 provided that the following condition is satisfied: there exists $\delta_0 > 0$ such that for any δ and λ with $0 < \delta < \delta_0 < \lambda$, $\sup_{\delta<|t|\leq\lambda t_n^\star} |\Lambda_n(t_n^\star + \mathrm{i}t)/\Lambda_n(t_n^\star)| = o(1/\sqrt{n})$, where the supremum is defined to be 0 if $\{t : \delta < |t| \leq \lambda t_n^\star\}$ is empty. In the case of Z_n being a sum of random variables, $\Lambda_n(t_n^\star + \mathrm{i}t)/\Lambda_n(t_n^\star)$ is the product of the characteristic functions of $\{X_i\}_{i=1}^n$. Since the supremum of a characteristic function on a compact interval not containing 0 is less than 1, this condition is thus satisfied.

We note that the lower bound in Theorem 15 for the general sequence of random variables $(X_i)_{i\in\mathbb{N}}$ suffices to establish the converse, Theorem 5. We do not particularly consider the case of lattice valued random variables (see e.g. [52, Theorem 3.5]).

C.2. **Proof of Proposition 7.**

Proposition 7 (A Sharp Converse Bound). *Consider a classical-quantum channel $\mathcal{W} : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ and a state $\sigma \in \mathcal{S}(\mathcal{H})$. Suppose the sequence $\mathbf{x}^n \in \mathcal{X}^n$ satisfies*

$$\nu \leq V\left(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}\right) < +\infty \tag{C.7}$$

for some $\nu > 0$, and suppose the sequence of rates $(R_n)_{n\in\mathbb{N}}$ satisfies $D_0(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}) < R_n < D(\mathcal{W}\|\sigma|P_{\mathbf{x}^n})$. Then, there exists an $N_0 \in \mathbb{N}$ such that, for all $n \geq N_0$,

$$\widehat{\alpha}_{\exp\{-nR_n\}}(W_{\mathbf{x}^n}^{\otimes n}\|\sigma^{\otimes n}) \geq \frac{A}{s_n^\star\sqrt{n}} \exp\left\{-nE_{\text{sp}}^{(2)}\left(R_n - c_n, P_{\mathbf{x}^n}, \sigma\right)\right\}, \tag{C.8}$$

where $c_n = \frac{K\log n}{n}$ and $A, K > 0$ are finite constants independent of the sequence $\mathbf{x}^n$, and

$$s_n^\star := \arg\max_{s\geq 0}\left\{E_{\text{h}}(s, P_{\mathbf{x}^n}, \sigma) - sR_n\right\}. \tag{C.9}$$

Proof of Proposition 7. Let $\rho^n := W_{\mathbf{x}^n}^{\otimes n}$, $\sigma^n := \sigma^{\otimes n}$, $p^n := \bigotimes_{i=1}^n p_{x_i}$, and $q^n := \bigotimes_{i=1}^n q_{x_i}$, where p_{x_i}, q_{x_i} are Nussbaum-Szkoła distributions [58] of W_{x_i}, σ for every $i \in [n]$. Let $\tilde{R}_n := R_n - \gamma_n$, where $\gamma_n := \frac{\log n}{2n} + \frac{x}{n}$ for some $x \in \mathbb{R}$. The choice of x and the rate back-off term γ_n will become evident later. Let $N_1 \in \mathbb{N}$ such that $\tilde{R}_n \geq D_0(\mathcal{W}\|\sigma|P_{\mathbf{x}^n})$ for all $n \geq N_1$. Subsequently, we choose such $n \geq N_1$ onwards.

Since $D_\alpha(W_{x_i}\|\sigma) = D_\alpha(p_{x_i}\|q_{x_i})$, for $\alpha \in (0,1]$, we use the notation

$$\phi_n(\tilde{R}_n) := E_{\text{sp}}^{(2)}(\tilde{R}_n, P_{\mathbf{x}^n}, \sigma) = \sup_{0<\alpha\leq 1} \frac{1-\alpha}{\alpha}\left(\sum_{x\in\mathcal{X}} P_{\mathbf{x}^n}(x) D_\alpha(p_{x_i}\|q_{x_i}) - \tilde{R}_n\right), \tag{C.10}$$

where $P_{\mathbf{x}^n}$ denotes the empirical distribution of $\mathbf{x}^n = x_1, \ldots x_n$. Moreover, the condition in Eq. (C.7) implies that $W_x \ll \sigma$, for all $x \in \texttt{supp}(P_{\mathbf{x}^n})$, and thus $p^n \ll q^n$. Without loss of generality, we let $q_{x_i}(\omega) = 0$, $\omega \notin \texttt{supp}(p_{x_i})$ since they won't contribute to $\phi_n(\tilde{R}_n)$.

We apply Nagaoka's argument [59]: for any $0 \leq Q_n \leq \mathbb{1}$, choosing $\delta = \exp\{n\tilde{R}_n - n\phi_n(\tilde{R}_n)\}$ yields:

$$\alpha\left(Q_n; \rho^n\right) + \delta\beta\left(Q_n; \sigma^n\right) \geq \frac{1}{2}\left(\alpha\left(\mathcal{U}; p^n\right) + \mathrm{e}^{n\tilde{R}_n - n\phi_n(\tilde{R}_n)}\beta\left(\mathcal{U}; q^n\right)\right), \tag{C.11}$$

where

$$\alpha\left(\mathcal{U}; p^n\right) := \sum_{\omega\in\mathcal{U}^{\text{c}}} p^n(\omega); \quad \beta\left(\mathcal{U}; q^n\right) := \sum_{\omega\in\mathcal{U}} q^n(\omega), \tag{C.12}$$

and

$$\mathcal{U} := \left\{\omega : p^n(\omega)\mathrm{e}^{n\phi_n(\tilde{R}_n)} > q^n(\omega)\mathrm{e}^{n\tilde{R}_n}\right\}. \tag{C.13}$$

In the following, we will employ Theorem 15, to further lower bound $\alpha\left(\mathcal{U}; p^n\right)$ and $\beta\left(\mathcal{U}; q^n\right)$. Before proceeding, we need to introduce some notation. Define the *tilted distributions*, for every $i \in [n]$ and $t \in [0,1]$, to be

$$\hat{q}_{x_i,t}(\omega) := \frac{p_{x_i}(\omega)^{1-t} q_{x_i}(\omega)^t}{\sum_{\omega\in\texttt{supp}(p_{x_i})} p_{x_i}(\omega)^{1-t} q_{x_i}(\omega)^t}, \quad \omega \in \texttt{supp}(p_{x_i}). \tag{C.14}$$

Let

$$\Lambda_{0,x_i}(t) := \log \mathbb{E}_{p_{x_i}}\left[\mathrm{e}^{t\log\frac{q_{x_i}}{p_{x_i}}}\right], \quad \Lambda_{1,x_i}(t) := \log \mathbb{E}_{q_{x_i}}\left[\mathrm{e}^{t\log\frac{p_{x_i}}{q_{x_i}}}\right], \tag{C.15}$$

Since p^n and q^n share the same support, it can be verified that the maps $t \mapsto \Lambda_{j,x_i}(t)$, $j \in \{0,1\}$ are differential for all $t \in [0,1]$. One can immediately verify the following partial derivatives with respect to t:

$$\Lambda'_{0,x_i}(t) = \mathbb{E}_{\hat{q}_{x_i,t}}\left[\log\frac{q_{x_i}}{p_{x_i}}\right], \quad \Lambda'_{1,x_i}(t) = \mathbb{E}_{\hat{q}_{x_i,1-t}}\left[\log\frac{p_{x_i}}{q_{x_i}}\right]; \tag{C.16}$$

$$\Lambda''_{0,x_i}(t) = \mathrm{Var}_{\hat{q}_{x_i,t}}\left[\log\frac{q_{x_i}}{p_{x_i}}\right], \quad \Lambda''_{1,x_i}(t) = \mathrm{Var}_{\hat{q}_{x_i,1-t}}\left[\log\frac{p_{x_i}}{q_{x_i}}\right]. \tag{C.17}$$

Note that Eqs. (C.15), (C.16), and (C.17) ensure that

$$\Lambda_{0,x_i}(t) = \Lambda_{1,x_i}(1-t), \quad \Lambda'_{0,x_i}(t) = -\Lambda'_{1,x_i}(1-t), \quad \Lambda''_{0,x_i}(t) = \Lambda''_{1,x_i}(1-t). \tag{C.18}$$

With $\Lambda_{j,x_i}(t)$ in Eq. (C.15), we can define

$$\Lambda_{j,P_{\mathbf{x}^n}}(t) := \sum_{x\in\mathcal{X}} P_{\mathbf{x}^n}(x)\Lambda_{j,x}(t), \qquad j \in \{0,1\}; \tag{C.19}$$

$$\Lambda^*_{j,P_{\mathbf{x}^n}}(z) := \sup_{t\in\mathbb{R}} \left\{tz - \Lambda_{j,P_{\mathbf{x}^n}}(t)\right\}, \quad j \in \{0,1\}, \tag{C.20}$$

where $\Lambda^*_{j,P_{\mathbf{x}^n}}(z)$ in Eq. (C.20) are the *Legendre-Fenchel transform* of $\Lambda_{j,P_{\mathbf{x}^n}}(t)$. The quantities $\Lambda^*_{j,P_{\mathbf{x}^n}}(z)$ would appear in the lower bounds of $\alpha(\mathcal{U}; p^n)$ and $\beta(\mathcal{U}; q^n)$ obtained by Theorem 15 as shown later.

In the following, we will relate the Legendre-Fenchel transform $\Lambda^*_{j,P_n}(z)$ to the desired error-exponent function $\phi_n(\tilde{R}_n)$. Such a relationship is stated in the following lemma whose proof was presented in [18].

Lemma 16 ([18, Lemma 17])**.** *The following holds for all sequences $\mathbf{x}^n$ satisfying Eq. (C.7) and all $r \in (D_0(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}), D(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}))$:*

(a) $\Lambda''_{0,P_{\mathbf{x}^n}}(t) > 0$ for all $t \in [0,1]$.
(b) $\Lambda^*_{0,P_{\mathbf{x}^n}}(\phi_n(r) - r) = \phi_n(r)$.
(c) $\Lambda^*_{1,P_{\mathbf{x}^n}}(r - \phi_n(r)) = r$.
(d) Let $s^\star$ be the optimizer of $E^{(2)}_{\text{sp}}(r, P_{\mathbf{x}^n}, \sigma)$, c.f. (C.9). The optimizer of $\Lambda^*_{0,P_{\mathbf{x}^n}}(z)$, denoted by $t^\star$, is unique and satisfies $t^\star = \frac{s^\star}{1+s^\star} \in (0,1)$ and $\Lambda'_{0,P_{\mathbf{x}^n}}(t^\star) = \phi_n(r) - r$.

Since the item (d) in Lemma 16 shows that the optimizer t in Eq. (C.20) always lies in the compact set $[0,1]$, by invoking Eq. (C.18) we define the following quantity:

$$V_{\min}(\nu) := \min_{t\in[0,1],\, P_{\mathbf{x}^n}\in\mathcal{P}_\nu(\mathcal{X})} \Lambda''_{0,P_n}(t), \tag{C.21}$$

where $\mathcal{P}_\nu(\mathcal{X}) := \{P_{\mathbf{x}^n} \in \mathcal{P}(\mathcal{X}) : \nu \le V(\mathcal{W}\|\sigma|P_{\mathbf{x}^n}) < +\infty\}$ is a compact set owing to the continuity of the map $P \mapsto V(\mathcal{W}\|\sigma|P)$; see Eq. (2.16).

Further, from the definitions in Eqs, (C.17), $\Lambda''_{0,(\cdot)}(\cdot)$ is continuous functions in $[0,1] \times \mathcal{P}(\mathcal{X})$. The minimization in the above definitions are well-defined and finite. Further, the quantity $V_{\min}(\nu)$ is bounded away from zero owing to item (a) in Lemma 16.

Now, we are ready to derive the lower bounds to $\alpha(\mathcal{U}; p^n)$ and $\beta(\mathcal{U}; q^n)$. Fix an arbitrary $\eta \in (0,1)$. Applying Theorem 15 to $X_i = \log q_i - \log p_i$ with probability measure p_i, and threshold $T_n = \tilde{R}_n - \phi_n(\tilde{R}_n)$ gives, for all sufficiently large n, say $n \ge N_2 \in \mathbb{N}$,

$$\begin{aligned}
\alpha(\mathcal{U}; p^n) &:= \sum_{\omega\in\mathcal{U}^c} p^n(\omega) && \text{(C.22)}\\
&= \Pr\left\{\frac{1}{n}\sum_{i=1}^n Z_i \ge \tilde{R}_n - \phi_n(\tilde{R}_n)\right\} && \text{(C.23)}\\
&\ge \frac{1-\eta}{t^\star_n\sqrt{2\pi n V_{\min}(\nu)}} \exp\left\{-n\Lambda^*_{0,P_{\mathbf{x}^n}}\left(\phi_n(\tilde{R}_n) - \tilde{R}_n\right)\right\}, && \text{(C.24)}
\end{aligned}$$

where

$$t^\star_n := \arg\max_{t\in\mathbb{R}} \left\{tz_n - \Lambda_{0,P_{\mathbf{x}^n}}(t)\right\} \tag{C.25}$$

Similarly, applying again Theorem 15 to $X_i = \log p_i - \log q_i$ with probability measure $= q_i$, and threshold $\phi_n(\tilde{R}_n) - \tilde{R}_n$ yields, for all sufficiently large n, say $n \geq N_3 \in \mathbb{N}$,

$$\beta\left(\mathfrak{U}; q^n\right) := \sum_{\omega \in \mathfrak{U}} q^n(\omega) \tag{C.26}$$

$$= \Pr\left\{ \frac{1}{n}\sum_{i=1}^{n} Z_i \geq \phi_n(\tilde{R}_n) - \tilde{R}_n \right\} \tag{C.27}$$

$$\geq \frac{1-\eta}{(1-t_n^\star)\sqrt{2\pi n V_{\min}(\nu)}} \exp\left\{ -n\Lambda^*_{1,P_{\mathbf{x}^n}}\left(\tilde{R}_n - \phi_n(\tilde{R}_n)\right)\right\} \tag{C.28}$$

$$\geq \frac{1-\eta}{\sqrt{2\pi n V_{\min}(\nu)}} \exp\left\{ -n\Lambda^*_{1,P_{\mathbf{x}^n}}\left(\tilde{R}_n - \phi_n(\tilde{R}_n)\right)\right\}, \tag{C.29}$$

where the term $1-t_n^\star$ in Eq (C.28) comes from the symmetry in Eq. (C.18), and the last inequality (C.29) follows from $t_n^\star \in (0,1)$ in item (d) of Lemma 16.

Continuing from Eq. (C.24) and item (b) in Lemma 16 gives

$$\alpha\left(\mathfrak{U}; p^n\right) \geq \frac{1-\eta}{t_n^\star\sqrt{2\pi n V_{\min}(\nu)}} \exp\{-n\phi_n(\tilde{R}_n)\}. \tag{C.30}$$

Eq. (C.29) together with item (c) in Lemma 16 yields

$$\beta\left(\mathfrak{U}; q^n\right) \geq \frac{1-\eta}{\sqrt{2\pi n V_{\min}(\nu)}} \exp\{-n\tilde{R}_n\} = 2\exp\{-nR_n\}, \tag{C.31}$$

where we choose $x = \log 2\sqrt{2\pi V_{\min}(\nu)} - \log(1-\eta)$ in the rate back-off $\gamma_n = \frac{\log n}{2n} + \frac{x}{n}$. Thus we can bound the left-hand side of Eq. (C.11) from below. If for any test $0 \leq Q_n \leq \mathbb{1}$ such that

$$\beta(Q_n; \sigma^n) \leq \exp\{-nR_n\}, \tag{C.32}$$

holds, it implies that

$$\alpha(Q_n; \rho^n) \geq \frac{1-\eta}{t_n^\star 2\sqrt{2\pi n V_{\min}(\nu)}} \exp\{-n\phi_n(\tilde{R}_n)\}. \tag{C.33}$$

Finally, let $A := (1-\eta)/(2\sqrt{2\pi V_{\min}(\nu)})$ and choose a constant $K > 0$ such that for all $n \geq N_0 := \max\{N_1, N_2, N_3\}$,

$$\gamma_n = \frac{\log n}{2n} + \frac{\log 2\sqrt{2\pi V_{\min}(\nu)} - \log(1-\eta)}{n} \leq \frac{K\log n}{n} =: c_n. \tag{C.34}$$

Since the map $r \mapsto \phi_n(r)$ is monotonically decreasing [30, Section 5], Eqs. (C.32), (C.33), and (C.34) conclude our result: for all $n \geq N_0$,

$$\widehat{\alpha}_{\exp\{-nR\}}\left(\rho^n \| \sigma^n\right) \geq \frac{A}{t_n^\star\sqrt{n}} \exp\left\{ -nE_{\text{sp}}^{(2)}\left(R_n - c_n, P_{\mathbf{x}^n}, \sigma\right)\right\} \tag{C.35}$$

$$\geq \frac{A}{s_n^\star\sqrt{n}} \exp\left\{ -nE_{\text{sp}}^{(2)}\left(R_n - c_n, P_{\mathbf{x}^n}, \sigma\right)\right\}, \tag{C.36}$$

where the last inequality follows from item (d) in Lemma 16: $t_n^\star = s_n^\star/(1+s_n^\star) \in (0,1)$. □

Appendix D. Proof of Proposition 8

Proposition 8 (Error Exponent around Capacity). *Let $(b_n)_{n\in\mathbb{N}}$ be a sequence of real numbers with $\lim_{n\to+\infty} b_n = 0$ and let $(\delta_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers with $\lim_{n\to+\infty}\delta_n = 0$. Suppose the sequence of distributions $(P_n)_{n\in\mathbb{N}}$ satisfies*

$$C_{\mathcal{W}} - \delta_n < D(\mathcal{W}\|P^\star\mathcal{W}|P_n) \leq C_{\mathcal{W}} - b_n. \tag{D.1}$$

The following holds:

$$\limsup_{n\to+\infty} \frac{E_{\rm sp}^{(2)}\left(C_{\mathcal{W}}-\delta_n, P_n, P^\star\mathcal{W}\right)}{\delta_n^2} \leq \limsup_{n\to+\infty} \frac{(\delta_n - b_n)^2}{2V_{\mathcal{W}}\delta_n^2}; \tag{D.2}$$

$$\limsup_{n\to+\infty} \frac{\widetilde{E}_{\rm sp}\left(C_{\mathcal{W}}-\delta_n, P_n, P^\star\mathcal{W}\right)}{\delta_n^2} \leq \limsup_{n\to+\infty} \frac{(\delta_n - b_n)^2}{2\widetilde{V}_{\mathcal{W}}\delta_n^2}; \tag{D.3}$$

$$\limsup_{n\to+\infty} \frac{s_n^\star}{\delta_n} \leq \frac{1}{V_{\mathcal{W}}}, \tag{D.4}$$

where

$$s_n^\star := \arg\max_{s\geq 0} \left\{E_{\rm h}(s, P_n, P^\star\mathcal{W}) - s\left(C_{\mathcal{W}} - \delta_n\right)\right\}. \tag{D.5}$$

Proof of Proposition 8. We only prove Eqs. (D.2) and (D.4), since Eq. (D.3) follows from the same argument and Proposition 3.

Recall the error-exponent function $E_{\rm sp}^{(2)}$:

$$E_{\rm sp}^{(2)}\left(C_{\mathcal{W}}-\delta_n, P, P^\star\mathcal{W}\right) = \sup_{s\geq 0}\left\{-s\left(C_{\mathcal{W}}-\delta_n\right) + E_{\rm h}(s, P, P^\star\mathcal{W})\right\}. \tag{D.6}$$

In the following, we fix $\sigma = P^\star\mathcal{W}$ in the definition of $E_{\rm h}$ (Eq. (2.26)) and denote by

$$E_{\rm h}(s, P) := E_{\rm h}(s, P, P^\star\mathcal{W}) = sD_{\frac{1}{1+s}}\left(\mathcal{W}\|P^\star\mathcal{W}|P\right). \tag{D.7}$$

for notational convenience. We define a *critical rate* for a c-q channel $\mathcal{W}$ to be

$$r_{\rm cr} := \max_{P\in\mathcal{P}(\mathcal{X})} \left.\frac{\partial E_{\rm h}(s, P)}{\partial s}\right|_{s=1}. \tag{D.8}$$

Let N_0 be the smallest integer such that $C_{\mathcal{W}} - \delta_n > r_{\rm cr}$, $\forall n \geq N_0$. Since the map $r \mapsto E_{\rm sp}^{(2)}(r, \cdot, \cdot)$ is non-increasing [30, Section 5], the maximization over s in Eq. (D.6) can be restricted to the set $[0, 1]$ for any rate above $r_{\rm cr}$, i.e.,

$$E_{\rm sp}^{(2)}\left(C_{\mathcal{W}}-\delta_n, P_n, P^\star\mathcal{W}\right) = \max_{0\leq s\leq 1}\left\{-s\left(C_{\mathcal{W}}-\delta_n\right) + E_{\rm h}(s, P_n)\right\}. \tag{D.9}$$

For every $n\in\mathbb{N}$, let $s_n^\star$ attain the maxima in Eq. (D.9) at a rate of $C_{\mathcal{W}} - \delta_n \geq 0$. In the following lemma, we discuss the asymptotic behavior of $\{s_n^\star\}_{n\in\mathbb{N}}$.

Lemma 17. *Let $s_n^\star$ attain the maxima in Eq. (D.9) and P_n satisfy Eq. (D.1). We have*

(a) The limit point of $\{P_n\}_{n\in\mathbb{N}}$ is capacity achieving.

(b) $s_n^\star > 0$ for all $n \in \mathbb{N}$ and $\lim_{n\to+\infty} s_n^\star = 0$.

Proof of Lemma 17. Let $\{P_{n_k}\}_{k\geq 1}$ and $\{s_{n_k}^\star\}_{k\geq 1}$ be arbitrary subsequences. Since $\mathcal{P}(\mathcal{X})$ and $[0, 1]$ are compact, we may assume that

$$\lim_{k\to+\infty} P_{n_k} = P_o, \quad \lim_{k\to\infty} s_{n_k}^\star = s_o, \tag{D.10}$$

for some $P_o \in \mathcal{P}(\mathcal{X})$ and $s_o \in [0, 1]$.

(17-(a)) Let $k \to +\infty$. Eq. (D.1) implies that

$$D(\mathcal{W}\|P^\star\mathcal{W}|P_o) = C_{\mathcal{W}}, \tag{D.11}$$

which guarantees that P_o is capacity-achieving by the dual representation of the information radius, see e.g. [61], [9, Theorem 2].

(17-(b)) One can observe from Eq. (D.9) that $s_n^\star = 0$ if and only if $C_{\mathcal{W}} - \delta_n \geq D(\mathcal{W}\|P^\star\mathcal{W}|P_n)$. However, this violates the assumption in Eq. (D.1). Hence, we have $s_n^\star > 0$ for all $n \in \mathbb{N}$.

Since P_o is capacity achieving, the uniqueness of the divergence center implies that $P_o\mathcal{W} = P^\star\mathcal{W}$. Item (c) in Proposition 2 shows that

$$\left.\frac{\partial^2 E_{\rm h}\left(s, P_o\right)}{\partial s^2}\right|_{s=0} = -V\left(\mathcal{W}\|P^\star\mathcal{W}|P_o\right) = -V(P_o, \mathcal{W}) \leq -V_{\mathcal{W}} < 0, \tag{D.12}$$

where the last inequality follows since $V_{\mathcal{W}} > 0$. Then, Eq. (D.12) implies that the first-order derivative $\partial E_{\mathrm{h}}(s, P_o)/\partial s$ is strictly decreasing around $s = 0$. Moreover, item (d) in Proposition 2 gives

$$\left.\frac{\partial E_{\mathrm{h}}(s, P_o)}{\partial s}\right|_{s=s_o} \leq D(\mathcal{W}\|P^\star\mathcal{W}|P_o) = C_{\mathcal{W}}, \tag{D.13}$$

This, together with items (b) and (c) in Proposition 2, shows that the first inequality in Eq. (D.13) becomes an equality if and only if $s_o = 0$. Since the subsequence was arbitrary, item (b) is shown.

□

Now we are ready to prove this proposition. We start with proving Eq. (D.4). Since $s \mapsto E_{\mathrm{h}}(s, \cdot)$ is concave from item (b) in Proposition 2, the maximizer $s_n^\star$ must satisfy

$$\left.\frac{\partial E_{\mathrm{h}}(s, P_{n_k})}{\partial s}\right|_{s=s_{n_k}^\star} = C_{\mathcal{W}} - \delta_{n_k}. \tag{D.14}$$

Further, item (c) in Proposition 2 gives

$$\left.\frac{\partial E_{\mathrm{h}}\left(s, P_{n_k}^\star\right)}{\partial s}\right|_{s=0} = D\left(\mathcal{W}\|P^\star\mathcal{W}|P_{n_k}^\star\right). \tag{D.15}$$

The mean value theorem states that there exists a number $\hat{s}_{n_k} \in \left(0, s_{n_k}^\star\right)$, for each $k \geq \mathbb{N}$, such that

$$-\left.\frac{\partial^2 E_{\mathrm{h}}(s, P_{n_k})}{\partial s^2}\right|_{s=\hat{s}_{n_k}} = \frac{D(\mathcal{W}\|P^\star\mathcal{W}|P_{n_k}) - C_{\mathcal{W}} + \delta_{n_k}}{s_{n_k}^\star} \tag{D.16}$$

$$\leq \frac{\delta_{n_k}}{s_{n_k}^\star}, \tag{D.17}$$

where the last inequality is again due to $D\left(\mathcal{W}\|P^\star\mathcal{W}|P_{n_k}^\star\right) \leq C_{\mathcal{W}}$. When k approaches infinity, items (a) and (e) in Proposition 2 give

$$\lim_{k\to+\infty} \left.\frac{\partial^2 E_{\mathrm{h}}(s, P_{n_k})}{\partial s^2}\right|_{s=\hat{s}_{n_k}} = \left.\frac{\partial^2 E_{\mathrm{h}}(s, P_o)}{\partial s^2}\right|_{s=0} = -V(P_o, \mathcal{W}) \leq -V_{\mathcal{W}}. \tag{D.18}$$

Combining Eqs. (D.17) and (D.18) leads to

$$\limsup_{k\to+\infty} \frac{s_{n_k}^\star}{\delta_{n_k}} \leq \frac{1}{V_{\mathcal{W}}}. \tag{D.19}$$

Since the subsequence was arbitrary, the above result establishes Eq. (D.4).

Next, for any sufficiently large $n \geq N_0$, we apply Taylor's theorem to the map $s_n^\star \mapsto E_{\mathrm{h}}(s_n^\star, P_n)$ at the original point to obtain

$$\begin{aligned}
&E_{\mathrm{sp}}^{(2)}(C_{\mathcal{W}} - \delta_n, P_n, P^\star\mathcal{W}) \\
&= -s_n^\star(C_{\mathcal{W}} - \delta_n) + E_{\mathrm{h}}(s_n^\star, P_n) && \text{(D.20)} \\
&= s_n^\star(\delta_n + D(\mathcal{W}\|P^\star W|P_n) - C_{\mathcal{W}}) - \frac{(s_n^\star)^2}{2} V(P_n, \mathcal{W}) + \frac{(s_n^\star)^3}{6} \left.\frac{\partial^3 E_{\mathrm{h}}(s, P_n)}{\partial s^3}\right|_{s=\bar{s}_n} && \text{(D.21)}
\end{aligned}$$

for some $\bar{s}_n \in [0, s_n^\star]$. Let

$$\Upsilon = \max_{(s,P)\in[0,1]\times\mathcal{P}(\mathcal{X})} \left|\frac{\partial^3 E_{\mathrm{h}}(s, P)}{\partial s^3}\right|. \tag{D.22}$$

Continuing from Eq. (D.21) gives

$$E_{\text{sp}}^{(2)}\left(C_{\mathcal{W}}-\delta_n, P_n, P^{\star}\mathcal{W}\right) \leq s_n^{\star}(\delta_n - b_n) - \frac{(s_n^{\star})^2}{2} V\left(P_n, \mathcal{W}\right) + \frac{(s_n^{\star})^3 \Upsilon}{6} \tag{D.23}$$

$$\leq \sup_{s\geq 0}\left\{ s(\delta_n - b_n) - \frac{s^2}{2} V(P_n, \mathcal{W})\right\} + \frac{(s_n^{\star})^3 \Upsilon}{6} \tag{D.24}$$

$$= \frac{(\delta_n - b_n)^2}{2V(P_n, \mathcal{W})} + \frac{(s_n^{\star})^3 \Upsilon}{6}, \tag{D.25}$$

where the first line follows from the assumption $D\left(\mathcal{W}\|P^{\star}\mathcal{W}|P_n\right) \leq C_{\mathcal{W}} - b_n$ in Eq. (D.1) and Eq. (D.22). Finally, Eq. (D.25), along with item (b) in Lemma 17 and Eq. (D.19), implies that

$$\limsup_{n\to+\infty} \frac{E_{\text{sp}}^{(2)}\left(C_{\mathcal{W}}-\delta_n, P_n, P^{\star}\mathcal{W}\right)}{\delta_n^2} \leq \limsup_{n\to+\infty} \frac{(\delta_n - b_n)^2}{2V(P_n, \mathcal{W})\delta_n^2} \tag{D.26}$$

$$\leq \limsup_{n\to+\infty} \frac{(\delta_n - b_n)^2}{2V_{\mathcal{W}}\delta_n^2}, \tag{D.27}$$

where the last inequality follows from the continuity of $V(\cdot, \mathcal{W})$ on $\mathcal{P}(\mathcal{X})$ (Eq. (2.20)); the fact that $\{P_n\}_{n\in\mathbb{N}}$ is capacity achieving (item (a) in Lemma 17); and the definition of $V_{\mathcal{W}}$ in Eq. (2.22). □

References

[1] V. Strassen, "Asymptotische abschätzungen in Shannon's informationstheorie," *Transactions of the Third Prague Conference on Information Theory*, pp. 689–723, 1962.

[2] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[3] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[4] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 4947–4966, Nov 2009.

[5] M. Tomamichel and M. Hayashi, "A hierarchy of information quantities for finite block length analysis of quantum tasks," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7693–7710, Nov 2013.

[6] K. Li, "Second-order asymptotics for quantum hypothesis testing," *The Annals of Statistics*, vol. 42, no. 1, pp. 171–189, Feb 2014.

[7] M. Tomamichel and V. Y. F. Tan, "Second-order asymptotics for the classical capacity of image-additive quantum channels," *Communications in Mathematical Physics*, vol. 338, no. 1, pp. 103–137, May 2015.

[8] M. Tomamichel, M. Berta, and J. M. Renes, "Quantum coding with finite resources," *Nature Communications*, vol. 7, p. 11419, May 2016.

[9] M. Tomamichel and V. Y. F. Tan, "A tight upper bound for the third-order asymptotics for most discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7041–7051, Nov 2013.

[10] V. Y. F. Tan, "Asymptotic estimates in information theory with non-vanishing error probabilities," *Foundations and Trends® in Communications and Information Theory*, vol. 10, no. 4, pp. 1–184, 2014.

[11] V. Y. F. Tan and M. Tomamichel, "The third-order term in the normal approximation for the AWGN channel," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2430–2438, May 2015.

[12] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell System Technical Journal*, vol. 38, no. 3, pp. 611–656, May 1959.

[13] R. Gallager, *Information Theory and Reliable Communication.* Wiley, 1968.

[14] R. M. Fano, *Transmission of Information, A Statistical Theory of Communications.* The MIT Press, 1961.

[15] R. E. Blahut, *Principles and practice of information theory.* Addison-Wesley, 1987.

[16] E. A. Haroutunian, M. E. Haroutunian, and A. N. Harutyunyan, “Reliability criteria in information theory and in statistical hypothesis testing,” *Foundations and Trends® in Communications and Information Theory*, vol. 4, no. 2–3, pp. 97–263, 2007.

[17] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Cambridge University Press (CUP), 2011.

[18] H.-C. Cheng, M.-H. Hsieh, and M. Tomamichel, “Quantum Sphere-Packing Bounds with Polynomial Prefactors,” `arXiv:1704.05703 [quant-ph]`.

[19] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications.* Springer, 1998.

[20] Y. Altuğ and A. B. Wagner, “Moderate deviation analysis of channel coding: Discrete memoryless case,” in *2010 IEEE International Symposium on Information Theory.* Jun 2010.

[21] ——, “Moderate deviations in channel coding,” *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4417–4426, Aug 2014.

[22] Y. Polyanskiy and S. Verdu, “Channel dispersion and moderate deviations limits for memoryless channels,” in *2010 48th Annual Allerton Conference on Communication, Control, and Computing.* Sep 2010.

[23] A. Winter, “Coding theorems of quantum information theory,” *PhD Thesis, Universität Bielefeld,* 1999.

[24] I. Sason, “Moderate deviations analysis of binary hypothesis testing,” in *2012 IEEE International Symposium on Information Theory Proceedings.* Jul 2012.

[25] S. Watanabe and M. Hayashi, “Finite-length analysis on tail probability for Markov chain and application to simple hypothesis testing,” `arXiv:1401.3801`.

[26] C. Rouzé and N. Datta, “Finite blocklength and moderate deviation analysis of hypothesis testing of correlated quantum states and application to classical-quantum channels with memory,” `arXiv:1612.01464`.

[27] M. Hayashi, “Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding,” *Physical Review A*, vol. 76, no. 6, dec 2007.

[28] H.-C. Cheng, M.-H. Hsieh and M. Tomamichel, “Sphere-Packing Bound for Symmetric Classical-Quantum Channels,” `arXiv:1701.02957`.

[29] M. Dalai and A. Winter, “Constant compositions in the sphere packing bound for classical-quantum channels,” in *2014 IEEE International Symposium on Information Theory.* Jun 2014.

[30] K. M. R. Audenaert, M. Nussbaum, A. Szkoła, and F. Verstraete, “Asymptotic error rates in quantum hypothesis testing,” *Communications in Mathematical Physics*, vol. 279, no. 1, pp. 251–283, Feb 2008.

[31] C. T. Chubb, V. Y. F. Tan, and M. Tomamichel, “Moderate deviation analysis for classical communication over quantum channels,” `arXiv:1701.03114 [quant-ph]`.

[32] L. Wang and R. Renner, “One-shot classical-quantum capacity and hypothesis testing,” *Physical Review Letters*, vol. 108, no. 20, May 2012.

[33] L. V. Rozovsky, “Estimate from below for large-deviation probabilities of a sum of independent random variables with finite variances,” *Journal of Mathematical Sciences*, vol. 109, pp. 2192–2209, 2002.

[34] H. Nagaoka and M. Hayashi, “An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses,” *IEEE Transactions on Information Theory*, vol. 53, no. 2, pp. 534–549, Feb 2007.

[35] M. V. Burnashev and A. S. Holevo, “On the reliability function for a quantum communication channel,” *Problems of information transmission*, vol. 34, no. 2, pp. 97–107, 1998.

[36] A. Holevo, “Reliability function of general classical-quantum channel,” *IEEE Transaction on Information Theory*, vol. 46, no. 6, pp. 2256–2261, 2000.

[37] J. I. Fujii, R. Nakamoto, and K. Yanagi,“Remarks on concavity of the auxiliary function appearing in quantum reliability function in classical-quantum channels,” in *2005 IEEE International Symposium on Information Theory.* Jun 2005.

[38] J. I. Fujii, R. Nakamoto, and K. Yanagi, “Concavity of the auxiliary function appearing in quantum reliability function,” *IEEE Transaction on Information Theory*, vol. 52, no. 7, pp. 3310–3313, 2006.

[39] H.-C. Cheng and M.-H. Hsieh, “Concavity of the auxiliary function for classical-quantum channels,” *IEEE Transactions on Information Theory*, vol. 62, no. 10, pp. 5960 – 5965, 2016.

[40] A. Winter, “Coding theorems of quantum information theory,” 1999, (PhD Thesis, Universität Bielefeld). arXiv:quant-ph/9907077.

[41] C. Shannon, R. Gallager, and E. Berlekamp, “Lower bounds to error probability for coding on discrete memoryless channels. I,” *Information and Control*, vol. 10, no. 1, pp. 65–103, Jan 1967.

[42] J.-C. Bourin, “Matrix versions of some classical inequalities,” Linear Algebra and its Applications, vol. 446, no. 2–3, pp. 890–907, July 2006.

[43] D. Petz, “Quasi-entropies for finite quantum systems,” *Reports on Mathematical Physics*, vol. 23, no. 1, pp. 57–65, Feb 1986.

[44] M. Mosonyi and T. Ogawa, “Two approaches to obtain the strong converse exponent of quantum hypothesis testing for general sequences of quantum states,” *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6975–6994, Dec 2015.

[45] S. M. Lin and M. Tomamichel, “Investigating properties of a family of quantum rényi divergences,” *Quantum Information Processing*, vol. 14, no. 4, pp. 1501–1512, Feb 2015.

[46] M. Tomamichel, *Quantum Information Processing with Finite Resources.* Springer International Publishing, 2016.

[47] N. J. Higham, *Functions of Matrices: Theory and Computation.* SIAM, Jan 2008.

[48] M. Mosonyi and T. Ogawa, “Strong converse exponent for classical-quantum channel coding,” 2014 arXiv:1409.3562 [quant-ph].

[49] R. Bhatia, *Matrix Analysis*, Springer New York, 1997.

[50] R. E. Blahut, “Hypothesis testing and information theory,” *IEEE Transaction on Information Theory*, vol. 20, no. 4, pp. 405–417, Jul 1974.

[51] N. Sharma and N. A. Warsi, “Fundamental bound on the reliability of quantum information transmission,” *Physical Review Letters*, vol. 110, no. 8, Feb 2013.

[52] N. R. Chaganty and J. Sethuraman, “Strong Large Deviation and Local Limit Theorems,” *The Annals of Probability*, vol. 21, no. 3, pp. 1671–1690, 1993.

[53] R. Sibson, “Information radius,” *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 14, no. 2, pp. 149–160, 1969.

[54] I. Csiszár, “Generalized cutoff rates and Rényi’s information measures,” *IEEE Transactions on Information Theory*, vol. 41, no. 1, pp. 26–34, 1995.

[55] M. Mosonyi and F. Hiai, “On the quantum Rényi relative entropies and related capacity formulas,” *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2474–2487, Apr 2011.

[56] M. M. Wilde, A. Winter, and D. Yang, “Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy,” *Communications in Mathematical Physics*, vol. 331, no. 2, pp. 593–622, Jul 2014.

[57] F. Hiai and D. Petz, *Introduction to Matrix Analysis and Applications.* Springer International Publishing, 2014.

[58] M. Nussbaum and A. Szkoła, “The Chernoff lower bound for symmetric quantum hypothesis testing,” *Annals of Statistics*, vol. 37, no. 2, pp. 1040–1057, apr 2009.

[59] H. Nagaoka, “The converse part of the theorem for quantum Hoeffding bound,” 2006 arXiv:quant-ph/0611289.

[60] Y. Altuğ and A. B. Wagner, “Refinement of the sphere-packing bound: Asymmetric channels,” *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1592–1614, Mar 2014.

[61] B. Schumacher and M. D. Westmoreland, “Optimal signal ensembles,” *Physical Review A*, vol. 63, no. 2, Jan 2001.

[62] X. Wang, W. Xie, and R. Duan, “Semidefinite programming strong converse bounds for classical capacity,” arXiv:1610.06381 [quant-ph].

Analog quantum error correction with encoding a qubit into an oscillator

Kosuke Fukui[1] * Akihisa Tomita[1] † Atsushi Okamoto[1]

[1] *Graduate School of Information Science and Technology, Hokkaido University*
Kita14-Nishi9, Kita-ku, Sapporo 060-0814, Japan

Abstract. To implement fault-tolerant quantum computation with continuous variables, Gottesman-Kitaev-Preskill (GKP) qubits have been recognized as an important technological element. However, the analog outcome of GKP qubits, which includes beneficial information to improve the error tolerance, has been wasted, because the GKP qubits have been treated as only discrete variables in quantum error-correcting codes. In this paper, we propose a hybrid quantum error correction approach that combines digital information with the analog information of the GKP qubits using the maximum-likelihood method. As an example, a concatenated code known as Knill's C4/C6 code can achieve the hashing bound for the quantum capacity of the Gaussian quantum channel. To the best of our knowledge, this approach is the first attempt to draw both digital and analog information from a single quantum state itself to improve quantum error correction performance.

Keywords: Continuous variables

1 Introduction

Quantum computation (QC) has a great deal of potential [1, 2]. Although a small-scale quantum computation (QC) with various quantum systems have been demonstrated, a large-scale QC is still a significant experimental challenge for most candidates of quantum systems. In continuous variable quantum computation (CV-QC), squeezed vacuum states with the optical setting have shown great potential, because scalable entangled states can be generated by only beam splitter (BS) coupling between two squeezed vacuum states [3]. Hence, CV-QC has attracted a lot of attention toward a large-scale QC. However, a large-scale computation with squeezed vacuum states has been shown to be difficult to achieve because of the accumulation of errors during the QC process, even though the states are created with perfect experimental apparatus [4]. Therefore, fault-tolerant (FT) protection from noise is required that uses the quantum error correcting code. Because noise accumulation originates from the "continuous" nature of the CV-QC, it can be circumvented by encoding CVs into digitized variables using an appropriate code, such as Gottesman–Kitaev–Preskill (GKP) code [5], which are referred to as GKP qubits. Menicucci showed that CV-FTQC is possible within the framework of measurement-based QC using squeezed vacuum cluster states with GKP qubits [4]. Moreover, GKP qubits keep the advantage of squeezed vacuum states on optical implementation that they can be entangled by only BS coupling. Hence, GKP qubits offer a promising element for the implementation of CV-FTQC.

To be practical, the squeezing level required for FTQC should be experimentally achievable. Unfortunately, Menicucci's scheme still requires a 14.8 dB squeezing level to achieve the FT threshold $p_{\mathrm{FT}} = 2 \times 10^{-2}$ [6, 7]. Thus, another twist is necessary to reduce the required squeezing level. It is analog information contained in the GKP qubit that has been overlooked. The effect of noise on CV states is observed as a deviation in an analog measurement outcome, which includes beneficial information for quantum error correction (QEC). Despite this, the analog information from the GKP qubit has been wasted because the GKP qubit has been treated as only a discrete variable (DV) qubit, for which the measurement outcomes are described by bits. Harnessing the wasted information for the QEC will improve the error tolerance compared with using the conventional method based on only bit information.

2 Likelihood function

To utilize analog information from the GKP qubits, we introduced likelihood function as shown in Fig. 1. We make a decision on the bit value $k(=0,1)$ from the measurement outcome of the GKP qubit $q_{\mathrm{m}} = q_k + \Delta_{\mathrm{m}}$ to minimize the deviation $|\Delta_{\mathrm{m}}|$, where $q_k(k=0,1)$ is defined as $(2t+k)\sqrt{\pi}(t = 0, \pm 1, \pm 2, \cdots.)$, shown in Fig. 1 (a). If we consider only digital information k, as in conventional QEC, we waste the analog information contained in Δ_{m}. Instead, we propose a likelihood method to improve our decision for the QEC using analog information. We define the true deviation $|\bar{\Delta}|$ as the difference between the measurement outcome and true peak value $\bar{q}_k$, that is, $|\bar{\Delta}| = |\bar{q}_k - q_m|$. We consider the following two possible events: one is the correct decision, where the true deviation value $|\bar{\Delta}|$ is less than $\sqrt{\pi}/2$ and equals to $|\Delta_{\mathrm{m}}|$ as shown in Fig.1 (b). The other is the incorrect decision, where $|\bar{\Delta}|$ is greater than $\sqrt{\pi}/2$ and satisfies $|\bar{\Delta}| + |\Delta_m| = \sqrt{\pi}$, as shown in Fig.1(c). Because the true deviation value obeys the Gaussian distribution function $f(\bar{\Delta})$, we can evaluate the probabilities of the two events by

$$f(\bar{\Delta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\bar{\Delta}^2/(2\sigma^2)}. \qquad (1)$$

In our method, we regard function $f(\bar{\Delta})$ as a likelihood function. Using this function, the likelihood of the correct decision is calculated by $f(\bar{\Delta}) = f(\Delta_{\mathrm{m}})$. The likelihood of the incorrect decision, whose $|\bar{\Delta}|$ is $\sqrt{\pi} - |\Delta_{\mathrm{m}}|$, is calculated by $f(\bar{\Delta}) = f(\sqrt{\pi} - |\Delta_{\mathrm{m}}|)$. We can reduce the decision error on the entire code word by considering the likelihood of the joint event and choosing the most likely candidate.

[*] fukui.opt@gmail.com

[†] tomita@ist.hokudai.ac.jp

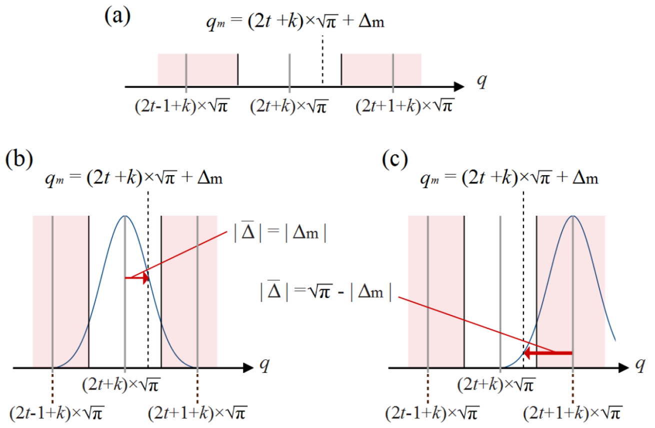

(a)
$q_m = (2t + k)\times\sqrt{\pi} + \Delta\mathrm{m}$
q
$(2t\text{-}1+k)\times\sqrt{\pi}$
$(2t+k)\times\sqrt{\pi}$
$(2t+1+k)\times\sqrt{\pi}$
(b)
$q_m = (2t + k)\times\sqrt{\pi} + \Delta\mathrm{m}$
$|\bar{\Delta}| = |\Delta\mathrm{m}|$
q
$(2t+k)\times\sqrt{\pi}$
$(2t\text{-}1+k)\times\sqrt{\pi}$
$(2t+1+k)\times\sqrt{\pi}$
(c)
$q_m = (2t + k)\times\sqrt{\pi} + \Delta\mathrm{m}$
$|\bar{\Delta}| = \sqrt{\pi} - |\Delta\mathrm{m}|$
q
$(2t+k)\times\sqrt{\pi}$
$(2t\text{-}1+k)\times\sqrt{\pi}$
$(2t+1+k)\times\sqrt{\pi}$

Figure 1: Introduction of a likelihood function. (a) Measurement outcome and deviation from the peak value in q (position) quadrature. The dotted line shows the measurement outcome q_{m} equal to $(2t+k)\sqrt{\pi}+\Delta_{\mathrm{m}}$ $(t=0,\pm1,\pm2,\cdots,\ k=0,1)$, where k is defined as the bit value that minimizes the deviation Δ_{m}. The red areas indicate the area that yields code word $(k+1)$ mod 2, whereas the white area denotes the area that yields the codeword k. (b) and (c) Gaussian distribution functions as likelihood functions of the true deviation value $\bar{\Delta}$ represented by the arrows. (b) refers to the case of the correct decision, where the amplitude of the true deviation value is $|\bar{\Delta}| < \sqrt{\pi}/2$, whereas (c) the case of the incorrect decision $\sqrt{\pi}/2 < |\bar{\Delta}| < \sqrt{\pi}$.

3 Concatenated code with analog information

We demonstrate that the proposed likelihood method improves the error tolerance on a concatenated code, which is indispensable for achieving FTQC. The use of a maximum-likelihood method for a concatenated code was proposed with a message-passing algorithm by Poulin [8], and later Goto and Uchikawa [9] for Knill's C_4/C_6 [6]. However, because previous proposals have been based on the probability of the correct decision given by

$$p_{\mathrm{corr}} = \int_{-\frac{\sqrt{\pi}}{2}}^{\frac{\sqrt{\pi}}{2}} dx \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-x^2/2\sigma^2). \tag{2}$$

The probability p_{corr} is the portion of a normalized Gaussian of a variance σ^2 that lies between $-\sqrt{\pi}/2$ and $\sqrt{\pi}/2$ [4].

We apply our method to the C_4/C_6 code modified with a message-passing algorithm proposed by Goto and Uchikawa [9]. The error correction in the C_4/C_6 code is based on quantum teleportation, where the logical qubit $|\widetilde{\psi}\rangle_{\mathrm{L}}$ encoded by the C_4/C_6 code is teleported to the fresh encoded Bell state. The quantum teleportation process refers to the outcome of the Bell measurement on the encoded qubits and determines the amount of displacement. If this feedforward is performed correctly, the error is successfully corrected. From Bell measurement, we obtain the outcomes of bit values for the physical GKP qubits of the encoded data qubit and encoded qubit of the encoded Bell state. In addition to bit values, we also obtain deviation values for the physical GKP qubits. Therefore, we can improve the error tolerance of the code by introducing the likelihood method to the Bell measurement.

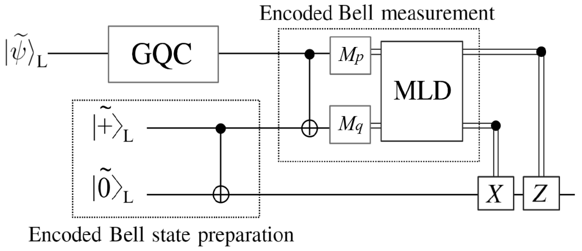

Encoded Bell measurement
$|\widetilde{\psi}\rangle_{\mathrm{L}}$
GQC
M_p
MLD
M_q
$|\widetilde{+}\rangle_{\mathrm{L}}$
$|\widetilde{0}\rangle_{\mathrm{L}}$
X
Z
Encoded Bell state preparation

Figure 2: Error correction by quantum teleportation. The encoded data qubit $|\widetilde{\psi}\rangle_{\mathrm{L}}$, two encoded qubits $|\widetilde{+}\rangle_{\mathrm{L}}$, and $|\widetilde{0}\rangle_{\mathrm{L}}$ are encoded by C_4/C_6 code. GQC and MLD denote the GQC and a maximum-likelihood decision, respectively.

To validate our method, we numerically simulated the quantum teleportation process against Gaussian quantum channel (GQC) [5, 10] for the C_4/C_6 code with the conventional [9] and proposed method using the Monte Carlo method. The error correction in the C_4/C_6 code is based on quantum teleportation, where the logical qubit $|\widetilde{\psi}\rangle_{\mathrm{L}}$ encoded by the C_4/C_6 code is teleported to the fresh encoded Bell state, as shown in Fig.2. The quantum teleportation process refers to the outcomes M_p and M_q of the Bell measurement on the encoded qubits, and determines the amount of displacement. We obtain the Bell measurement outcomes of bit values m_{pi} and m_{qi} for the i-th physical GKP qubit of the encoded data qubit and encoded qubit of the encoded Bell state, respectively. In addition to bit values, we also obtain deviation values Δ_{pmi} and Δ_{qmi} for the i-th physical GKP qubit. Therefore, the proposed likelihood method can improve the error tolerance of the Bell measurement.

As a simple example to explain our method for the Bell measurement, we describe the level-1 C_4/C_6 code, that is, the C_4 code. The C_4 code is the $[[4,2,2]]$ code and consists of four physical GKP qubits to encode a level-1 qubit pair; thus, it is not the error-correcting code but the error-detecting code in the conventional method. The logical bit value of the C_4 code is k (=0,1) when the bit value of the level-1 qubit pair is (k,0) or (k,1), that is, the bit value of the first qubit k defines a logical bit value of a qubit pair. As the parity check of the Z operator for the first and second qubits $ZIZI$ and $IIZZ$ indicates, the bit value of the level-1 qubit pair (0,0) corresponds to the bit value of the physical GKP qubits $(m_{q1}, m_{q2}, m_{q3}, m_{q4})$ = (0,0,0,0) or (1,1,1,1) [6]. The bit values of the pairs (0,1), (1,0), and (1,1) correspond to the bit values of the physical GKP qubits (0,1,0,1) or (1,0,1,0), (0,0,1,1) or (1,1,0,0), and (0,1,1,0) or (1,0,0,1), respectively. Therefore, if the measurement outcome of the physical GKP qubits is (0,0,1,0) for the Z basis, then we consider two error patterns, assuming the level-1 qubit pair (0,0). The first pattern is a single error on the physical qubit 3 and the second pattern is the triple errors on the physical qubits 1, 2, and 4. We then calculate the likelihood for the level-1 qubit pair (0,0) $F_{0,0}$ as

$$\begin{aligned} F_{0,0} &= f(\Delta_{q\mathrm{m}1}) f(\Delta_{q\mathrm{m}2}) f(\sqrt{\pi} - |\Delta_{q\mathrm{m}3}|) f(\Delta_{q\mathrm{m}4}) \\ &+ f(\sqrt{\pi} - |\Delta_{q\mathrm{m}1}|) f(\sqrt{\pi} - |\Delta_{q\mathrm{m}2}|) f(\Delta_{q\mathrm{m}3}) f(\sqrt{\pi} - |\Delta_{q\mathrm{m}4}|). \end{aligned} \tag{3}$$

We similarly calculate the $F_{0,1}, F_{1,0}$, and $F_{1,1}$ likelihood for

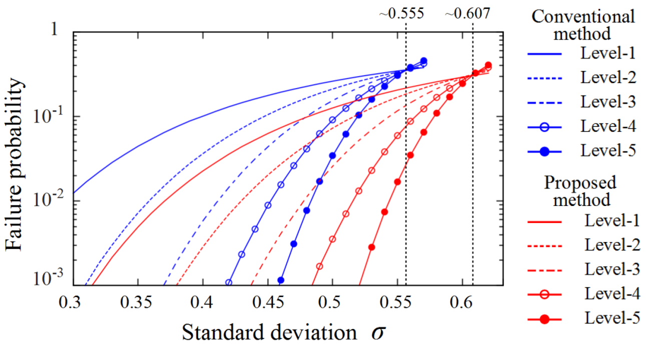


Figure 3: Simulation results for the failure probabilities of the C_4/C_6 code using the conventional and proposed method. The failure probabilities using the conventional method (blue line) and proposed method (red line) are represented for the concatenated level-1 (solid), level-2 (dashed), level-3 (dashed-dotted), level-4 (open circles), and level-5 (filled circles).

the bit value of qubit pairs (0,1), (1,0), and (1,1). Finally, we determine the level-1 logical bit value for the Z basis by comparing $F_{0,0}+F_{0,1}$ with $F_{1,0}+F_{1,1}$, which refer to the likelihood functions for the logical bit values zero and one, respectively. If $F_{0,0}+F_{0,1} > F_{1,0}+F_{1,1}$, then we determine that the level-1 logical bit value for the Z basis is zero, and vice versa. The level-1 logical bit value for the X basis can be determined by the parity check of the X operator for the first and second qubits $XXII$ and $IXIX$ in a similar manner. In the conventional likelihood method [8, 9] $F_{0,0}$, $F_{0,1}$, $F_{1,0}$, and $F_{1,1}$ are given by the same joint probability

$$p_{\mathrm{corr}}^3(1-p_{\mathrm{corr}})+p_{\mathrm{corr}}(1-p_{\mathrm{corr}})^3, \tag{4}$$

where the probability p_{corr} is defined by Eq. (2) in the main text. Because $F_{0,0}+F_{0,1}=F_{1,0}+F_{1,1}$, the C_4 code is not error-correcting code but error-detecting code in the conventional method, whereas it is the error-correcting code in our method. For higher levels of concatenation, the likelihood for the level-l ($l \geqq 2$) bit value can be calculated by the likelihood for the level-$(l-1)$ bit value in a similar manner.

In Fig.3, the failure probabilities up to level-5 of the concatenation are plotted as a function of the data qubit's deviation. The results confirm that our method suppresses errors more effectively than the conventional method. It is also remarkable that our method achieves the hashing bound of the standard deviation for the quantum capacity of the GQC ~ 0.607, which corresponds to the squeezing level of 1.3 dB and has been conjectured to be an attainable value using the optimal method [5, 10]. The quantum capacity is defined as the supremum of all achievable rates at which quantum information can be transmitted over the quantum channel and the hashing bound of the standard deviation is the maximum value of the condition that yields the non-zero positive quantum capacity. By contrast, the concatenated code with only digital information achieves the hashing bound ~ 0.555 [5, 10], which corresponds to the squeezing level of 2.1 dB. This implies that QEC using our method provides an optimal performance against GQC, while QEC using only digital information provides suboptimal performance.

4 Conclusion

We proposed a new approach to maximize QEC performance with digitized CV states. To our knowledge, our approach is the first attempt to draw both digital and analog information from a single CV state to improve QEC performance. Our method can reduce the threshold of squeezing level required for CV-FTQC, which will encourage the experimental developments by alleviating the burden to implement. Our method can be applied to not only C_4/C_6 code, but also surface code [11], color code [12], and other QECs. Furthermore, our method is a versatile tool for decision, which can incorporate with GKP qubit, cat code, and other various codes used to digitize CV states. We believe this work will open up a new approach to QEC with digitized CV states, which will be indispensable to construct CV-FTQC.

Although several methods to implement GKP qubits have been proposed [13, 14, 15], it is still difficult to experimentally generate GKP qubits with the squeezing level required for FTQC. Our method can alleviate this requirement, and will encourage experimental developments.

This work was funded by ImPACT Program of Council for Science, Technology and Innovation.

References

[1] P. Shor, In Proceeding of 35th IEEE FOCS, pp.124-134, Santa Fe, NM, Nov 20-22 (1994).

[2] L. Grover, STOC'96, pp.212-219, Philadelphia, Pennsylvania, United States, May 22-24 (1996).

[3] J. Yoshikawa, S. Yokoyama, T. Kaji, C. Sornphiphatphong, Y. Shiozawa, K. Makino, and A. Furusawa, APLPhotonics **1** 060801 (2016).

[4] N. C. Menicucci, Phys. Rev. Lett. **112**, 120504 (2014).

[5] D. Gottesman, A. Kitaev, and J. Preskill, Phys. Rev. A **64**, 012310 (2001).

[6] E. Knill, Nature, **434**, 39-44 (2005).

[7] K. Fujii and K. Yamamoto, Phys. Rev. A **81**, 042324 (2010).

[8] D. Poulin, Phys. Rev. A **74**, 052333 (2006).

[9] H. Goto and H. Uchikawa, Sci. Rep. **3**, 2044 (2013).

[10] J. Harrington and J. Preskill, Phys. Rev. A **64**, 062301 (2001).

[11] A. Y. Kitaev, Ann. Phys. **303**, 2 (2003).

[12] H. Bombin and M. A. Mrtin-Delgado, Phys. Rev. Lett. **97**, 180501 (2006).

[13] H. M. Vasconcelos, L. Sanz, and S. Glancy, Opt. Lett.**35**, 3261 (2010).

[14] B.M.Terhal and D. Weigand, Phys. Rev. A **93**, 012315 (2016).

[15] K. R. Motes, B. Q. Baragiola, A. Gilchrist, N. C. Menicucci, Phys. Rev. A **95**, 053819 (2017).

Entanglement in interactive proof systems

Zhengfeng Ji

UTS, Sydney, Australia

Abstract: The area of interactive proof system studies the procedure of proof verification from the computer science perspective and has played a key role in computational complexity theory. This talk will focus on the study of interactive proof systems through the lens of quantum information processing. Entanglement, the key of quantum weirdness, has been both the cause and the solution of many problems in quantum proof systems. In this talk, we will discuss several such examples and highlight the role of entanglement in quantum interactive proof systems.

Irreversibility of Asymptotic Entanglement Manipulation Under PPT-preserving Operations

Xin Wang[1] [*] Runyao Duan[1] [2] [†] [‡]

[1] *Centre for Quantum Software and Information,*
Faculty of Engineering and Information Technology,
University of Technology Sydney, NSW 2007, Australia

[2] *UTS-AMSS Joint Research Laboratory for Quantum Computation and Quantum Information Processing,*
Academy of Mathematics and Systems Science,
Chinese Academy of Sciences, Beijing 100190, China

Abstract. We demonstrate the irreversibility of asymptotic entanglement manipulation under quantum operations that completely preserve positivity of partial transpose (PPT), which resolves a major open problem in quantum information theory. To be more specific, we show that for any rank-two mixed state supporting on the $3 \otimes 3$ antisymmetric subspace, the amount of distillable entanglement by PPT operations is strictly smaller than one entanglement bit (ebit) while its entanglement cost under PPT operations is exactly one ebit. As a byproduct, we find that for this class of quantum states, both the Rains' bound and its regularization, are strictly less than the asymptotic relative entropy of entanglement with respect to PPT states. So, in general, there is no unique entanglement measure for the manipulation of entanglement by PPT operations. We further present a feasible sufficient condition for the irreversibility of entanglement manipulation under PPT operations.

Keywords: Distillable entanglement, Entanglement measure, Irreversibility, Entanglement cost

Introduction Entanglement plays a crucial role in quantum physics and is the key resource in quantum information processing. So it is quite natural and important to develop a theoretical framework to describe and quantify it. In spite of a series of remarkable recent progress in the theory of entanglement (for reviews see, e.g., [1, 2, 3, 4]), many fundamental challenges still remain open. One of the most significant task is to determine the *distillable entanglement* E_D, i.e. the highest rate at which one can obtain maximally entangled states from an entangled state by local operations and classical communication (LOCC) [5, 6]. This fundamental measure fully captures the ability of given state shared between distant parties to generate strongly correlated qubits in order to allow reliable quantum teleportation or quantum cryptography. However, how to calculate E_D for general quantum states still remains unknown. Another fundamental measure in entanglement theory is entanglement cost E_C [5, 7], which quantifies the rate for converting maximally entangled states to the given state by LOCC alone. Entanglement cost is also difficult to evaluate [8] and it is known only for a few of quantum states [9, 10, 11].

A well-known upper bound of the distillable entanglement is the relative entropy of entanglement w.r.t PPT states [12, 13, 14], i.e., $E_{R,PPT}(\rho) = \min S(\rho\|\sigma)$ s.t. $\sigma, \sigma^{T_B} \geq 0, \operatorname{Tr}\sigma = 1$, where $S(\rho\|\sigma) = \operatorname{Tr}(\rho \log_2 \rho - \rho \log_2 \sigma)$ denotes the relative Von Neumann entropy and the optimal solution σ is called the closest PPT state of ρ. An improved bound is the Rains' bound [15], which is arguably the best known upper bound of distillable entanglement and refined in [16] as a convex optimization problem as $R(\rho) = \min S(\rho\|\tau)$ s.t. $\tau \geq 0, \operatorname{Tr}|\tau^{T_B}| \leq 1$. As Rains' bound is proved to be equal to the asymptotic relative entropy of entanglement for Werner states [17] and orthogonally invariant states [16], one open problem is to determine whether these two quantities always coincide [1].

Another fundamental problem in entanglement theory is the irreversibility in entanglement manipulations. The manipulation of entanglement under LOCC is generally irreversible in the finite-copy regime [5]. Surprisingly, in the asymptotic settings where the number of copies tend to infinite, this process of entanglement manipulation for bipartite pure states is shown to be reversible [18]. In contrast, for mixed states, this asymptotic reversibility under LOCC operations does not hold anymore

[*] `xin.wang-8@student.uts.edu.au`

[†] `runyao.duan@uts.edu.au`

[‡] This submission is based on arxiv:1606.09421.

[19, 20, 9, 21, 22, 23]. Various approaches have been considered to enlarge the class of operations to ensure reversible interconversion of entanglement in the asymptotic setting. A natural candidate is the class of quantum operations that completely preserve positivity of partial transpose (PPT) [15]. A remarkable result is that any state with a nonpositive partial transpose (NPT) is distillable under this class of operations [24]. This suggests the possibility of reversibility under PPT operations and there are examples of mixed states which can be reversibly converted into pure states in the asymptotic setting, e.g. the class of antisymmetric states [25]. However, the reversibility under PPT operations remained unsolved [25, 26, 1, 27] and it is one of the major open problems in quantum information theory [28].

The main difficulty of the problems above is that the regularized quantities are usually extremely difficult to determine or estimate. Note that the asymptotic relative entropy of entanglement w.r.t PPT states is given by $E^{\infty}_{R,PPT}(\rho) = \inf_{n\geq 1} E_{R,PPT}(\rho^{\otimes n})/n$. To figure out whether Rains' bound always coincides with $E^{\infty}_{R,PPT}$, one necessarily has to evaluate $E^{\infty}_{R,PPT}(\rho)$ of an explicit state ρ. The problem of irreversibility under PPT operations is more intractable, one not only has to evaluate the PPT distillable entanglement, but also needs to determine the PPT entanglement cost.

In this paper, we resolve the open problems mentioned above via convex optimization approach. In particular, we utilize semidefinite programming techniques to overcome the difficulty of evaluating regularized quantities. The main results of this work are as follows:

(i) Rains' bound and its regularization can be strictly smaller than the asymptotic relative entropy of entanglement;

(ii) Asymptotic entanglement manipulation under PPT operations is irreversible, i.e., PPT operations are not sufficient to ensure asymptotically reversibly interconversion (see FIG. 1).

Lower bound for asymptotic relative entropy of entanglement To see results (i) and (ii), the key approach is to introduce a single-letter additive SDP lower bound for E^{∞}_{R} and construct an explicit class of states. We will first introduce the SDP lower bound and show the separation between the regularized Rains' bound and $E^{\infty}_{R,PPT}$ later in Eq. (2). We define $D(\rho)$ as the set of quantum states supporting on supp(ρ) and denote the set of PPT states by Γ. Then, the problem can be relaxed to the minimization of the relative entropy distance between $D(\rho)$ and the set Γ. Applying some properties of quantum relative entropy, the problem can be further relaxed to minimizing $-\log \operatorname{Tr} P_{AB}\sigma$ over all PPT states σ, where P_{AB} is the projection onto supp(ρ). Noting that this is SDP-computable, we can further use SDP techniques to lower bound the regularized quantity, i.e.,

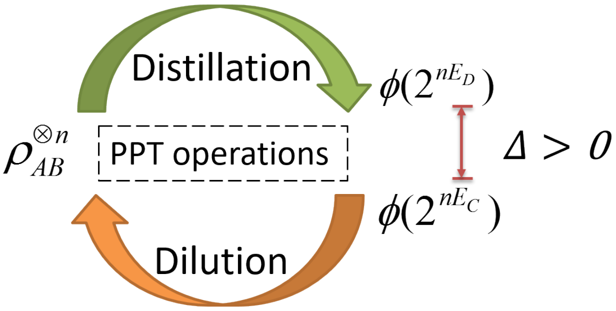


Figure 1: The amount of Bell states distilled from the state is not enough to reproduce the given state under PPT operations in the asymptotic regime.

$$\begin{aligned} E^{\infty}_{R,PPT}(\rho) &= \inf_{n\geq 1} \frac{E_{R,PPT}(\rho^{\otimes n})}{n} \\ &\geq E_{\eta}(\rho) = \max -\log_2 \|Y_{AB}^{T_B}\|_{\infty}, \\ &\qquad \text{s.t.} \ \ -Y_{AB} \leq P_{AB}^{T_B} \leq Y_{AB}. \end{aligned} \tag{1}$$

The key idea here is utilizing the duality theory of SDP to prove that E_{η} is additive under tensor product. It is worth noting that E_{η} provides an efficiently computable lower bound for entanglement cost, i.e.,

$$E_C(\rho) \geq E_{C,PPT}(\rho) \geq E_{\eta}(\rho).$$

One can also obtain SDP lower bound for entanglement cost of quantum channels [29] via E_{η}.

Irreversibility of asymptotic entanglement manipulation under PPT operations We use a $3\otimes 3$ state to show the irreversibility under PPT operations and then present an SDP-computable sufficient condition for the irreversibility. The state we use is $\rho_v = \frac{1}{2}(|v_1\rangle\langle v_1| + |v_2\rangle\langle v_2|)$ with $|v_1\rangle = (|01\rangle - |10\rangle)/\sqrt{2}, |v_2\rangle = (|02\rangle - |20\rangle)/\sqrt{2}$. It is clear that ρ_v is a rank-two state supporting on the $3\otimes 3$ antisymmetric subspace. On one hand, we apply the lower bound E_{η} to estimate the PPT entanglement cost and prove that $E_{C,PPT}(\rho_v) = E^{\infty}_{R,PPT}(\rho_v) = E_{\eta}(\rho_v) = 1$. On the other hand, we evaluate the PPT distillable entanglement of ρ_v by the Rains bound and the SDP characterization of the one-copy PPT deterministic distillable entanglement [30], i.e.,

$E_{D,PPT}(\rho_v) = R^\infty(\rho_v) = \log_2(1+1/\sqrt{2})$. Hence,

$$\begin{aligned} E_{C,PPT}(\rho_v) &= E^\infty_{R,PPT}(\rho_v) = 1 \\ &> \log_2(1+1/\sqrt{2}) \\ &= R^\infty(\rho_v) = E_{D,PPT}(\rho_v). \end{aligned} \tag{2}$$

We further show that for any rank-two mixed state supporting on the $3\otimes 3$ antisymmetric subspace, the PPT distillable entanglement is strictly smaller than one entanglement bit (ebit) while its PPT entanglement cost is exactly one ebit.

As a byproduct, from Eq. (2), it is clear for ρ_v, both the Rains' bound and its regularization, are strictly less than the asymptotic relative entropy of entanglement, which resolve the second problem. So in general there is no unique entanglement measure under PPT operations.

Finally, we present an SDP-computable sufficient condition for the irreversibility of entanglement manipulation under PPT operations. For a bipartite state ρ, if $E_\eta(\rho) > E_W(\rho) = \min_{X_{AB}\geq\rho} \log \|X_{AB}^{T_B}\|_1$, then

$$E_{D,PPT}(\rho) \leq E_W(\rho) < E_\eta(\rho) \leq E_{C,PPT}(\rho), \tag{3}$$

where E_W is an improved SDP upper bound on PPT distillable entanglement in our previous work [30]. As a example, we show the irreversibility under PPT operations for a class of $3\otimes 3$ states defined by $\rho^{(\alpha)} = (|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|)/2$, where $|\psi_1\rangle = \sqrt{\alpha}|01\rangle - \sqrt{1-\alpha}|10\rangle$ and $|\psi_2\rangle = \sqrt{\alpha}|02\rangle - \sqrt{1-\alpha}|20\rangle$ with $0.42 \leq \alpha \leq 0.5$. It is worth noting that Brandão and Plenio [31, 27] have shown that multipartite entangled states can be reversibly interconverted under asymptotically non-entangling operations. Our results may imply that this set of operations is minimal for bipartite entanlged states to ensure the reversibility.

RD would like to thank Andreas Winter for many inspirational discussions on the potential gap between the regularized Rains' bound and the asymptotic relative entropy of entanglement. The authors also thank Jonathan Oppenheim for helpful suggestions. This work was partly supported by the Australian Research Council under Grant Nos. DP120103776 and FT120100449.

References

[1] M. B. Plenio and S. S. Virmani, "An introduction to entanglement measures," *Quant. Inf. Comp.*, vol. 7, no. 1, pp. 1–51, 2007.

[2] J. Eisert, "Entanglement in quantum information theory," *arXiv preprint quant-ph/0610253*, 2006.

[3] M. Christandl, "The structure of bipartite quantum states-Insights from group theory and cryptography," *arXiv preprint quant-ph/0604183*, 2006.

[4] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Reviews of Modern Physics*, vol. 81, no. 2, p. 865, 2009.

[5] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed-state entanglement and quantum error correction," *Physical Review A*, vol. 54, no. 5, p. 3824, 1996.

[6] E. M. Rains, "Rigorous treatment of distillable entanglement," *Physical Review A*, vol. 60, no. 1, p. 173, 1999.

[7] P. M. Hayden, M. Horodecki, and B. M. Terhal, "The asymptotic entanglement cost of preparing a quantum state," *Journal of Physics A: Mathematical and General*, vol. 34, no. 35, p. 6891, 2001.

[8] Y. Huang, "Computing quantum discord is NP-complete," *New journal of physics*, vol. 16, no. 3, p. 33027, 2014.

[9] G. Vidal, W. Dür, and J. I. Cirac, "Entanglement cost of bipartite mixed states," *Physical Review Letters*, vol. 89, no. 2, p. 27901, 2002.

[10] F. Yura, "Entanglement cost of three-level antisymmetric states," *Journal of Physics A Mathematical and General*, vol. 36, no. 15, 2003.

[11] K. Matsumoto and F. Yura, "Entanglement cost of antisymmetric states and additivity of capacity of some quantum channels," *Journal of Physics A: Mathematical and General*, vol. 37, no. 15, p. L167, 2004.

[12] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, "Quantifying entanglement," *Physical Review Letters*, vol. 78, no. 12, p. 2275, 1997.

[13] V. Vedral and M. B. Plenio, "Entanglement measures and purification procedures," *Physical Review A*, vol. 57, no. 3, p. 1619, 1998.

[14] V. Vedral, M. B. Plenio, K. Jacobs, and P. L. Knight, "Statistical inference, distinguishability of quantum states, and quantum entanglement," *Physical Review A*, vol. 56, no. 6, p. 4452, 1997.

[15] E. M. Rains, "A semidefinite program for distillable entanglement," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2921–2933, 2001.

[16] K. Audenaert, B. De Moor, K. G. H. Vollbrecht, and R. F. Werner, "Asymptotic relative entropy of entanglement for orthogonally invariant states," *Physical Review A*, vol. 66, no. 3, p. 32310, 2002.

[17] K. Audenaert, J. Eisert, E. Jané, M. B. Plenio, S. Virmani, and B. De Moor, "Asymptotic relative entropy of entanglement," *Physical Review Letters*, vol. 87, no. 21, p. 217902, 2001.

[18] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, "Concentrating partial entanglement by local operations," *Physical Review A*, vol. 53, no. 4, p. 2046, 1996.

[19] G. Vidal and J. I. Cirac, "Irreversibility in asymptotic manipulations of entanglement," *Physical Review Letters*, vol. 86, no. 25, p. 5803, 2001.

[20] ——, "Irreversibility in asymptotic manipulations of a distillable entangled state," *Physical Review A*, vol. 65, no. 1, p. 12323, 2001.

[21] K. G. H. Vollbrecht, R. F. Werner, and M. M. Wolf, "Irreversibility of entanglement distillation for a class of symmetric states," *Physical Review A*, vol. 69, no. 6, p. 62304, 2004.

[22] D. Yang, M. Horodecki, R. Horodecki, and B. Synak-Radtke, "Irreversibility for all bound entangled states," *Physical Review Letters*, vol. 95, no. 19, p. 190501, 2005.

[23] M. F. Cornelio, M. C. de Oliveira, and F. F. Fanchini, "Entanglement irreversibility from quantum discord and quantum deficit," *Physical Review Letters*, vol. 107, no. 2, p. 20502, 2011.

[24] T. Eggeling, K. G. H. Vollbrecht, R. F. Werner, and M. M. Wolf, "Distillability via protocols respecting the positivity of partial transpose," *Physical Review Letters*, vol. 87, no. 25, p. 257902, 2001.

[25] K. Audenaert, M. B. Plenio, and J. Eisert, "Entanglement cost under positive-partial-transpose-preserving operations," *Physical Review Letters*, vol. 90, no. 2, p. 27901, 2003.

[26] M. Horodecki, J. Oppenheim, and R. Horodecki, "Are the Laws of Entanglement Theory Thermodynamical?" *Physical Review Letters*, vol. 89, no. 24, p. 240403, nov 2002.

[27] F. G. S. L. Brandão and M. B. Plenio, "A reversible theory of entanglement and its relation to the second law," *Communications in Mathematical Physics*, vol. 295, no. 3, pp. 829–851, 2010.

[28] M. Plenio, "Problem 20," *Problem 20 of the list https://qig.itp.uni-hannover.de/qiproblems/Open_Problems.*

[29] M. Berta, F. G. S. L. Brandao, M. Christandl, and S. Wehner, "Entanglement cost of quantum channels," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6779–6795, 2013.

[30] X. Wang and R. Duan, "Improved semidefinite programming upper bound on distillable entanglement," *Physical Review A*, vol. 94, no. 5, p. 050301, nov 2016.

[31] F. G. S. L. Brandao and M. B. Plenio, "Entanglement theory and the second law of thermodynamics," *Nature Physics*, vol. 4, no. 11, pp. 873–877, 2008.

Non-asymptotic entanglement distillation

Kun Fang[1] * Xin Wang[1] † Marco Tomamichel[1] ‡ Runyao Duan[1 2] § ¶

[1] *Centre for Quantum Software and Information,*
Faculty of Engineering and Information Technology,
University of Technology Sydney, NSW 2007, Australia

[2] *UTS-AMSS Joint Research Laboratory for Quantum Computation and Quantum Information Processing,*
Chinese Academy of Sciences, Beijing 100190, China

Abstract. Non-asymptotic entanglement distillation studies the trade-off between three parameters: the distillation rate, the number of independent and identically distributed prepared states, and the fidelity of the distillation. We first study the one-shot ε-infidelity distillable entanglement under quantum operations that completely preserve positivity of the partial transpose (PPT) and characterize it as a semidefinite program (SDP). For isotropic states, it can be further simplified to a linear program. The one-shot ε-infidelity PPT-assisted distillable entanglement can be transformed to a quantum hypothesis testing problem. Moreover, we show efficently computable second-order upper and lower bounds for the non-asymptotic distillable entanglment with a given infidelity tolerance. Utilizing these bounds, we obtain the second order asymptotic expansions of the optimal distillation rates for pure states and some classes of mixed states. In particular, this result recovers the second-order expansion of LOCC distillable entanglement for pure states in [Datta/Leditzky, *IEEE Trans. Inf. Theory* **61**:582, 2015]. Furthermore, we provide an algorithm for calculating the Rains bound and present direct numerical evidence (not involving any other entanglement measures, as in [Wang/Duan, *Phys. Rev. A* **95**:062322, 2017]), showing that the Rains bound is not additive under tensor products.

Keywords: entanglement distillation, Rains bound, hypothesis testing, semidefinite program

1 Introduction

Quantum entanglement is a striking feature of quantum mechanics and is a key ingredient in many quantum information processing tasks, including the teleportation [1], superdense coding [2], and numerous uses in quantum cryptography protocols [3, 4]. All these protocols necessarily rely on entanglement resources, especially the maximally entangled states.

In general, the task of entanglement distillation aims at obtaining maximally entangled states from less-entangled bipartite states shared between two parties and it allows them to perform LOCC. The concept of *distillable entanglement* characterizes the rate at which one can asymptotically obtain maximally entangled states from a collection of identically and independently distributed (i.i.d) prepared entangled states by LOCC [5, 6]. Distillation from non-i.i.d prepared states has also been considered recently [7]. Distillable entanglement is a fundamental entanglement measure which captures the resource character of entanglement. Up to now, how to calculate distillable entanglement for general quantum states remains unknown and various approaches [8, 9, 10, 11, 12, 13, 14, 15] have been developed to evaluate this important quantity.

However, in a realistic setting, the resources are finite and the number of independent and identically distributed (i.i.d.) prepared states is necessarily limited. More importantly, it is notoriously hard to do coherent state manipulation over a very large numbers of qubits in the current status or near future. Therefore, it is important to characterize how well we can distill maximally entangled states from finite copies of prepared states. Under the non-asymptotic setting, one also has to make a trade-off between the distillation rate and infidelity tolerance.

The non-asymptotic analysis of entanglement distillation will help us better exploit the power of entanglement in a realistic setting. Previously, the one-shot distillable entanglement was studied in Refs. [16, 17], but their bounds are not efficiently computable. The Rains bound [11] and the hashing bound [18] are arguably the best general upper and lower bound for distillable entanglement, respectively. However, these bounds do not provide sufficiently good evaluation about entanglement distillation with finite resources.

* `kun.fang-1@student.uts.edu.au`

† `xin.wang-8@student.uts.edu.au`

‡ `marco.tomamichel@uts.edu.au`

§ `runyao.duan@uts.edu.au`

¶ This submission is based on arXiv:1706.06221.

2 Overview of results

In this work, we study the entanglement distillation with finite resources and provide efficiently computable estimation of the non-asymptotic distillable entanglement. Our approach utilizes the techniques of convex optimization and second-order expansion of hypothesis testing.

We first give the SDP characterization of one-shot PPT distillable entanglement with a given infidelity tolerance and connect it to a hypothesis testing problem. For given bipartite state ρ_{AB} and infidelity tolerance $\varepsilon > 0$, we show that

$$
\begin{aligned}
E^{(1)}_{\Gamma,\varepsilon}(\rho_{AB}) = -\log\min\ \eta & \\
\text{s.t. } & \operatorname{Tr}\rho_{AB}M_{AB} \ge 1-\varepsilon, \\
& 0 \le M_{AB} \le \mathbb{1}_{AB}, \\
& -\eta\mathbb{1}_{AB} \le M_{AB}^{T_B} \le \eta\mathbb{1}_{AB}
\end{aligned}
\tag{1}
$$

and

$$
E^{(1)}_{\Gamma,\varepsilon}(\rho_{AB}) = \min_{\|C^{T_B}\|_1 \le 1} D_H^{\varepsilon}(\rho\|C). \tag{2}
$$

As the one-shot PPT distillable entanglement can be represented in the form of hypothesis testing relative entropy, we further derive a second-order upper bound of the non-asymptotic distillable entanglement. Specifically, for any bipartite states ρ_{AB} and given infidelity tolerance $\varepsilon \in (0,1)$, we show that

$$
E^{(1)}_{\Gamma,\varepsilon}(\rho^{\otimes n}) \le nR(\rho) + \sqrt{nV_R(\rho)}\Phi^{-1}(\varepsilon) + O(\log n), \tag{3}
$$

$$
\text{where } V_R(\rho) = \begin{cases} \max_{\sigma\in\mathcal{S}_\rho} V(\rho\|\sigma) & \text{if } \ 0<\varepsilon\le 1/2 \\ \min_{\sigma\in\mathcal{S}_\rho} V(\rho\|\sigma) & \text{if } \ 1/2<\varepsilon<1 \end{cases}, \tag{4}
$$

and $V(\rho\|\sigma) = \operatorname{Tr}\rho(\log\rho - \log\sigma)^2 - D(\rho\|\sigma)^2$, $\mathcal{S}_\rho$ is the set of operators that achieve the minimum of $R(\rho) = \min_{\sigma\in\mathrm{PPT}'} D(\rho\|\sigma)$ and Φ^{-1} is the cumulative normal distribution function.

We also observe that $R(\rho)$ and $V_R(\rho)$ can be efficiently computed via the cutting-plane method [19] or rational (Padé) approximations [20]. This allows us to efficiently compute the second-order converse bound of non-asymptotic distillable entanglement in Eq. (3). Moreover, one can use these algorithms to verify the non-additivity of Rains bound [21] as shown in Fig. 4.

On the other hand, to estimate the achievability of non-asymptotic entanglement distillation, we give the second order expansion of the 1-LOCC hashing lower bound [18]. To be specific, for any bipartite state ρ and infidelity tolerance $\varepsilon \in (0,1)$,

$$
E^{(1)}_{\to,\varepsilon}(\rho_{AB}^{\otimes n}) \ge nI(A\rangle B)_\rho + \sqrt{nV(A\rangle B)_\rho}\Phi^{-1}(\varepsilon) + O(\log n). \tag{5}
$$

Finally, we use our results to study some particular classes of bipartite quantum states, including pure states, isotropic states and some other classes of mixed states.

(i) For any bipartite pure state ψ_{AB}, denote the reduced state as $\rho_A = \operatorname{Tr}_B \psi_{AB}$, then

$$
\begin{aligned}
& E^{(1)}_{\to,\varepsilon}(\psi^{\otimes n}) = E^{(1)}_{\Gamma,\varepsilon}(\psi^{\otimes n}) = nS(\rho_A) \\
& + \sqrt{n\left[\operatorname{Tr}\rho_A(\log\rho_A)^2 - S(\rho_A)^2\right]}\Phi^{-1}(\varepsilon) + O(\log n).
\end{aligned}
$$

(ii) For the bipartite quantum state $\rho_{AB} = p|v_1\rangle\langle v_1| + (1-p)|v_2\rangle\langle v_2|$, where $|v_1\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, $|v_2\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$, its second-order distillable entanglement is

$$
\begin{aligned}
E^{(1)}_{\to,\varepsilon}(\rho_{AB}^{\otimes n}) = E^{(1)}_{\Gamma,\varepsilon}(\rho_{AB}^{\otimes n}) &= n(1-h_2(p)) \\
&+ \sqrt{np(1-p)\left(\log\frac{1-p}{p}\right)^2} + O(\log n).
\end{aligned}
$$

(iii) For the isotropic states $\rho_F = (1-F)\frac{\mathbb{1}-\Phi(d)}{d^2-1} + F\cdot\Phi(d)$, we have the linear program:

$$
\begin{aligned}
E^{(1)}_{\Gamma,\varepsilon}(\rho_F^{\otimes n}) = -\log\min\ \eta & \\
\text{s.t. } & 0 \le m_i \le 1,\ i = 0,1,\cdots,n, \\
& \sum_{i=0}^{n}\binom{n}{i}F^i(1-F)^{n-i}m_i \ge 1-\varepsilon, \\
& -\eta \le \sum_{i=0}^{n} x_{i,k}m_i \le \eta,\ k = 0,1,\cdots,n.
\end{aligned}
\tag{6}
$$

We further show the numerical estimation of non-asymptotic distillable entanglement of $3\otimes 3$ isotropic state ρ_F $(F = 0.9)$ with infidelity tolerance $\varepsilon = 0.001$ in the following figures. Fig. 1 shows that the 1-LOCC hashing bound cannot be achieved by coherently manipulating 100 copies of the isotropic state while such manipulation is already hard to perform in practice. Fig. 2 shows the estimation of non-asymptotic (1-LOCC, LOCC, SEP, PPT-assisted) distillable entanglement of the isotropic state. The finite blocklength distillable entanglement will lie between two dashed lines while asymptotic distillation rates lie between the two solid lines. Fig. 3 shows that the fitting curve of the series of points $\frac{1}{n}E^{(1)}_{\Gamma,\varepsilon}(\rho_F^{\otimes n})$ $(1 \le n \le 100)$ almost coincides with the second-order upper bound in large n ($\geq 10^3$) and converges to its Rains bound. This may indicates that $E_\Gamma(\rho_F) = R(\rho_F)$.

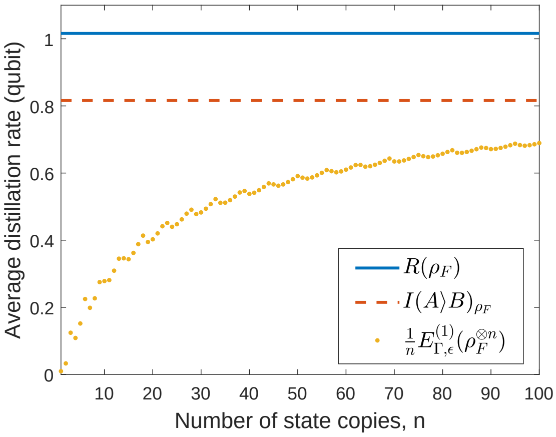


Figure 1:

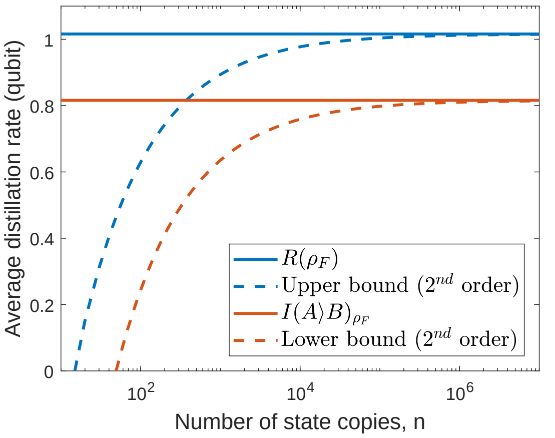


Figure 2:

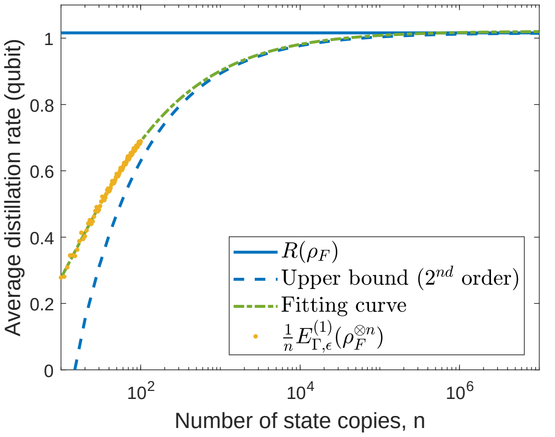


Figure 3:

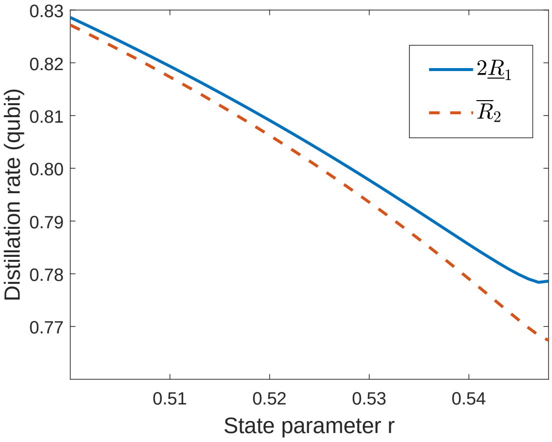


Figure 4: Non-additivity of Rains bound.

References

[1] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Physical Review Letters*, vol. 70, no. 13, p. 1895, 1993.

[2] C. H. Bennett and S. J. Wiesner, "Communication via one-and two-particle operators on Einstein-Podolsky-Rosen states," *Physical Review Letters*, vol. 69, no. 20, p. 2881, 1992.

[3] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *International Conference on Computers, Systems & Signal Processing, Bangalore, India, Dec 9-12, 1984*, 1984, pp. 175–179.

[4] A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Physical Review Letters*, vol. 67, no. 6, pp. 661–663, aug 1991.

[5] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, "Purification of noisy entanglement and faithful teleportation via noisy channels," *Physical Review Letters*, vol. 76, no. 5, p. 722, 1996.

[6] E. M. Rains, "Rigorous treatment of distillable entanglement," *Physical Review A*, vol. 60, no. 1, p. 173, 1999.

[7] S. Waeldchen, J. Gertis, E. T. Campbell, and J. Eisert, "Renormalizing Entanglement Distillation," *Physical Review Letters*, vol.

116, no. 2, p. 020502, jan 2016. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.116.020502

[8] V. Vedral and M. B. Plenio, "Entanglement measures and purification procedures," *Physical Review A*, vol. 57, no. 3, p. 1619, 1998.

[9] E. M. Rains, "Bound on distillable entanglement," *Physical Review A*, vol. 60, no. 1, p. 179, 1999.

[10] G. Vidal and R. F. Werner, "Computable measure of entanglement," *Physical Review A*, vol. 65, no. 3, p. 32314, 2002.

[11] E. M. Rains, "A semidefinite program for distillable entanglement," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2921–2933, 2001.

[12] M. Horodecki, P. Horodecki, and R. Horodecki, "Asymptotic manipulations of entanglement can exhibit genuine irreversibility," *Physical Review Letters*, vol. 84, no. 19, p. 4260, 2000.

[13] M. Christandl and A. Winter, "Squashed entanglement: An additive entanglement measure," *Journal of Mathematical Physics*, vol. 45, no. 3, pp. 829–840, 2004.

[14] F. Leditzky, N. Datta, and G. Smith, "Useful states and entanglement distillation," *arXiv:1701.0308*, jan 2017.

[15] X. Wang and R. Duan, "Improved semidefinite programming upper bound on distillable entanglement," *Physical Review A*, vol. 94, no. 5, p. 050301, nov 2016.

[16] F. Buscemi and N. Datta, "Distilling entanglement from arbitrary resources," *Journal of Mathematical Physics*, vol. 51, no. 10, p. 102201, oct 2010.

[17] F. G. S. L. Brandao and N. Datta, "One-Shot Rates for Entanglement Manipulation Under Non-entangling Maps," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1754–1760, mar 2011.

[18] I. Devetak and A. Winter, "Distillation of secret key and entanglement from quantum states," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 461, no. 2053. The Royal Society, 2005, pp. 207–235.

[19] Y. Zinchenko, S. Friedland, and G. Gour, "Numerical estimation of the relative entropy of entanglement," *Physical Review A*, vol. 82, no. 5, p. 52336, 2010.

[20] H. Fawzi and O. Fawzi, "Relative entropy optimization in quantum information theory via semidefinite programming approximations," *arXiv preprint arXiv:1705.06671*, pp. 1–14, may 2017.

[21] X. Wang and R. Duan, "Rains' bound is not additive," *arXiv:1605.00348*, 2016.

Superadditivity of the classical capacity with limited entanglement assistance (arXiv:1704.06955)

Elton Yechao Zhu[1] [2] [*] Quntao Zhuang[1] [3] [†] Peter W. Shor[2] [4] [‡]

[1] *Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[2] *Center For Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[3] *Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[4] *Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

Abstract. Finding the optimal encoding strategies can be challenging for communication using quantum channels, as classical and quantum capacities may be superadditive. Pre-shared entanglement assistance can often simplify this task, as the entanglement-assisted classical capacity for any channel is additive. If the entanglement assistance is limited, the picture is much more unclear. If the classical capacity is additive, it is unknown if superadditivity can still be developed with limited entanglement assistance. We show this is possible, by providing an example. We construct a channel for which, the classical capacity is additive, but that with limited entanglement assistance can be superadditive.

Keywords: quantum Shannon theory, quantum channels, entanglement, trade-off capacities, superadditivity

Channel capacities describe the maximum rate at which a channel can transmit information. A classical channel can only transmit classical information, and the maximum communication rate is fully characterized by its capacity [1], which is described by a simple formula.

For quantum channels, however, the story is very different. First, a quantum channel can transmit both classical and quantum information. Hence there are a few different types of capacity, such as the classical capacity $\mathcal{C}$ [2, 3] and the quantum capacity $\mathcal{Q}$ [4, 5, 6]. Moreover, the capacity formulae are of a “multi-letter” nature, optimized over inputs of an infinite number of channel uses. This is easy to understand, as quantum inputs can be entangled across channels. Sometimes such inputs can improve the communication rate, leading to the “superadditivity” phenomenon. Many channels with such phenomenon have been found [7, 8, 9].

Although the “superadditivity” phenomenon is itself interesting, it is not desirable for communication purpose. If the capacity formula of a channel is superadditive, it means computing the capacity and finding the capacity-achieving input states get very hard. Without knowing such input states, one cannot find the optimal encoding strategy. If the capacity formula of a channel is additive (*i.e.*“single-letter”), then one only has to optimize over inputs of a single channel use, and this is tractable. Hence an important goal in quantum information theory is to characterize channels with additive capacities [10, 11, 12]. More recently, this has been extended to approximate channels by those with additive capacities [13].

There are also vast differences between classical and quantum communication, in terms of the auxiliary resources that can be used, and how they enhance the capacities. In the classical setting, it can be shown that pre-shared randomness does not increase the capacity of a channel. In the quantum setting, the most common resource is quantum entanglement. Unlike classical communication, pre-shared entanglement between the sender and receiver can enhance communication, with the most prominent example being superdense coding [14]. Remarkably, in the presence of unlimited pre-shared entanglement, the capacity formula of an arbitrary noisy channel becomes “single-letter” [15, 16]. Hence, with unlimited pre-shared entanglement assistance, quantum Shannon theory greatly simplifies.

However, in many cases, unlimited pre-shared entanglement between the sender and receiver can be unrealistic. Thus it makes sense to study the trade-off capacity region between entanglement and classical/quantum communication, or even among these three resources.

The first such work is given by Shor [17], who examined the case where only finite pre-shared entanglement is available and obtained a trade-off curve that illustrates how the optimal rate of classical communication depends on the amount of entanglement assistance (CE trade-off). Subsequently, many other trade-off capacities were studied [12, 18, 19, 20, 21].

This work aims to study the additivity properties of CE trade-off capacity, *i.e.* the classical capacity with limited entanglement assistance. This is important both from a theoretical point of view, as an extension of numerous former studies on additivity of the classical capacity [22, 7], and also from a practical point of view, since trade-off capacities naturally occur in future quantum communication.

The CE trade-off capacity formula given by Shor is “multi-letter”. This comes naturally, as the formula must reduce to the classical capacity formula when the entanglement assistance is set to zero. Hence, one does not expect it to be additive in general. This intuition can be made sharper, by considering the CE trade-off capacity of a channel with superadditive classical capacity. As the classical capacity must vary continuously with the

[*] eltonzhu@mit.edu
[†] quntao@mit.edu
[‡] shor@math.mit.edu

amount of entanglement, superadditivity will be retained when entanglement assistance is small.

What if we assume the classical capacity of a channel is additive? In this case, the picture is much unclearer. Intuitively, one expect the CE trade-off capacity to be additive. This comes from the fact that with unlimited pre-shared entanglement, the classical capacity is additive. So one would hope that limited entanglement assistance wouldn't complicate matters.

However, naively proving additivity of the CE trade-off capacity from its additive classical capacity fails for simple channels like the depolarizing channel, and entanglement-breaking channels. One of the reasons is that the trade-off capacity beats the time-sharing strategy. For CE trade-off, the time-sharing strategy is simply to distil the pre-shared entanglement into Bell pairs, and use them for classical communication over a fraction of the channel uses. For the other channel uses, one just performs classical communication without entanglement assistance. Since the trade-off protocol can beat this strategy, additivity of the CE trade-off capacity does not immediately follow from that of the classical capacity.

We show that this is simply not possible. There exist channels with a superadditive trade-off capacity, even when the classical capacity is additive. We show this by constructing an example. Our example is a switch channel, where part of the input acts as a switch register and determines which of the two sub-channels are used. One of the sub-channels is a classical channel. The second sub-channel is a quantum channel with a superadditive classical capacity. The classical channel has a larger classical capacity. Hence without entanglement assistance, the classical channel is always used and the capacity is additive. However, when pre-shared entanglement is available, the quantum channel is more favorable. Hence superadditivity develops. We also require the quantum channel to have its CE trade-off curve strictly concave, otherwise the trade-off protocol reduces to time-sharing. The whole argument is made precise in Ref.[23].

Our work implies that additivity is a very non-robust notion, and can be lost when resources start to trade. Even though unlimited pre-shared entanglement simplies quantum Shannon theory, limited entanglement can potentially complicate it.

Interestingly, the quantum capacity with limited entanglement assistance (QE trade-off) does not have this weird behavior, because its trade-off protocol does not beat the time-sharing strategy.

Recognizing the difficulty in obtaining an additive CE trade-off region in Shor's original framework, efforts have been made to give an additive CE trade-off capacity, by imposing a different constraint [24]. This can be used to bound an eavesdropper's information gain in two-way quantum key distribution protocols. The above switch channel framework can also be used to study the additivity property of other trade-off capacities [25].

References

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, July 1948.

[2] A. S. Holevo, "The capacity of the quantum channel with general signal states," *IEEE Transactions on Information Theory*, vol. 44, pp. 269–273, Jan 1998.

[3] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A*, vol. 56, pp. 131–138, Jul 1997.

[4] S. Lloyd, "Capacity of the noisy quantum channel," *Phys. Rev. A*, vol. 55, pp. 1613–1622, Mar 1997.

[5] P. W. Shor, "The quantum channel capacity and coherent information," *MSRI Workshop on Quantum Computation*, 2002.

[6] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Transactions on Information Theory*, vol. 51, pp. 44–55, Jan 2005.

[7] M. B. Hastings, "Superadditivity of communication capacity using entangled inputs," *Nat Phys*, vol. 5, pp. 255–257, 04 2009.

[8] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, "Quantum-channel capacity of very noisy channels," *Phys. Rev. A*, vol. 57, pp. 830–839, Feb 1998.

[9] G. Smith and J. A. Smolin, "Degenerate quantum codes for pauli channels," *Phys. Rev. Lett.*, vol. 98, p. 030501, Jan 2007.

[10] C. King, "Additivity for unital qubit channels," *Journal of Mathematical Physics*, vol. 43, no. 10, pp. 4641–4653, 2002.

[11] P. W. Shor, "Additivity of the classical capacity of entanglement-breaking quantum channels," *Journal of Mathematical Physics*, vol. 43, no. 9, pp. 4334–4340, 2002.

[12] I. Devetak and P. W. Shor, "The capacity of a quantum channel for simultaneous transmission of classical and quantum information," *Communications in Mathematical Physics*, vol. 256, no. 2, pp. 287–303, 2005.

[13] D. Sutter, V. B. Scholz, and R. Renner, "Approximate degradable quantum channels," in *2015 IEEE International Symposium on Information Theory (ISIT)*, pp. 2767–2771, June 2015.

[14] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on einstein-podolsky-rosen states," *Phys. Rev. Lett.*, vol. 69, pp. 2881–2884, Nov 1992.

[15] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted classical capacity of noisy quantum channels," *Phys. Rev. Lett.*, vol. 83, pp. 3081–3084, Oct 1999.

[16] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem," *IEEE Transactions on Information Theory*, vol. 48, pp. 2637–2655, Oct 2002.

[17] P. W. Shor, "The classical capacity achievable by a quantum channel assisted by a limited entanglement.," *Quantum Information and Computation*, vol. 4, no. 6, pp. 537 – 545, 2004.

[18] I. Devetak, A. W. Harrow, and A. J. Winter, "A resource framework for quantum shannon theory," *IEEE Transactions on Information Theory*, vol. 54, pp. 4587–4618, Oct 2008.

[19] M. M. Wilde and M.-H. Hsieh, "The quantum dynamic capacity formula of a quantum channel," *Quantum Information Processing*, vol. 11, no. 6, pp. 1431–1463, 2012.

[20] M.-H. Hsieh and M. M. Wilde, "Entanglement-assisted communication of classical and quantum information," *IEEE Transactions on Information Theory*, vol. 56, pp. 4682–4704, Sept 2010.

[21] M.-H. Hsieh and M. M. Wilde, "Trading classical communication, quantum communication, and entanglement in quantum shannon theory," *IEEE Transactions on Information Theory*, vol. 56, pp. 4705–4730, Sept 2010.

[22] P. W. Shor, "Equivalence of additivity questions in quantum information theory," *Communications in Mathematical Physics*, vol. 246, no. 3, pp. 453–472, 2004.

[23] E. Y. Zhu, Q. Zhuang, and P. W. Shor, "Superadditivity of the classical capacity with limited entanglement assistance," 2017.

[24] Q. Zhuang, E. Y. Zhu, and P. W. Shor, "Additive classical capacity of quantum channels assisted by noisy entanglement," *Phys. Rev. Lett.*, vol. 118, p. 200503, May 2017.

[25] E. Y. Zhu, Q. Zhuang, M.-H. Hsieh, and P. W. Shor, "Work in progress," 2017.

The quantum monad on relational structures

Samson Abramsky[1] * Rui Soares Barbosa[1] † Nadish de Silva[2] ‡ Ovtavio Zapata[2] §

[1] *Department of Computer Science, University of Oxford*
[2] *Department of Computer Science, University College London*

Abstract. A central theme in quantum computation is to show how quantum resources can be used to gain advantage in information processing tasks. In particular, non-local games have been used to exhibit quantum advantage in boolean constraint satisfaction, and to obtain quantum versions of graph invariants such as the chromatic number, and more broadly, of graph homomorphisms.

We introduce a general notion of non-local games for homomorphisms between relational structures, which play a central role in finite model theory, constraint satisfaction and database theory. We show how quantum strategies for such games can be viewed as Kleisli morphisms for a quantum monad on the (classical) category of relational structures and homomorphisms, removing the two-player non-local element of the game. We use these results to exhibit a wide range of examples of contextuality-powered quantum advantage, and to unify several apparently diverse strands of previous work. In particular, we spell out an equivalence between state-independent strong contextuality, quantum homomorphisms, and quantum advantage in constraint satisfaction, showing in particular that state-independent strong contextuality proofs can always be underwritten by non-locality arguments.

The full version of this paper is available from `arXiv:1705.07310[cs.LO]`.

Keywords: non-local games, quantum advantage, monads

Finite relational structures and the homomorphisms between them form a mathematical core common to finite model theory [10], constraint satisfaction [5], and relational database theory [9]. Moreover, much of graph theory can be formulated in terms of the existence of graph homomorphisms, as expounded e.g. in the influential text [6]. Thus, implicitly at least, the mathematical setting for all these works is categories of σ-structures and homomorphisms, for relational vocabularies σ.

What could it mean to quantize these structures? More precisely, with the advent of quantum computing, we can now consider the consequences of using quantum resources for carrying out various information-processing tasks. A major theme of current research is to delineate the scope of the *quantum advantage* which can be gained by the use of quantum resources. How can this be related to these fundamental structures?

Our starting point is the notion of *quantum graph homomorphism* introduced in [11] as a generalization of the notion of quantum chromatic number [2]. Consider the following game, played by Alice and Bob cooperating against a Verifier. Their goal is to establish the existence of a homomorphism $G \to H$ for given graphs G and H. Verifier provides vertices $v_1, v_2 \in V(G)$ to Alice and Bob respectively. They produce outputs $w_1, w_2 \in V(H)$ in response. No communication between Alice and Bob is permitted during the game. They win if the following conditions hold: $v_1 = v_2 \;\Rightarrow\; w_1 = w_2$ and $v_1 \sim v_2 \;\Rightarrow\; w_1 \sim w_2$, where we write $\sim$ for the adjacency relation.

If only classical resources are permitted, then the existence of a *perfect strategy* for Alice and Bob — one in which they win with probability 1 — is equivalent to the existence of a graph homomorphism in the standard sense. However, using quantum resources, in the form of an entangled bipartite state where Alice and Bob can each perform measurements on their part, there are perfect strategies in cases where no classical homomorphism exists, thus exhibiting quantum advantage.

Alice–Bob games have also been studied for other tasks, notably for *constraint systems.* Consider the following system of linear equations over $\mathbb{Z}_2$:

$$\begin{aligned} A \oplus B \oplus C &= 0 & B \oplus E \oplus H &= 0 \\ A \oplus D \oplus G &= 0 & G \oplus H \oplus I &= 0 \\ D \oplus E \oplus F &= 0 & C \oplus F \oplus I &= 1 \end{aligned}$$

Of course, this system is not satisfiable in the standard sense, as we can see by summing over the left- and right-hand sides. Now consider the following Alice–Bob game. The Verifier sends Alice an equation, and Bob a variable. Alice returns an assignment to the variables in the equation, and Bob returns an assignment for his variable. They win if Bob's assignment agrees with Alice's, and moreover Alice's assignment satisfies the given equation. Classically, the existence of a perfect strategy is equivalent to the existence of a satisfying assignment for the whole system. Using quantum resources, there is a perfect strategy for the above system, which corresponds to Mermin's "magic square" construction [12]. This can be generalized to a notion of quantum perfect strategies for a broad class of constraint systems [4, 3], which have strong connections both to the study of contextuality in quantum mechanics, and to a number of challenging mathematical questions [16, 15]. Clearly, these games are analogous to those for graph homomorphisms. What is the precise relationship?

In [11], generalizing results in [2], the existence of a quantum perfect strategy for the homomorphism game from G to H is characterized in terms of the existence of a family $\{E_{vw}\}_{v \in V(G), w \in V(H)}$ of projectors in

*`samsom.abramsky@cs.ox.ac.uk`
†`rui.soares.barbosa@cs.ox.ac.uk`
‡`nadish.desilva@utoronto.ca`
§`ocbzapata@gmail.com`

d-dimensional Hilbert space for some d, subject to certain conditions. Analogous results for constraint systems are proved in [4]. This characterization eliminates the two-person aspect of the game, and the shared state, leaving a "projector-valued relation" as the witness for existence of a quantum perfect strategy. We shall henceforth call these witnesses *quantum graph homomorphisms*. An important further step is taken in [11]. A construction $H \mapsto \mathsf{M}H$ on graphs is introduced, such that the existence of a quantum graph homomorphism from G to H is equivalent to the existence of a *standard* graph homomorphism $G \to \mathsf{M}H$.

Our contribution begins at this point. We describe a general notion of non-local game for witnessing homomorphisms between structures for any relational signature. We show that the use of quantum resources in these games can be characterized by a notion of *quantum homomorphism*, removing the two-player non-local and the state-dependent aspects of the game. Moreover, quantum homomorphisms can in turn be characterized as the Kleisli morphisms for a *quantum monad* on the (classical) category of relational structures and homomorphisms. Monads are used in computer science – particularly in functional programming and semantics of programming language – to express computational effects [14]. This monad is *graded* [13] by the dimension of the Hilbert space.

Our account refines and generalizes the ideas from both [2, 11] and [4]. We characterize quantum solutions for general constraint satisfaction problems, showing as a special case that these subsume the binary constraint systems of [4].

We also show how quantum witnesses for state-independent strong contextuality in the sense of [1] are characterized by quantum homomorphisms. This establishes a link between state-independent contextuality and non-locality, showing that that state-independent strong contextuality proofs can always be underwritten by non-locality arguments. This can be seen as a general form of constructions for turning Kochen–Specker contextuality proofs into Bell non-locality arguments [7]. The rôle of the entangled state and of Bob in the non-local game is to provide an operational or physical underpinning for the compatibility or generalized no-signalling assumption which is made for empirical models [1].

The precise relationship with the quantum graph homomorphisms of [11] turns out to be more subtle. By adapting a construction from [8], we show that their notion is characterized by a quantum solution in our sense for a related boolean constraint system. Overall, we show that a wide range of notions of quantum advantage is captured in a uniform way by the quantum monad, applied directly to the standard classical structures.

References

[1] Samson Abramsky and Adam Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(11):113036, 2011.

[2] Peter J. Cameron, Ashley Montanaro, Michael W. Newman, Simone Severini, and Andreas Winter. On the quantum chromatic number of a graph. *Electronic Journal of Combinatorics*, 14(1):R81, 2007.

[3] Richard Cleve, Li Liu, and William Slofstra. Perfect commuting-operator strategies for linear system games. *Journal of Mathematical Physics*, 58(1):012202, 2017.

[4] Richard Cleve and Rajat Mittal. Characterization of binary constraint system games. In Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias, editors, *Proceedings of the 41st International Colloquium on Automata, Languages, and Programming, Part I (ICALP 2014)*, pages 320–331. Springer, 2014.

[5] Tomás Feder and Moshe Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: A study through Datalog and group theory. *SIAM Journal on Computing*, 28(1):57–104, 1998.

[6] Pavol Hell and Jaroslav Nešetřil. *Graphs and homomorphisms*, volume 28 of *Oxford Lecture Series in Mathematics and Its Applications*. Oxford University Press, 2004.

[7] Peter Heywood and Michael L. G. Redhead. Nonlocality and the Kochen–Specker paradox. *Foundations of physics*, 13(5):481–499, 1983.

[8] Zhengfeng Ji. Binary constraint system games and locally commutative reductions, 2013. Available as arXiv:1310.3794 [quant-ph].

[9] Phokion G. Kolaitis and Moshe Y. Vardi. Conjunctive-query containment and constraint satisfaction. *Journal of Computer and System Sciences*, 61(2):302–332, 2000.

[10] Leonid Libkin. *Elements of finite model theory*. Texts in Theoretical Computer Science. Springer, 2004.

[11] Laura Mančinska and David E. Roberson. Quantum homomorphisms. *Journal of Combinatorial Theory, Series B*, 118:228–267, 2016.

[12] N. David Mermin. Simple unified form for the major no-hidden-variables theorems. *Physical Review Letters*, 65(27):3373–3376, Dec 1990.

[13] Stefan Milius, Dirk Pattinson, and Lutz Schröder. Generic trace semantics and graded monads. In Lawrence S. Moss and Pawel Sobocinski, editors, *6th Conference on Algebra and Coalgebra in Computer Science (CALCO 2015)*, volume 35 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 253–269. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2015.

[14] Eugenio Moggi. Notions of computation and monads. *Information and computation*, 93(1):55–92, 1991.

[15] Volkher B. Scholz and Reinhard F. Werner. Tsirelson's problem, 2008. Available as arXiv:0812.4305 [math-ph].

[16] William Slofstra. Tsirelson's problem and an embedding theorem for groups arising from non-local games, 2016. Available as arXiv:1606.03140 [quant-ph].

Converting multilevel nonclassicality into genuine multipartite entanglement

Bartosz Regula[1] * Marco Piani[2] Marco Cianciaruso[1] Thomas R. Bromley[1]
Alexander Streltsov[3] [4] Gerardo Adesso[1]

[1] *Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems (CQNE),*
School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom
[2] *SUPA and Department of Physics, University of Strathclyde, Glasgow G4 0NG, UK*
[3] *Faculty of Applied Physics and Mathematics, Gdańsk University of Technology, 80-233 Gdańsk, Poland*
[4] *National Quantum Information Center in Gdańsk, 81-824 Sopot, Poland*

Abstract. Characterizing genuine quantum resources and determining operational rules for their manipulation are crucial steps to appraise possibilities and limitations of quantum technologies. Two such key resources are nonclassicality, manifested as quantum superposition between reference states of a single system, and entanglement, capturing quantum correlations among two or more subsystems. Here we present a general formalism for the conversion of nonclassicality into multipartite entanglement, showing that a faithful reversible transformation between the two resources is always possible, within a precise resource-theoretic framework. Specializing to quantum coherence between the levels of a quantum system as an instance of nonclassicality, we introduce explicit protocols for such a mapping. We further show that the conversion relates multilevel coherence and multipartite entanglement not only qualitatively, but also quantitatively, restricting the amount of entanglement achievable in the process and in particular yielding an equality between the two resources when quantified by geometric measures.

1 Introduction

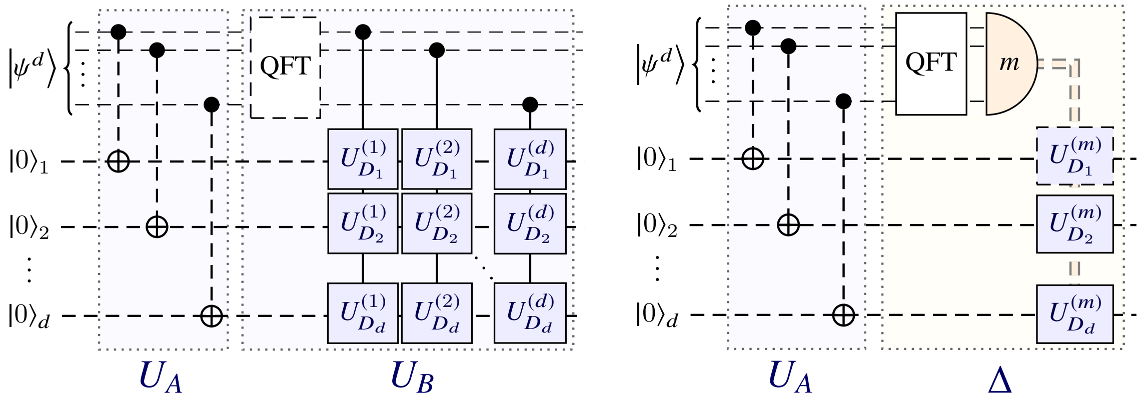
Signature features of the quantum world have been recently recognized as resources that can be harnessed for disruptive technologies [1]. One such resource, embodying the nonclassicality of quantum mechanics, is the possibility for a quantum system to exist in a *superposition* of "classical" states. The latter are usually determined based on physical considerations; for instance, in continuous-variable systems they can be identified with the Glauber-Sudarshan coherent states [2, 3], while in discrete-variable systems they can be taken to form a reference orthonormal basis (e.g. the energy eigenbasis), so that superposition manifests as quantum coherence [4–12].

Superposition underlies other nonclassical phenomena such as quantum entanglement among parts of a quantum system [13]. These two resources enjoy different uses in quantum technologies, and it thus becomes particularly relevant to investigate the connection between them beyond a merely conceptual standpoint, and to devise operational schemes that allow the dynamical transformation of one into the other. Several works have analyzed this problem. In quantum optics, nonclassicality gets mapped into entanglement by a beam splitter [14–18], while, in the discrete-variable scenario, it is the controlled NOT (CNOT) gate [19, 20] that plays a similar role. The quantitative interplay between the degree of nonclassicality and the bipartite entanglement obtained from it has been investigated as well [17, 21–26]. These studies have advanced our understanding of nonclassicality as a resource in systems of arbitrary dimension [7, 12, 25–31].

In this work [32], we show that there always exists a state-independent unitary mapping, realized by operations which alone cannot create nonclassicality, such that the presence of k-level nonclassicality in the state of a single d-level system is necessary and sufficient to create $k+1$-partite entanglement between the system and k ancillas. To exemplify such a conversion procedure, we specialize to quantum coherence as an instance of nonclassicality [12], and introduce an explicit physical protocol which directly converts k-level coherence into $k+1$-body multipartite entanglement. The protocol entangles a d-level system (qudit) with up to d qubits by a sequential application of generalized CNOT gates. The protocol can be further extended via the decoupling of the qudit system, realizable by either unitary or LOCC operations, to provide a mapping of k-coherence into multipartite entanglement of the ancillary qubits alone. This process can also be seen as a toy model for decoherence [33] due to the interaction with a many-body environment, with information about the superposition leaking into the environment in the form of multipartite entanglement.

Further, we explicitly show that the amount of k-coherence in the initial system places a quantitative restriction on the amount of entanglement that can be converted from it. In particular, the fidelity-based geometric measure of $k+1$-partite entanglement [34, 35] at the output of the protocol is exactly equal to the initial coherence of the qudit system, quantified by the fidelity-based geometric measure of k-coherence — a computable quantifier of multi-level coherence introduced here, extending previous work [7, 23].

*bartosz.regula@gmail.com



$|\psi^d\rangle$
$|0\rangle_1$
$|0\rangle_2$
$|0\rangle_d$
QFT
$U_{D_1}^{(1)}$ $U_{D_1}^{(2)}$ $U_{D_1}^{(d)}$
$U_{D_2}^{(1)}$ $U_{D_2}^{(2)}$ $U_{D_2}^{(d)}$
$U_{D_d}^{(1)}$ $U_{D_d}^{(2)}$ $U_{D_d}^{(d)}$
U_A
U_B
m
$U_{D_1}^{(m)}$
$U_{D_2}^{(m)}$
$U_{D_d}^{(m)}$
Δ

(a) (b)

Figure 1: Schemes of two protocols to convert k-coherence into multipartite entanglement. Both protocols begin with the global unitary operation U_A which sequentially entangles each level of the qudit system in the state $|\psi^d\rangle$ with a corresponding ancillary qubit by generalized CNOT gates, resulting in a $k+1$-partite entangled state. One can then decouple the qudit system either (a) by a unitary transformation U_B, consisting of a Fourier transform and a disentangling unitary U_D, or (b) via a one-way LOCC operation Δ. Both protocols result in genuine k-partite entanglement between the ancillary qubits.

2 Nonclassicality converison

The nonclassicality of a state ρ is a notion that depends on our chosen set of states that we take to be "classical". Choosing a finite set of states $\{|\chi_i\rangle\}$ (not necessarily orthogonal) which spans the whole Hilbert space $\mathcal{H}$ to constitute the pure classical states, one asks whether ρ can be represented as a convex combination of classical states only. If this is not possible — that is, if one has to consider superpositions of $\{|\chi_i\rangle\}$ — then ρ is a nonclassical state. Therefore, the set of all classical states $\mathcal{C}$ is formed by the convex hull of $\{|\chi_i\rangle\}$.

This formalism leads to a natural measure of the level of nonclassicality of a state. For a pure state, one can indeed define the *nonclassical rank* ($\mathrm{R_N}$) [21, 25]: $\mathrm{R_N}(|\psi\rangle) = \min\left\{r \;\middle|\; |\psi\rangle = \sum_{i=1}^{r} c_i |\chi_i\rangle\,,\ |\chi_i\rangle \in \mathcal{C}\right\}$ with nonzero complex coefficients c_i. This clearly resembles the definition of the Schmidt rank $\mathrm{R_S}(|\psi\rangle)$ of bipartite entangled states, and it can in fact be extended to mixed states in the same way as the latter is extended to the Schmidt number $\mathrm{N_S}(\rho)$ [36]. We thus define the *nonclassical number* ($\mathrm{N_N}$) as $\mathrm{N_N}(\rho) = \min_{\{p_i,|\psi_i\rangle\}} \max_i\ \mathrm{R_N}(|\psi_i\rangle)$ where the minimization is performed over all pure-state convex decompositions of ρ into $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$.

Killoran et al. [25] showed that in this formalism there always exists an isometry, consisting of adding an ancilla system and applying a global unitary operation, which maps a pure state of nonclassical rank k into a bipartite entangled pure state of Schmidt rank k. As one of the main results of this paper, we show that an analogous faithful conversion of multilevel nonclassicality into genuine multipartite entanglement is always possible.

Theorem 1 *Let $\mathcal{H}$ be a d-dimensional Hilbert space, and $\mathcal{H}_{anc}$ the Hilbert space of an ancillary system. Then if the classical pure states $\{|\chi_i\rangle\}_{i=1}^{d} \in \mathcal{H}$ form a linearly independent set spanning $\mathcal{H}$, there exists an isometry $\Lambda : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}_{anc}^{\otimes d}$ such that for any state $\rho \in \mathcal{D}(\mathcal{H})$ with nonclassical number $\mathrm{N_N}(\rho) = k$, $\Lambda\rho\Lambda^\dagger$ is genuinely $k+1$-partite entangled iff ρ is nonclassical ($2 \le k \le d$) and $\Lambda\rho\Lambda^\dagger$ is fully separable iff ρ is classical ($k = 1$).*

Theorem 1 shows that one can always faithfully map the k-level nonclassicality of a quantum system in two ways, but we note that the specifics of the mappings are not fixed by the theorems — in particular, this means that one can construct different mappings which map multilevel nonclassicality into qualitatively different types of entanglement. We give two examples of such mappings: one which only use two levels of each ancillary system, resulting in entanglement akin to that of W states [37], and another protocol which instead uses qudit ancillary systems and generates a generalized GHZ-type entanglement between the qudits. However, the choice of a W-type mapping in the theorem makes the conversion quite appealing in practice, as it only requires qubit ancillas, and, as we show below, enables one to create entanglement by a sequential application of two-body gates on the nonclassical system and each ancilla.

3 Coherence conversion

We will now specialize to the framework of quantum coherence [4, 7, 12]. Here, the classical states $\{|i\rangle\}_{i=1}^{d}$ are taken to form an orthonormal basis for $\mathcal{H}$. Analogously to nonclassicality, we can then define a hierarchy of coherence levels by considering the *coherence rank* $\mathrm{R_C}(|\psi\rangle)$, defined to be the number of non-zero coefficients c_i that a state $|\psi\rangle = \sum_i c_i |i\rangle$ has in this basis [5, 6], and extending it to mixed states as the *coherence number* $\mathrm{N_C}(\rho)$.

The k-coherence of a single qudit can be converted into multipartite entanglement in different physical ways. To show this, we design a protocol to convert k-coherence into $k+1$-partite entanglement between the qudit and k qubit ancillas (following Thm. 1), realizable by a sequential application of CNOT gates (see U_A in Fig. (a) and (b)).

We then provide a natural mapping of k-coherence into k-body entanglement, which can be accomplished by a second step which disentangles the qudit system — either by unitary transformations as in Fig. (a), or by one-way LOCC Δ as in Fig. (b). The LOCC protocol might lend itself to a more efficient implementation as it does not require global interactions. It also reflects an operational scenario in which input agents are constrained to the resource theory of k-coherence, having at disposal only incoherent ancillas and incoherent operations as used in the first step, while output agents are constrained to the resource theory of entanglement, being bound to use LOCC as in the second step.

Quantification. — In any resource theory, one can define a faithful class of quantifiers by considering the distance to the set of non-resource states [38, 39]. In the cases of bipartite entanglement and standard coherence (i.e., 2-coherence in our framework), the corresponding non-resource sets are the sets of separable states $\mathcal{S}$ and incoherent states $\mathcal{I}$, respectively [7, 40]. For the case of k-partite entanglement, one can define the non-resource set as the set of $k-1$-producible states $\mathcal{P}^{(k-1)}$ [41], i.e., states which are at most $k-1$-partite entangled. Similarly for k-coherence, we consider the set $\mathcal{C}^{(k-1)}$ of states which are at most $k-1$-coherent. We can then define quantifiers of k-partite entanglement $E_D^{(k)}$ and k-level coherence $C_D^{(k)}$, for which we obtain the following result.

Theorem 2 *Let D be any distance contractive under* CPTP *and G denote the choice of distance $1-F(\rho,\sigma)$. Given the protocol which converts the k-coherence of a state ρ into $k+1$-partite entanglement of $\rho' = U_A \rho U_A^\dagger$ or k-partite entanglement of $\rho'' = \Delta(\rho')$, we get:*

$$C_D^{(k)}(\rho) \geq E_D^{(k+1)}(\rho') \tag{1}$$

$$C_D^{(k)}(\rho) \geq E_D^{(k)}(\rho'') \tag{2}$$

$$C_G^{(k)}(\rho) = E_G^{(k+1)}(\rho') \tag{3}$$

The amount of k-coherence present in the initial state thus places quantitative constraints on the multipartite entanglement one can obtain from it. Remarkably, under the fidelity-based geometric quantifiers, the k-coherence of any system and the converted $k+1$-partite entanglement are actually *equal*, and the simple properties and computability of the geometric measure of k-coherence mean that the geometric measure of $k+1$-partite entanglement can be efficiently computed for states ρ' obtained from the conversion protocol.

4 Conclusions

Our work reveals a qualitative and quantitative connection between multilevel nonclassicality and multipartite entanglement, generalizing previous results in the resource theory of quantum coherence [7, 12, 23], and further contributing towards the formalization of nonclassicality as a resource [25, 26, 31]. In particular, by proving the convertibility of the two resources in general multipartite settings, the results provide on one hand a further advance towards establishing a unified framework for the quantification of fundamental quantum phenomena as resources, and reveals on the other hand feasible protocols to interchange such resources experimentally to realize efficient hybrid approaches to quantum technologies.

References

[1] J. P. Dowling and G. J. Milburn, Philos. T. Roy. Soc. A **361**, 1655 (2003).

[2] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).

[3] E. C. G. Sudarshan, Phys. Rev. Lett. **10**, 277 (1963).

[4] J. Aberg, (2006), arXiv:quant-ph/0612146 .

[5] B. Witt and F. Mintert, New J. Phys. **15**, 093020 (2013).

[6] F. Levi and F. Mintert, New J. Phys. **16**, 033007 (2014).

[7] T. Baumgratz, M. Cramer, and M. B. Plenio, Phys. Rev. Lett. **113**, 140401 (2014).

[8] T. R. Bromley, M. Cianciaruso, and G. Adesso, Phys. Rev. Lett. **114**, 210401 (2015).

[9] A. Winter and D. Yang, Phys. Rev. Lett. **116**, 120404 (2016).

[10] C. Napoli, T. R. Bromley, M. Cianciaruso, M. Piani, N. Johnston, and G. Adesso, Phys. Rev. Lett. **116**, 150502 (2016).

[11] E. Chitambar and G. Gour, Phys. Rev. Lett. **117**, 030401 (2016).

[12] A. Streltsov, G. Adesso, and M. B. Plenio, (2016).

[13] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

[14] M. S. Kim, W. Son, V. Bužek, and P. L. Knight, Phys. Rev. A **65**, 032323 (2002).

[15] X.-b. Wang, Phys. Rev. A **66**, 024303 (2002).

[16] M. M. Wolf, J. Eisert, and M. B. Plenio, Phys. Rev. Lett. **90**, 047904 (2003).

[17] J. K. Asbóth, J. Calsamiglia, and H. Ritsch, Phys. Rev. Lett. **94**, 173602 (2005).

[18] J. S. Ivan, S. Chaturvedi, E. Ercolessi, G. Marmo, G. Morandi, N. Mukunda, and R. Simon, Phys. Rev. A **83**, 032118 (2011).

[19] A. Streltsov, H. Kampermann, and D. Bruß, Phys. Rev. Lett. **106**, 160401 (2011).

[20] M. Piani, S. Gharibian, G. Adesso, J. Calsamiglia, P. Horodecki, and A. Winter, Phys. Rev. Lett. **106**, 220403 (2011).

[21] J. Sperling and W. Vogel, Phys. Scr. **90**, 074024 (2015).

[22] W. Vogel and J. Sperling, Phys. Rev. A **89**, 052302 (2014).

[23] A. Streltsov, U. Singh, H. S. Dhar, M. N. Bera, and G. Adesso, Phys. Rev. Lett. **115**, 020403 (2015).

[24] W. Ge, M. E. Tasgin, and M. S. Zubairy, Phys. Rev. A **92**, 052328 (2015).

[25] N. Killoran, F. E. S. Steinhoff, and M. B. Plenio, Phys. Rev. Lett. **116**, 080402 (2016).

[26] T. Theurer, N. Killoran, D. Egloff, and M. B. Plenio, (2017), arXiv:1703.10943 .

[27] F. G. S. L. Brandão and M. B. Plenio, Nat. Phys **4**, 873 (2008).

[28] Y.-R. Zhang, L.-H. Shao, Y. Li, and H. Fan, Phys. Rev. A **93**, 012334 (2016).

[29] J. Xu, Phys. Rev. A **93**, 032111 (2016).

[30] K. C. Tan, T. Volkoff, H. Kwon, and H. Jeong, (2017), arXiv:1703.01067 .

[31] C. Mukhopadhyay, S. Das, S. Bhattacharya, A. Sen De, and U. Sen, arXiv:1705.04343 [quant-ph] (2017), arXiv:1705.04343 [quant-ph] .

[32] B. Regula, M. Piani, M. Cianciaruso, T. R. Bromley, A. Streltsov, and G. Adesso, (2017), arXiv:1704.04153 .

[33] W. H. Zurek, Rev. Mod. Phys. **75**, 715 (2003).

[34] T.-C. Wei and P. M. Goldbart, Phys. Rev. A **68**, 042307 (2003).

[35] A. Streltsov, H. Kampermann, and D. Bruß, New J. Phys. **12**, 123004 (2010).

[36] B. M. Terhal and P. Horodecki, Phys. Rev. A **61**, 040301 (2000).

[37] W. Dür, G. Vidal, and J. I. Cirac, Phys. Rev. A **62**, 062314 (2000).

[38] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, 2007).

[39] G. Adesso, T. R. Bromley, and M. Cianciaruso, J. Phys. A: Math. Theor. **49**, 473001 (2016).

[40] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275 (1997).

[41] O. Gühne, G. Tóth, and H. J. Briegel, New J. Phys. **7**, 229 (2005).

Simultaneous hollowisation separability criterion in general multipartite systems

A. Neven[1] *   T. Bastin[1]

[1] *Institut de Physique Nucléaire, Atomique et de Spectroscopie, CESAM Research Unit, University of Liège, Liège 4000, Belgium*

Abstract. We use the generalized concurrence approach to investigate the general multipartite separability problem. To this aim, we first show how to generate a set containing all the independent generalized concurrences for any multipartite system. Then, by extending the preconcurrence matrix formalism to these systems, we show that the separability problem is equivalent to a pure matrix analysis problem that consists in determining whether a set of given symmetric matrices is simultaneously unitarily congruent to hollow matrices, i.e., to matrices whose main diagonal is composed only of zeroes.

Keywords: Quantum entanglement, separability criteria, generalized concurrences.

Quantum entanglement is at the heart of quantum mechanics and intimately linked to its nonlocal feature [1]. It is a key resource in many promising applications, like, to cite a few, quantum cryptography [2], quantum communication [3], quantum imaging [4], or also quantum sensing [5]. In this context, the ability to distinguish both experimentally and theoretically between entangled and separable states is a crucial issue. Theoretically, this issue is entirely solved in the pure state case where general and practical necessary and sufficient separability criteria have been identified (see, e.g., Ref. [6]). For mixed states, the question is much more involved and remains open in the very general case. Still various necessary but not sufficient conditions of separability have been stated [1, 7], such as the positive partial transpose (PPT) criterion [8], combinatorially independent permutation criteria [9, 10], Bell-type inequalities [11], or criteria based on entanglement witnesses [12, 13]. In some restricted cases, some of these above-cited criteria turn to be also sufficient conditions of separability. This happens for example for the PPT criterion in low-dimensional or low-rank cases, such as for qubit-qubit or qubit-qutrit systems [12], for $\mathbb{C}^m \otimes \mathbb{C}^n (m \leq n)$ bipartite states with rank at most n [14], or even for general multipartite mixed states with rank at most 3 [15].

The concurrence [16] is another tool that proved to provide a necessary and sufficient condition (NSC) of separability in 2-qubit systems. It is defined for pure states $|\psi\rangle$ as

$$C(\psi) \equiv |\langle\psi|S|\psi^*\rangle|, \tag{1}$$

where $S = \sigma_y \otimes \sigma_y$ is the 2-qubit spin-flip operator with σ_y the second Pauli matrix and where $|\psi^*\rangle$ is the complex conjugate of $|\psi\rangle$ expressed in the computational basis. For mixed states ρ, the concurrence is defined via the standard convex-roof construction :

$$C(\rho) = \inf_{\{p_i,|\psi_i\rangle\}} \sum_i p_i C(\psi_i), \tag{2}$$

where the infinimum is computed over all possible decompositions of ρ, i.e., all sets $\{p_i, |\psi_i\rangle\}$ such that $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. The concurrence is an entanglement measure that vanishes only for separable states [16] and this provides an easy of separability : a state ρ is separable if and only if $C(\rho) = 0$. In general, the minimization implied by convex-roofs is a very challenging task. However, in the case of the concurrence, Eq. (2) simplifies to [16]

$$C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}, \tag{3}$$

with λ_i $(i = 1, \ldots, 4)$ the square roots of the eigenvalues of $\rho S \rho^* S$ sorted in decreasing order.

The 2-qubit concurrence has been generalized to more general bipartite [17] or even multipartite [6] systems by the introduction of a vector of generalized concurrences C_α defined similarly as in Eq.(1) but each with a specific generalized "spin-flip" operator S_α [17, 6]. The cancellation of all the generalized concurrences still provides an NSC of separability, however only for pure states. Though the extension to mixed states via the convex-roof construction yields a similar elegant result as in Eq. (3) for each C_α [17, 6], the cancellation of each of them only provides a necessary separability condition for mixed states [17, 6]. Here, we show that the missing element to get a necessary and sufficient condition of separability based on generalized concurrences can be formalized equivalently as a pure matrix analysis problem that consists in *determining whether a set of given symmetric matrices is simultaneously unitarily congruent to hollow matrices, i.e., to matrices whose main diagonal is composed only of zeroes.*

To this aim, we first refine the NSC of separability based on generalized concurrences for pure states by showing how to get an optimal non-redundant set of generalized "spin-flip" operators S_α for arbitrary multipartite systems. The generalized "spin-flip" operators S_α introduced in Refs. [6, 17] are generated either from tensor products of SO(n) generators [6] or from 2×2 minor equations from tensor matricizations [17, 18]. Both methods unfortunately produce highly redundant sets of operators. Here, we show how to extract from them the only independent operators. For this purpose we make use of the 2×2 minor equations method [17], that is better

*a.neven@ulg.ac.be

suited for this task.

We then extend the concept of preconcurrence matrices [19] to these independent operators and address the mixed state case. From there, we can prove our main result, which is to show that a general mixed states is separable if and only if all its preconcurrence matrices (which are complex symmetric matrices) are *simultaneously hollowisable*, i.e. simultaneously unitarily congruent to hollow matrices. In other words, we show that the separability problem is equivalent to the pure matrix analysis problem of finding whether a set of symmetric matrices is simultaneously hollowisabe or not.

Although related topics such as simultaneous unitary congruence of pairs of complex matrices have already been studied in the literature [20], very few seems to be known about the simultaneous hollowisation problem. The problem of simultaneous diagonalisation of symmetric matrices, which in a sense can be seen as the opposite problem to simultaneous hollowisation, is by contrast well known and can be solved using a simple commutation criterion [21].

With this mathematical reformulation of the separability problem, we wish to draw the attention of both the matrix analysis and the entanglement detection communities to the problem of simultaneous hollowisation. Progress on this topic could indeed provide interesting headway in the field of entanglement detection. This problem may be hard to solve but we would like to point that partial answers to the problem can already lead to separability criteria for particular classes of states. We illustrate this in the last part of our paper by showing that a criterion for simultaneous hollowisability of 2 × 2 symmetric matrices can be translated into a separability criterion for general mixed states of rank 2.

To do so, we first use theorem 1 from Ref. [22] to get the hollowisability condition for symmetric 2×2 matrices. With this condition, we can prove that a set of symmetric 2×2 symmetric matrices are simultaneously hollowisable if and only if these matrices are individually hollowisable and proportional to each other. Combined with the preconcurrence matrix formalism developed earlier, this simultaneous hollowisability condition can be used to prove that any rank 2 state is separable if and only if all its preconcurrence matrices are hollowisable and proportional to each other. As we already mentioned, the PPT criterion is also an NSC of separability for rank 2 states [15]. The criterion involving the preconcurrence matrices has however the advantage to provide a separable decomposition for separable states, directly computed from the preconcurrence matrices.

In conclusion, we show that the simultaneous hollowisation problem constitutes a new approach to the separability problem. Using this approach, we obtain an NSC of separability for general multipartite states of rank 2 that is independent from the PPT criterion (which is also an NSC of separability in that case). We hope that this reformulation will stimulate further research to find practical criteria for simultaneous hollowisation, which could lead to new NSC of separability.

References

[1] R. Horodecki, P. Horodecki, Horodecki M., and K. Horodecki. *Rev. Mod. Phys.*, **81**, 865 (2009).

[2] A. K. Ekert. *Phys. Rev. Lett.*, **67**, 661 (1991).

[3] R. Ursin *et al. Nat. Phys.*, **3**, 481 (2007).

[4] A. F. Abouraddy, B. E. Saleh, A. V. Sergienko, and M. C. Teich. *Phys. Rev. Lett.*, **87**, 123602 (2001).

[5] W. Wieczorek, R. Krischek, N. Kiesel, Ch. Schmid, and H. Weinfurter. *Proc. SPIE 7608, Quantum Sensing and Nanophotonic Devices VII*, 76080P (2010).

[6] C.-S. Yu and H.-S. Song. *Phys. Rev. A*, **73**, 022325 (2006).

[7] O. Gühne and Tóth. *Phys. Rep.*, **474**, 1-75 (2009).

[8] A. Peres. *Phys. Rev. Lett.*, **77**, 1413 (1996).

[9] M. Horodecki, P. Horodecki, and Horodecki R. *Open Syst. Inf. Dyn.*, **13**, 103 (2006).

[10] P. Wocjan and M. Horodecki. *Open Syst. Inf. Dyn.*, **12**, 331 (2005).

[11] M. Seevinck and J. Uffink. *Phys. Rev. A*, **78**, 032101 (2008).

[12] M. Horodecki, P. Horodecki, and R. Horodecki. *Phys. Lett. A*, **223**, 1 (1996).

[13] B. M. Terhal. *J. Theor. Comp. Science*, **287**, 313 (2002).

[14] P. Horodecki, M. Lewenstein, G. Vidal, and I. Cirac. *Phys. Rev. A*, **62**, 032310 (2000).

[15] L. Chen and D. Đokovic. *J. Phys. A: Math. Theor.*, **46**, 275304 (2013).

[16] W. K. Wootters. *Phys. Rev. Lett.*, **80**, 2245 (1998).

[17] K. Audenaert, F. Verstraete, and B. De Moor. *Phys. Rev. A*, **64**, 052304 (2001).

[18] J. Gillet. *Contribution to Entanglement Theory, Applications in Atomic Systems and Cavity QED.* PhD thesis, Université de Liège, 2011.

[19] P. Badziąg, P. Deuar, M. Horodecki, P. Horodecki, and R. Horodecki. *J. Mod. Opt.*, **49**, 1289 (2002).

[20] T. G. Gerasimova, R. A. Horn, and V. V. Sergeichuk. *Linear Algebra Appl.*, **438**, 3829-3835 (2013).

[21] R. A. Horn and C. R. Johnson. *Matrix analysis.* Cambridge University Press, second edition, 2013.

[22] R. C. Thompson. *Linear Algebr. Appl.*, **26**, 65 (1979).

Towards high-dimensional entanglement-based quantum communication in space

Fabian Steinlechner[1] * Sebastian Ecker[1] Matthias Fink[1] Bo Liu[1] Oliver de Vries[2]
Jessica Bavaresco[1] Marcus Huber[1] Erik Beckert[2] Thomas Scheidl[1] Rupert Ursin[1] †

[1]*Institute for Quantum Optics and Quantum Information (IQOQI)*
Austrian Academy of Sciences, Boltzmanngasse 3, A-1090 Vienna, Austria
[2]*Fraunhofer Institute for Applied Optics and Precision Engineering*
Albert-Einstein-Str.7, 07745 Jena, Germany

Abstract. Quantum entanglement is a fundamental resource in quantum information processing and its generation, manipulation and distribution between distant parties are all key challenges in the pursuit of global quantum communications. Increasing the dimensionality of entanglement has been shown to improve robustness and channel capacities in secure quantum communications. Here, we give an overview of our efforts towards exploiting high-dimensional entanglement in long-distance quantum communication. We report on the results a first feasibility study, in which we distribute genuine 4-dimensional hyperentanglement via a free-space link over the rooftops of Vienna. We discuss how this approach could be extended to applications such as large-alphabet quantum key distribution in space and conclude with a brief update on our progress in engineering a sutiable space-proof-entangled photon source.

Keywords: Quantum entanglement, Optical communication, Quantum cryptography, Quantum optics

Long-distance quantum communication with high-dimensional entanglement

The distribution of quantum information over long distances is a key challenge in the pursuit of global quantum communication. Optical satellite links allow overcoming the distance limitations of fiber-based transmission on ground and could thus play a central role in future quantum communication networks. The viability of this approach is backed by a long history of long-distance quantum optics experiments over terrestrial free-space links where the losses and atmospheric turbulence were similar (or worse) than for optical satellite links [1]. Free-space quantum communication has now reached a level of maturity that is most markedly reflected in the recent launch of dedicated quantum communication satellites [2–4]. Despite these remarkable developments, experiments in this field still focus on two-level photonic systems. Specifically, polarization qubits have been the system of choice for free-space quantum communications for over a decade.

High-dimensional entanglement and hyperentanglement have both been shown to improve the security and channel capacity quantum communications [5,6] and have been successfully exploited in the realization of advanced quantum information processing protocols in a laboratory setting, such as quantum teleportation of multiple degrees of freedom [7], quantum dense coding with increased channel capacity [8], and efficient entanglement purification [9,10]. Consequently, increasing the dimensionality of entangled quantum systems can be considered a key technological step towards the realization of more practical protocols in real world free-space link scenarios ultimately also linking to and from space.

In a recent feasibility study [11], we made a first step towards implementing advanced high-dimensional quantum communication protocols in long-distance free-space links. We used hyperentanglement – that is a quantum state entangled not only in polarization but also the arrival time of the photons – to realize a large state space [12] and distribute high-dimensional entanglement via an intra-city free-space link in turbulent atmosphere (Fig. 1). In order to verify the integrity of the atmospheric quantum communication channel for hyperentangled photons, we experimentally certified entanglement in both polarization and energy-time subspaces individually, as well as genuine 4-dimensional entanglement with a Bell-state Fidelity of 0.9419. The hyperentangled state of the entire system could thus be used to transmit 1.4671 ebits of entanglement of formation. Note, however, that the potential dimensionality of energy-time entanglement is orders of magnitudes larger than demonstrated in this first feasibility study. Future setups for free-space experiments could use several unbalanced interferometers, or additional photonic degrees of freedom to greatly increase the dimensionality and with it the resistance to inevitable background noise.

The coherent transmission of quantum information embedded in a genuine high-dimensional state space under real-world link conditions represents an important step towards long-distance quantum communications with more complex quantum systems. The methodology of implementation, as well as the remarkably high-transmission fidelity and pair-detection rates demonstrated in our proof-of-concept experiment, make the approach highly suitable for the exploitation of such states in existing proposals for satellite experiments with polarization-entangled photons. This could significantly extend the scope of future experiments in space: The additional possibility of analyzing high-dimensional energy-time entanglement not only allows for larger information

*fabian.steinlechner@oeaw.ac.at
†rupert.ursin@oeaw.ac.at

capacity in quantum communication, but could provide a platform for entirely new fundamental physics experiments, such as the evaluation of models for gravity-induced wave function collapse [13] or quantum information processing in a relativistic framework. We thus hope that our results will motivate both further theoretical research into energy-time entanglement experiments conceivable at relativistic scenarios with satellite links, as well as experimental research into the exploitation of hyperentanglement in long-distance free-space quantum communications.

Engineering a space-proof entangled photon source

The development of space-suitable sources and detection hardware represents another major challenge in the pursuit of global-scale quantum communication with satellite links. Robust and efficient entangled photon sources are not only a vital pre-requesite towards implementing advanced quantum protocols in space, but can also enable other challenging experiments on ground [14].

In collaboration with the Fraunhofer Instiut Jena, we are currently developing a power-efficient prototype entangled photon source (EPS) that can sustain the strong vibrations and thermal fluctuations of space flight and operation in space. We outline the main factors which led to the baseline optical design and opto-mechanical implementation (Fig. 2), as well as preliminary results on performance characterization and environmental testing of the EPS. We discuss some of the main challenges that are still to be addressed, such as further integreation, as well as sources with tailored spectral properties, such as ultra-narrowband or pulsed sources (e.g. for multi-photon experiments) and ultra-broadband sources (with strong correlations in time e.g. for clock synchronization).

References

[1] R Ursin, F Tiefenbacher, T Schmitt-Manderbach, H Weier, T Scheidl, M Lindenthal, B Blauensteiner, T Jennewein, J Perdigues, P Trojek, M Furst, M Meyenburg, J Rarity, Z Sodnik, C Barbieri, H Weinfurter, and A Zeilinger. Entanglement-based quantum communication over 144 km. *Nat Phys*, 3(7):481–486, July 2007.

[2] Juan Yin, Yuan Cao, Yu-Huai Li, Sheng-Kai Liao, Liang Zhang, Ji-Gang Ren, Wen-Qi Cai, Wei-Yue Liu, Bo Li, Hui Dai, et al. Satellite-based entanglement distribution over 1200 kilometers. *Science*, 356(6343):1140–1144, 2017.

[3] Hideki Takenaka, Alberto Carrasco-Casado, Mikio Fujiwara, Mitsuo Kitamura, Masahide Sasaki, and Morio Toyoshima. Satellite-to-ground quantum-limited communication using a 50-kg-class microsatellite. *Nature Photonics*, 2017.

[4] Zhongkan Tang, Rakhitha Chandrasekara, Yue Chuan Tan, Cliff Cheng, Kadir Durak, and Alexander Ling. The photon pair source that survived a rocket explosion. *Scientific reports*, 6, 2016.

[5] Nicolas J Cerf, Mohamed Bourennane, Anders Karlsson, and Nicolas Gisin. Security of quantum key distribution using d-level systems. *Physical Review Letters*, 88(12):127902, 2002.

[6] Irfan Ali-Khan, Curtis J Broadbent, and John C Howell. Large-alphabet quantum key distribution using energy-time entangled bipartite states. *Physical review letters*, 98(6):060503, 2007.

[7] Xi-Lin Wang, Xin-Dong Cai, Zu-En Su, Ming-Cheng Chen, Dian Wu, Li Li, Nai-Le Liu, Chao-Yang Lu, and Jian-Wei Pan. Quantum teleportation of multiple degrees of freedom of a single photon. *Nature*, 518(7540):516–519, 2015.

[8] Julio T Barreiro, Tzu-Chieh Wei, and Paul G Kwiat. Beating the channel capacity limit for linear photonic superdense coding. *Nature physics*, 4(4):282–286, 2008.

[9] Christoph Simon and Jian-Wei Pan. Polarization entanglement purification using spatial entanglement. *Physical review letters*, 89(25):257901, 2002.

[10] Yu-Bo Sheng and Fu-Guo Deng. One-step deterministic polarization-entanglement purification using spatial entanglement. *Physical Review A*, 82(4):044305, 2010.

[11] Fabian Steinlechner, Sebastian Ecker, Matthias Fink, Bo Liu, Jessica Bavaresco, Marcus Huber, Thomas Scheidl, and Rupert Ursin. Distribution of high-dimensional entanglement via an intra-city free-space link. *Nature Communications*, 8(15971), 2017.

[12] Paul G. Kwiat. Hyper-entangled states. *Journal of Modern Optics*, 44(11-12):2173–2184, 1997.

[13] Siddarth Koduru Joshi, Jacques Pienaar, Timothy C Ralph, Luigi Cacciapuoti, Will McCutcheon, John Rarity, Dirk Giggenbach, Vadim Makarov, Ivette Fuentes, Thomas Scheidl, et al. Space quest mission proposal: Experimentally testing decoherence due to gravity. *Preprint at http://arxiv.org/abs/1703.08036*, 2017.

[14] Matthias Fink, Ana Rodriguez-Aramendia, Johannes Handsteiner, Abdul Ziarkash, Fabian Steinlechner, Thomas Scheidl, Ivette Fuentes, Jacques Pienaar, Timothy C Ralph, and Rupert Ursin. Experimental test of photonic entanglement in accelerated reference frames. *Nature Communications*, 8(15304), 2017.

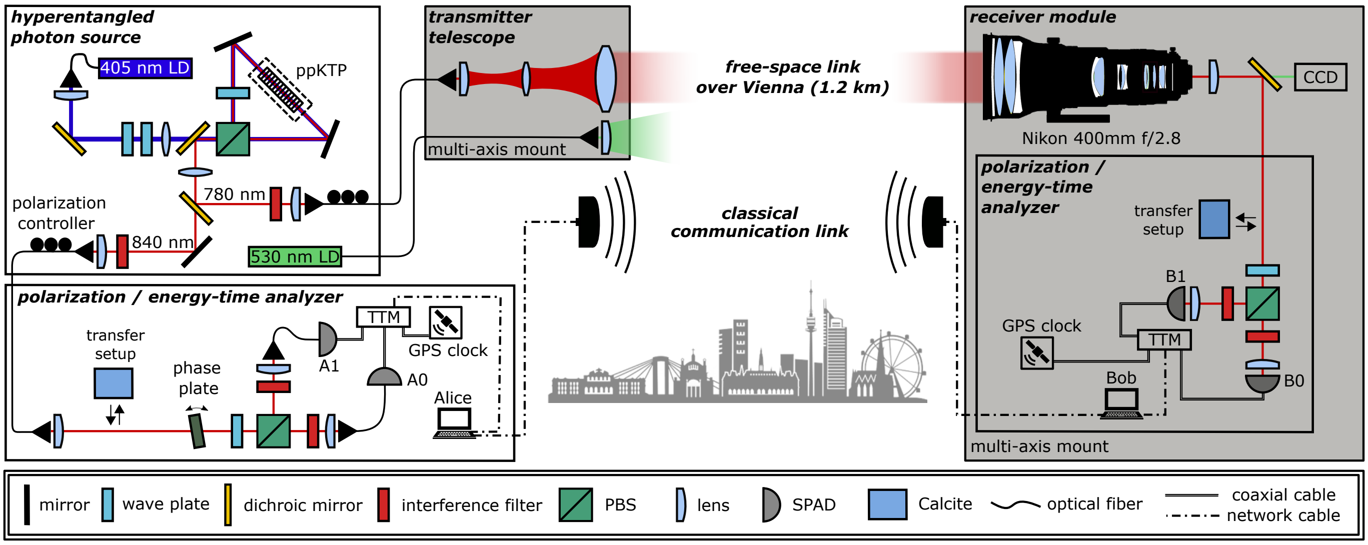


Figure 1: **High-dimensional entanglement distribution over the rooftops of Vienna.** An ultra-bright hyperentangled photon source was located in a laboratory at the Institute for Quantum Optics and Quantum Information Vienna (IQOQI). The source utilized spontaneous parametric down-conversion to produces polarization/energy-time hyperentangled photon pairs. The photons were distributed to Alice and Bob via a free-space link and and optical single-mode fiber, respectively. Bob's photons were collected using a telephoto objective and guided to a polarization detection module. A polarization-dependent delay was implemented for Franson interference measurements in the energy-time basis.

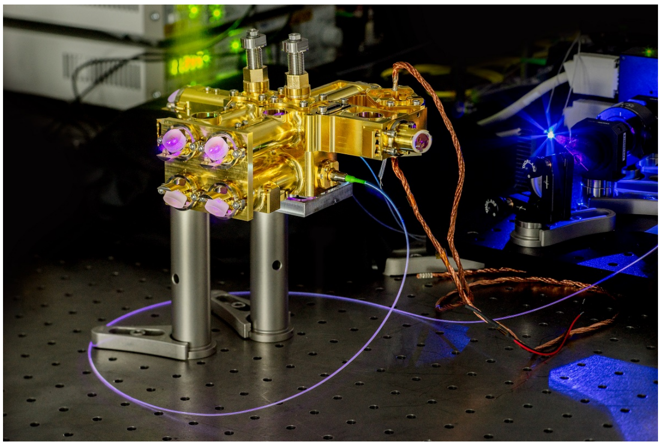


Figure 2: **Opto-mechanical implementation of the space-suitable entangled photon engineering model.** The main constraints of the opto-mechanical implementation are to provide a mechanically and temperature stable mounting structure for the optics of the EPS while maintaining finetuneability of the required degrees of freedom for alignment and fixation of individual components.

Physical-depth architectural requirements for generating universal photonic cluster states

Sam Morley-Short[1] [2] [*] Sara Bartolucci[3] Mercedes Gimeno-Segovia[1] [3] [4] Pete Shadbolt [3]
Hugo Cable[1] Terry Rudolph[3]

[1] *Quantum Engineering Technology Labs, H. H. Wills Physics Laboratory and Department of Electrical and Electronic Engineering, University of Bristol, BS8 1FD, UK*

[2] *Quantum Engineering Centre for Doctoral Training, H. H. Wills Physics Laboratory and Department of Electrical and Electronic Engineering, University of Bristol, Tyndall Avenue, BS8 1FD, UK*

[3] *Department of Physics, Imperial College London, London SW7 2AZ, UK*

[4] *Institute for Quantum Science and Technology, University of Calgary, Alberta T2N 1N4, Canada*

Abstract. Leading proposals for linear-optical quantum computing (LOQC) use cluster states as universal resources for measurement-based quantum computation. Results from percolation theory have shown that universal cluster states can be generated using schemes which exceed the critical percolation threshold, but these results consider states with unbounded size. Here we consider how percolation can be maintained using a fixed physical-depth architecture, assuming the state is continuously generated and measured such that only a finite portion is visible at any time. We show that universal LOQC can be implemented using a constant-size device with modest depth without the need for high-complexity algorithms.

Keywords: Linear Optical Quantum Computing, Quantum Computation Architectures, Percolation

1 Context

Over the two last decades architectures for linear optical quantum computation (LOQC) have matured from (technically) efficient, but ultimately unrealistic proof-of-principle designs [1] to scalable and increasingly feasible modern proposals [2]. Historically, the primary challenge for LOQC architectures was the inherent non-determinism of entangling operations between non-interacting photonic qubits (using only linear optics). By utilising the paradigm of measurement-based quantum computation (MBQC) [3, 4], modern architectures avoid the need to perform probabilistic entangling gates arbitrarily during the computation, allowing all entanglement to be generated prior to the desired quantum computation. As such, current LOQC architectures demand the generation of large-scale entangled cluster states that provide a universal resource for quantum computation.

In current LOQC architectures, large cluster states are created from small three-qubit GHZ states via probabilistic entangling "fusion" gates (that can be "boosted" to operate with an arbitrarily high success rate [5, 6]). Within such a model, Gimeno-Segovia, et. al. showed that when the fusion success rate exceeds some critical threshold, large-scale entanglement is produced in a percolated manner [7]. Single-qubit channels and states for quantum error correction are then produced from percolated lattice cluster states by "renormalization", whereby blocks of percolated physical qubits are abstracted to individual logical qubits with the idealised lattice structure [8]. One key advantage of such a model is that after the initial GHZ resource state generation [9], photons do not pass through active, high-loss components, producing a so-called "ballistic" architecture. This approach is contrasted to other modern LOQC proposals that generate entanglement with a "repeat-until-success" architecture, requiring large, snowflake-like resource states and many layers of active switching [10]. Once a large percolated cluster state lattice is constructed, identifying a percolation path spanning the cluster allows single-qubit channels and renormalization to be performed.

While this model is valid in an abstract computational space, any feasible architecture must also consider realistic constraints of a physical device. Due to the geometric and material constraints of any foreseeable optical platform, it is unrealistic to suggest that a LOQC device must create and store the full cluster state lattice needed for a quantum computation at any one time. Instead, we consider a "windowed" architecture whereby the device is continually being created and measured, storing only a finite slice of the full resource state at any one time. As such, current architectural methodssuch as identifying MBQC paths for single-qubit channelsmust be extended to a windowed architecture.

2 Presented work

Our submitted work [11] considers this architectural challenge. Specifically, we ask: "what is the smallest computational window required to produce a single-qubit channel from a percolated cluster state lattice, and what are the associated architectural trade-offs." Within the described LOQC architecture, this question maps to the challenge of finding paths through a maze given only a fixed "lookahead" and with no backtracking. By considering a simple algorithm for "limited-lookahead pathfinding" (LLP), we show that finite window lengths are sufficient to produce long-range, low-loss single-qubit channels. Moreover, we show that LLP can operate with surprisingly small window lengths even for lattices produced with entangling gate success rates

[*] sam.morley-short@bristol.ac.uk

only marginally above the percolation threshold. Furthermore, we demonstrate that at such window lengths, near-perfect LLP is achieved without any need for more sophisticated pathfinding "strategies". This is especially pertinent to LOQC as photonic qubits must be stored in loss-inducing delay-lines during all classical co-processes, and therefore any reductions in co-processor requirements may lead to significant reductions in qubit loss rates.

Our work also identifies heuristics techniques for simulating the performance LLP. By showing that easy-to-calculate percolation statistics can be used to approximate computationally expensive LLP, we provide a method for fast simulation of novel architectures; as architectural models become increasingly complex, such simulation heuristics will be crucial to rapid development and innovation within the field.

Finally, we summarise by providing a number of key implications this work has for LOQC and identifying directions for further study. Firstly, this work indicates that a device with fixed-depth can still allow for successful LOQC, providing a significant insight into the ultimate form of a physical LOQC device. Secondly, this work elucidates key resource trade-offs inherent to modern LOQC architectures. Our work shows that cluster states lattices must be produced with connectivity exceeding the percolation threshold, and that large reductions in resource costs occur when this rate is further increased. Lastly, the work provides a positive outlook for the maturation of LOQC from purely theoretical models to experimentally viable architectures.

References

[1] E. Knill, R. Laflamme & G. J. Milburn. A scheme for efficient quantum computation with linear optics. In *Nature*, **409**, pages 46–52, 2001.

[2] T. Rudolph Why I am optimistic about the silicon-photonic route to quantum computing. In *APL Photonics*, **2**, 030901, 2017.

[3] R. Raussendorf & H. J. Briegel. A One-Way Quantum Computer. In *Physical Review Letters*, **86**, pages 5188–5191, 2001.

[4] R. Raussendorf, D. E. Browne & H. J. Briegel. Measurement-based quantum computation on cluster states. In *Physical Review A*, **68**, 022312, 2003.

[5] W. Grice. Arbitrarily complete Bell-state measurement using only linear optical elements. In *Physical Review A*, **84**, pages 1–6, 2011.

[6] F. Ewert & P. van Loock. 3/4-Efficient Bell Measurement with Passive Linear Optics and Unentangled Ancillae. In *Physical Review Letters*, **113**, pages 1–5, 2014.

[7] M. Gimeno-Segovia, P. Shadbolt, D. E. Browne & T. Rudolph. From Three-Photon Greenberger-Horne-Zeilinger States to Ballistic Universal Quantum Computation. In *Physical Review Letters*, **115**, 020502, 2015.

[8] K. Kieling, T. Rudolph & J. Eisert. Percolation, renormalization, and quantum computing with non-deterministic gates. In *Physical Review Letters*, **99**, pages 2–5, 2007.

[9] M. Gimeno-Segovia, et al. Relative multiplexing for minimizing switching in linear-optical quantum computing. In *New Journal of Physics*, **19**, 063013, 2017.

[10] Y. Li, P. C. Humphreys, G. J. Mendoza & S. C. Benjamin. Resource costs for fault-tolerant linear optical quantum computing. In *Physical Review X*, **5**, pages 1–15, 2015.

[11] S. Morley-Short, et al. Physical-depth architectural requirements for generating universal photonic cluster states. *arXiv preprint*, 1706.07325, 2017.

Quantum simulation of the quantum Rabi model in a single trapped ion

Dingshun Lv[1] * Shuoming An[1] Zhenyu Liu[1] JN Zhang[1] Julen S. Pedernales[2] Lucas Lamata[2] Enrique Solano[2 3] Kihwan Kim[1]

[1] *Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, People's Republic of China*

[2]*Department of Physical Chemistry, University of the Basque Country UPV/EHU, Apartado 644, 48080 Bilbao, Spain*

[3] *IKERBASQUE, Basque Foundation for Science, Alameda Urquijo 36, 48011 Bilbao, Spain*

Abstract. The quantum Rabi model describes the most fundamental light-matter interaction of the dipolar coupling between a two-level system and a bosonic field mode. An analytical solution of the quantum Rabi model covering all coupling regimes such as the weak, the ultrastrong, and the deep-strong coupling regimes has only recently been proposed. Moreover, several physical systems have been pursued to implement the perturbative regime of ultrastrong coupling. However, it is still challenging to reach the dynamics of the nonperturbative ultrastrong coupling regime and the deep-strong coupling regime, which would show intriguing physical phenomena beyond intuitive features of the quantum Rabi model. Here, we implement the quantum simulation of the paradigmatic quantum Rabi model in a trapped-ion system, reproducing key features of all parameter regimes from the weak to the deep-strong coupling regimes.

Keywords: Quantum Rabi model, Jaynes-Cummings coupling, Ultrastrong coupling, Deep-Strong coupling, Adiabatic ground state preparation, Ion trap

The quantum Rabi model (QRM) describes the most fundamental light-matter interaction involving quantized light and quantized matter [1, 2], which is associated to the dipolar coupling between a two-level system and a bosonic field mode. The Hamiltonian of the QRM is written as

$$\hat{H}_{\mathrm{QRM}}(\phi) = \frac{\omega_0}{2}\hat{\sigma}_z + \omega_m \hat{a}^+\hat{a} + ig(\hat{\sigma}_+ - \hat{\sigma}_-)(\hat{a} + \hat{a}^+), \quad (1)$$

where ω_0, ω_m, and g are frequencies of the qubit, the mode, and the coupling strength. Although it plays a central role in the dynamics of a collection of quantum optics and condensed matter systems [3], such as cavity quantum electrodynamics (CQED), quantum dots, trapped ions, or circuit QED (cQED), an analytical solution of the QRM in all coupling regimes has only recently been proposed [4]. Typically, the coupling strength is much weaker than the mode frequency, which allows one to perform the rotating-wave approximation that leads to the Jaynes-Cummings (JC) model. When the interaction strength grows to a meaningful fraction or larger than the mode frequency, the ultrastrong coupling (USC) or the deep-strong coupling (DSC) regime is reached, respectively, where the rotating-wave approximation is no longer valid. Recently, several systems have been able to experimentally reach the perturbative USC regime of the QRM. However, it is still challenging to reach the nonperturbative USC regime, or DSC regime, which would show intriguing physical phenomena beyond intuitive features of the QRM.

Here, following Ref. [5], we implement the QRM and we experimentally simulate the dynamics of the model in all parameter regimes from the weak and USC to the DSC regime in a single trapped ion system, which is considered as one of the prominent platforms for building quantum simulators. In the experimental quantum simulation, a single atomic ion with the two internal levels of a qubit is confined in a radio-frequency Paul trap and its motional quantum state is cooled down to the ground state by standard sideband cooling. The general coupling of the QRM between the internal level and one of the radial motional modes is realized by a laser field with two frequencies in the resolved-sideband limit shown in Ref. [5]. By taking a suitable interaction picture associated with two inhomogeneously detuned laser beams, we address all parameter regimes of the QRM [5].

In the experiment, we have simulated various features of the QRM from the dynamics, spectrum, and ground states in all parameter regimes. Firstly, we simulate the dynamics of the QRM from the weak-coupling regime via a USC regime to the DSC regime. Figure 1 shows the spin dynamics under the QRM with respect to the ratio g/ω_m equal to 0.04, 0.6, and 1.2, respectively, where we can clearly observe collapses and partial revivals during the system evolution. In particular, in the DSC regime, we observe the phonon bounce forth and back within the same parity chains. Secondly, we measure the spectrum of the QRM up to the USC regime. Finally, we adiabatically prepare the ground state of the QRM in the DSC regime and measure its corresponding phonon distribution for different spin states as shown in Fig. 2, which reveals a complex entangled structure between the spin and the bosonic field. We also measure the fidelity of the state by adiabatically bringing it back to the original initial state and detecting the population in the original ground state.

Summarizing, we fully simulate the QRM in a trapped ion system. The present ideas are straightforwardly generalizable to many ions, opening the possibility of going from the more natural Tavis-Cummings model to the Dicke model.

*ywlds@163.com

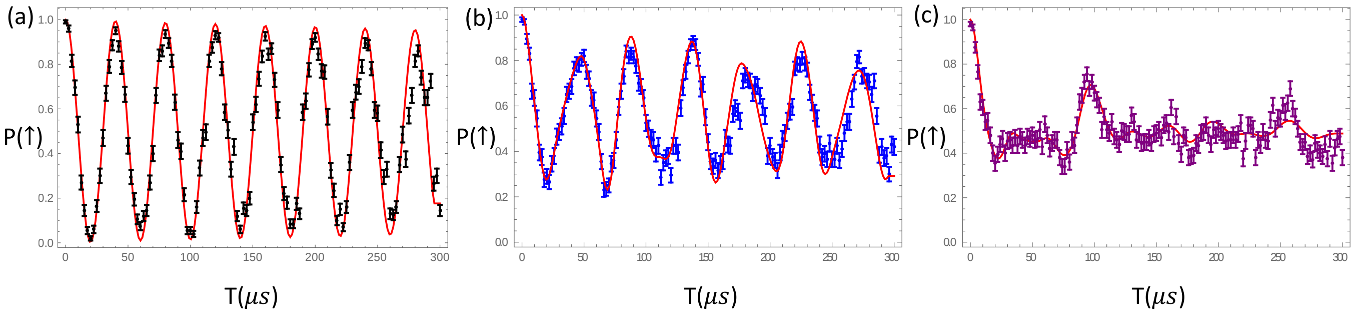


Figure 1: **Spin dynamics under the QRM in different coupling regimes.** (a) For the ratio $g/\omega_m = 0.04$, the system behaves as the JC model in the weak coupling regime. (b) For the ratio $g/\omega_m = 0.6$, namely, the ultrastrong coupling regime, we can clearly observe collapse and revival signatures. (c) For the ratio $g/\omega_m = 1.2$, the collapse is fast and the revival is only partial. For all the panels above, the theoretical curve is the solid line, while the points with error bars are the experiment results.

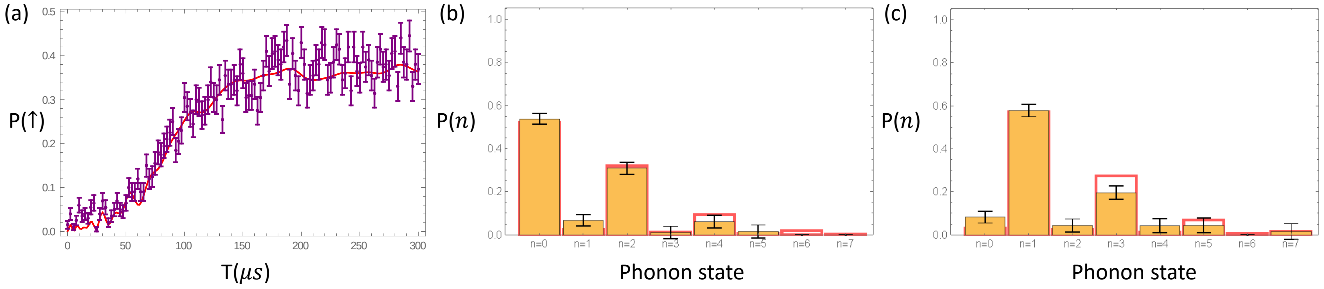


Figure 2: **Adiabatical ground-state preparation to the deep-strong coupling regime** $g/\omega_m = 1.2$**.** Panel (a) shows the spin dynamical evolution during the adiabatic preparation of the ground state for the DSC regime. Panels (b) and (c) depict the phonon distributions of the prepared ground state with respect to the spin down and spin up part, respectively. All the experimental results are plotted with error bars, while the simulation results are solid curve (panel (a)) or empty bars (panels (b) and (c)).

References

[1] Rabi, I. On the process of space quantization. Phys. Rev. 49, 324 (1936).

[2] Braak, D., Chen, Q.-H., Batchelor, M. T., and Solano, E. Semi-classical and quantum Rabi models: in celebration of 80 years. J. Phys. A: Math. Theor. 49, 300301 (2016).

[3] Arakawa, Y. et al. Focus on cavity and circuit quantum electrodynamics in solids. New Journal of Physics 17, 010201 (2015).

[4] Braak, D. Integrability of the Rabi model. Physical Review Letters 107, 100401 (2011).

[5] Pedernales, J. et al. Quantum Rabi model with trapped ions. Scientific Reports 5, 15472 (2015).

Holonomic surface codes for fault-tolerant quantum computation

Jian-Qiang You

CSRC, Beijing

Abstract: Surface codes can protect quantum information stored in qubits from local errors as long as the per-operation error rate is below a certain threshold. Imperfect control is a main source of errors that make the threshold a challenging task. By harnessing quantum holonomy, here we propose a method to suppress the errors caused by imperfect control in surface codes. In our scheme, the holonomic gates are built via auxiliary qubits rather than the auxiliary levels in multi-level systems used in conventional holonomic quantum computation. The key advantage of our scheme is that the auxiliary qubits are in their ground state before and after each gate operation, so they are not involved in the operation cycles of surface codes. This provides a new and advantageous way to implement surface codes for fault-tolerant quantum computation.

Group theory and non-local games

William Slofstra

Waterloo, Canada

Abstract: How much entanglement is required to play a non-local game optimally or near-optimally? This question has proven very difficult to answer in general. Recently we have found non-local games which cannot be played optimally with any finite amount of entanglement. In this talk, I will explain how this result arises out of a connection between linear system non-local games and group theory. This connection opens up a number of avenues for further research in entanglement requirements, self-testing, and complexity.

Classically testing the exponential nature of Hilbert space

Henry Yuen

Berkeley, USA

Abstract: As we create more sophisticated quantum systems (including, one day, quantum computers) it becomes imperative to characterize just how much "quantumness" is present in them. However, the exponentiality of Hilbert space, the very feature of Nature that we are trying to exploit in these systems, also poses a significant barrier to verifying quantum behavior. Bell tests offer a powerful solution to this challenge. By performing simple statistical tests on measurement outcomes of spatially separated systems, we can certify not only the presence of quantum behavior, in certain cases we can even *characterize* the quantum state of the systems, as well as the measurement operators. In recent years, Bell tests (also known as *non-local games*) have found widespread usage in quantum information processing, from randomness testing protocols to delegated quantum computation.In this talk, I will survey the recent progress in using Bell tests to certify *high dimensional* entanglement --- a setting where the dimensionality is an asymptotically growing parameter. I will also describe some new results on testing high dimensional entanglement in the presence of noise (joint work with Rotem Arnon-Friedman). These tests establish an important bridge connecting the classical world we live in to the exponentially vast Hilbert space of quantum states.

Fidelity of quantum strategies with applications to cryptography

Gus Gutoski[1] [*] Ansis Rosmanis[2] [3] [†] Jamie Sikora[2] [‡]

[1] *Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada*
[2] *Centre for Quantum Technologies, National University of Singapore, Singapore*
[3] *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore*

Abstract. We generalize the fidelity function for multi-round quantum strategies, which we call the *strategy fidelity*. We provide many interesting properties of the strategy fidelity, including a Fuchs-van de Graaf relationship with the strategy norm. And we illustrate an operational interpretation of the strategy fidelity in the spirit of Uhlmann's Theorem and discuss its application to the security analysis of quantum protocols for interactive cryptographic tasks such as bit-commitment and oblivious string transfer. Our analysis is very general in the sense that the actions of the protocol need not be fully specified, which is in stark contrast to most other security proofs.

Keywords: Quantum strategies, cryptography, fidelity, semidefinite programming

1 Setting and Definitions

In this paper we consider multiple-round interactions between two parties involving the exchange of quantum information. There is a natural asymmetry between the parties as only one of the parties can send the first message or receive the final message. Since we are not concerned about optimizing the number of messages exchanged, without loss of generality both of these tasks are done by the same party, which, for convenience, we call *Bob*. Let us call the other party *Alice*. The interaction between Alice and Bob decomposes naturally into a finite number r of *rounds* (see Figure 1). Such interactions are conveniently described by the formalism of quantum strategies introduced in Ref. [3].

Definition 1 (Pure strategy and pure co-strategy) *Let $r \geq 1$ and let $\mathcal{X}_1, \ldots, \mathcal{X}_r, \mathcal{Y}_1, \ldots, \mathcal{Y}_r, \mathcal{Z}_r, \mathcal{W}_r$ be complex Euclidean spaces and, for notational convenience, let $\mathcal{X}_{r+1} := \mathbb{C}$ and $\mathcal{Z}_0 := \mathbb{C}$. An* r-round pure strategy *$\tilde{A}$ having* input spaces *$\mathcal{X}_1, \ldots, \mathcal{X}_r$,* output spaces *$\mathcal{Y}_1, \ldots, \mathcal{Y}_r$, and* final memory space *$\mathcal{Z}_r$, consists of:*

1. *complex Euclidean spaces $\mathcal{Z}_1, \ldots, \mathcal{Z}_{r-1}$, called* intermediate memory spaces*, and*
2. *an r-tuple of linear isometries $(A_1, \ldots, A_r)$ of the form $A_i : \mathcal{X}_i \otimes \mathcal{Z}_{i-1} \to \mathcal{Y}_i \otimes \mathcal{Z}_i$.*

An r-round pure co-strategy *having input spaces $\mathcal{Y}_1, \ldots, \mathcal{Y}_r$, output spaces $\mathcal{X}_1, \ldots, \mathcal{X}_r$, and final memory space $\mathcal{W}_r$ is defined similarly (see Figure 1 for an illustration).*

A pure strategy and a pure co-strategy are said to be compatible *when the input spaces of one are the output spaces of the other, and vice versa. The* final state *of the interaction between $\tilde{A}$ and $\tilde{B}$ is denoted by*

$$|\psi(\tilde{A}, \tilde{B})\rangle := (I_{\mathcal{Z}_r} \otimes B_r)(A_r \otimes I_{\mathcal{W}_{r-1}}) \cdots \\ \cdots (I_{\mathcal{Z}_1} \otimes B_1)(A_1 \otimes I_{\mathcal{W}_0})|\beta\rangle \in \mathcal{Z}_r \otimes \mathcal{W}_r.$$

For compatible pure strategy $\tilde{A}$ and pure co-strategy $\tilde{B}$, let

$$\rho_A(\tilde{B}) := \operatorname{Tr}_{\mathcal{Z}_r}\left(|\psi(\tilde{A}, \tilde{B})\rangle\langle\psi(\tilde{A}, \tilde{B})|\right) \tag{1}$$

denote the reduced state of the final memory space $\mathcal{W}_r$ of $\tilde{B}$ after the interaction between $\tilde{A}$ and $\tilde{B}$.

Recall that the fidelity $\mathrm{F}(P, Q)$ between two positive semidefinite operators P and Q is defined as

$$\mathrm{F}(P, Q) := \left\| \sqrt{P}\sqrt{Q} \right\|_{\mathrm{Tr}}.$$

When applied to density operators ρ, ξ, the fidelity function $\mathrm{F}(\rho, \xi)$ is a useful distance measure for quantum states. We would like to construct a generalization of the fidelity function that can serve as a useful distance measure for quantum strategies.

Definition 2 (Strategy fidelity) *For any r-round strategies S and T having the same input and output spaces, the* strategy fidelity *is defined as*

$$\mathrm{F}_\mathrm{r}(S, T) := \min_{B} \mathrm{F}(\rho_S(\tilde{B}), \rho_T(\tilde{B}))$$

where the minimization is over all compatible co-strategies B and the states $\rho_S(\tilde{B}), \rho_T(\tilde{B})$ are as defined in (1)*.*

2 Properties of the Strategy Fidelity

We now list several properties of the strategy fidelity which we prove in the paper.

- (Fuchs-van de Graaf inequalities for strategies) For any r-round strategies S and T, it holds that
$$1 - \frac{1}{2}\|S - T\|_{\diamond\mathrm{r}} \leq \mathrm{F}_\mathrm{r}(S, T) \leq \sqrt{1 - \frac{1}{4}\|S - T\|_{\diamond\mathrm{r}}^2}.$$
- (Symmetry) For any r-round strategies S and T, it holds that $\mathrm{F}_\mathrm{r}(S, T) = \mathrm{F}_\mathrm{r}(T, S)$.

[*] gus.gutoski@isara.com
[†] ansis@ntu.edu.sg
[‡] jamiesikora@gmail.com

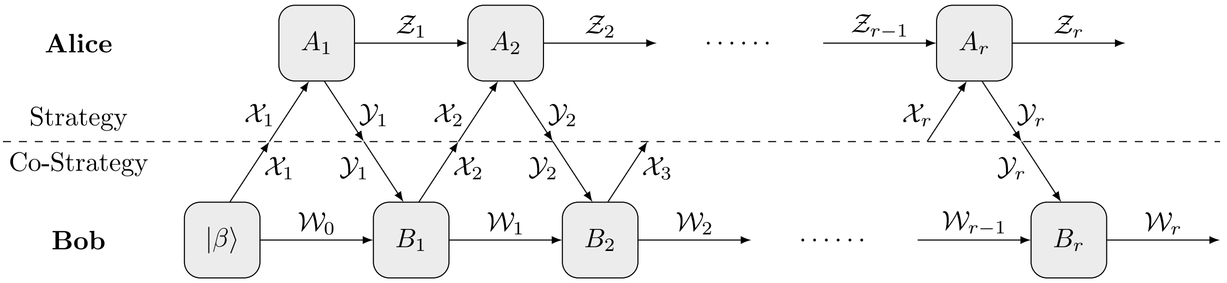


Figure 1: An r-round interaction between a pure strategy of Alice (the linear isometries above the dashed line) and a pure co-strategy of Bob (the linear isometries below the dashed line). Arrows crossing the dashed line represent messages exchanged between the parties, while horizontal arrows represent private memory.

- (Joint concavity) For any r-round strategies $S^1, \ldots, S^n$ and $T^1, \ldots, T^n$, and nonnegative scalars $\lambda_1, \ldots, \lambda_n$ satisfying $\sum_{i=1}^n \lambda_i = 1$, we have
$$\mathrm{F_r}\left(\sum_{i=1}^n \lambda_i S^i, \sum_{i=1}^n \lambda_i T^i\right) \geq \sum_{i=1}^n \lambda_i \, \mathrm{F_r}\left(S^i, T^i\right).$$
- (Bounds on the strategy fidelity) For any r-round strategies S and T, we have $0 \leq \mathrm{F_r}(S,T) \leq 1$. Moreover, $\mathrm{F_r}(S,T) = 1$ if and only if $S = T$ and $\mathrm{F_r}(S,T) = 0$ if and only if S and T are perfectly distinguishable.
- (Monotonicity) For all physically realizable maps Υ from r-round strategies to r'-round strategies, it holds that
$$\mathrm{F_{r'}}(\Upsilon(S), \Upsilon(T)) \geq \mathrm{F_r}(S,T).$$
- (Strategy generalization of Uhlmann's Theorem) Let S, T be r-round strategies and let $\tilde{S}, \tilde{T}$ be any purifications of S, T. Let $\psi(\tilde{S}, \tilde{B})$, $\psi(\tilde{T}, \tilde{B})$ be the density operators corresponding, respectively, to states $|\psi(\tilde{S}, \tilde{B})\rangle$, $|\psi(\tilde{T}, \tilde{B})\rangle$ in Definition 1. We have
$$\mathrm{F_r}(S,T)^2 = \max_{\Xi} \min_{B} \left\langle \ (\tilde{S}, \tilde{B}), \left(\Xi \otimes I_{\mathbf{L}(\mathcal{W}_r)}\right) \left(\ (\tilde{T}, \tilde{B})\right) \right\rangle$$
where the minimum is over all r-round pure co-strategies $\tilde{B}$ and the maximum is over all quantum channels Ξ acting on $\mathcal{Z}_r$ alone.
- (Semidefinite programming formulation of the strategy fidelity) This is detailed in the paper.

3 Applications to Cryptography

In the paper we discuss how the strategy version of the Fuchs-van de Graaf inequalities is crucial to our cryptographic applications. In particular, we show the impossibility of ideal quantum protocols for interactive *bit-commitment* and *oblivious string transfer*.

We define bit-commitment below, and include the definition of oblivious transfer in the paper.

Definition 3 *In bit-commitment, we require Alice and Bob to interact over two communication stages:*

- *Commit Phase: Alice chooses a uniformly random bit a and interacts with Bob using an r-round pure strategy $\tilde{A}^a$.*
- *Reveal Phase: Alice sends a to Bob and continues her interaction with him (so that Bob can test if she has cheated).*
- *Cheat Detection: Bob, knowing which pure strategy $\tilde{B}$ he has used, measures to check if the final state is consistent with Alice's pure strategy $\tilde{A}^a$. He aborts the protocol if this measurement detects the final state is not consistent with Alice's pure strategy $\tilde{A}^a$. If Alice is honest, he never aborts.*

Protocols are designed with the intention to achieve the following two important properties of interest:

- Binding: Alice cannot change her mind after the Commit Phase and reveal the other value of a (without being detected by Bob).
- Hiding: Bob cannot learn Alice's bit a before she reveals it during the Reveal Phase.

Finding a protocol with perfect binding and hiding properties is known to be impossible [6, 4, 5]. However, these security proofs rely on an assumption that we do not make, that Bob's actions are specified beforehand (see the paper for details).

We define the cheating probabilities of Alice and Bob as follows:

Bob:	The maximum probability with which a dishonest Bob can *learn* an honest Alice's committed bit $a \in \{0,1\}$ after the Commit Phase.
Alice:	The maximum probability Alice can change her commitment from 0 to 1 (or from 1 to 0) before the Reveal Phase.

Theorem 4 *In any interactive quantum protocol for bit-commitment, we have that Alice or Bob can cheat with probability at least $\frac{9-\sqrt{17}}{8} \approx 61\%$.*

Note that this is a similar bound to the one obtained in [2] for the interactive setting and the same as in [1] in the channel setting.

We also derive a lower bound on oblivious transfer, given below.

Theorem 5 *In any interactive quantum protocol for 1-out-of-2 oblivious string transfer, we have that Alice or Bob can cheat with probability at least* $\frac{9-\sqrt{17}}{8} \approx 61\%$.

References

[1] V. P. Belavkin, G. M. D'Ariano, and M. Raginsky. Operational distance and fidelity for quantum channels. *Journal of Mathematical Physics*, 46(6):062106, 2005. arXiv:quant-ph/0408159.

[2] G. Chiribella, G. M. D'Ariano, P. Perinotti, D. Schlingemann, and R. F. Werner. A short impossibility proof of quantum bit commitment. *Physics Letters A*, 377(15):1076–1087, 2013. arXiv:0905.3801v1 [quant-ph].

[3] G. Gutoski and J. Watrous. Toward a general theory of quantum games. In *Proceedings of the 39th ACM Symposium on Theory of Computing (STOC 2007)*, pages 565–574, 2007. arXiv:quant-ph/0611234.

[4] H.-K. Lo and H. F. Chau. Is quantum bit commitment really possible? *Physical Review Letters*, 78(17):3410–3413, 1997.

[5] H.-K. Lo and H. F. Chau. Why quantum bit commitment and ideal quantum coin tossing are impossible. *Physica D: Nonlinear Phenomena*, 120(1–2):177–187, Sept. 1998. Proceedings of the Fourth Workshop on Physics and Consumption.

[6] D. Mayers. Unconditionally secure quantum bit commitment is impossible. *Physical Review Letters*, 78(17):3414–3417, 1997.

Generalized entanglement entropies of quantum designs

[Long version arXiv:1703.08104]

Zi-Wen Liu,[1,*] Seth Lloyd,[2] Elton Yechao Zhu,[1] and Huangjun Zhu[3]

[1] *Center for Theoretical Physics and Department of Physics,*
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[2] *Department of Mechanical Engineering, Massachusetts*
Institute of Technology, Cambridge, Massachusetts 02139, USA
[3] *Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany*

Ensembles of quantum states or unitaries that reproduce the first α moments of completely random states or unitaries (drawn from the Haar measure) are called α designs. Entropic functions of the α-th power of a density operator are called α entropies (e.g. Rényi and Tsallis). We reveal strong connections between designs and generalized (in particular Rényi) entropies of the same order, by showing that the Rényi-α entanglement entropies averaged over α designs are generically almost maximal. Moreover, we find that the min entanglement entropies become maximal for designs of an order logarithmic in the dimension of the system, which implies that they are indistinguishable from Haar-random by the entanglement spectrum.

The entanglement properties of random quantum states or dynamics play important roles in various disciplines of physics, not only restricted to quantum information. For example, the notion of 'scrambling', which originates from the study of black holes and quantum gravity [1–3], describes the phenomenon that initially localized quantum information spreads over the entire system via global entanglement, so that the state of the system is effectively randomized and the information is lost from the perspective of any local observer. The relation between the degree of entanglement and randomness also plays key roles in many other important fields, such as quantum chaos [4–6], quantum thermalization and many-body localization [7], and quantum data hiding [8, 9].

It has long been noticed that a random state is typically highly entangled [10, 11]. This observation is formalized by the Page's theorem [12–15], which states that the average von Neumann entanglement entropy of a completely random state (drawn from the Haar measure) is very close to maximum. Similar observations for the entanglement in random unitary channels are recently made in Ref. [6]. However, such results are not tight from a complexity point of view. Designs are pseudorandom distributions of quantum states or unitaries that mimic the Haar measure up to certain moments. The complexity of Haar randomness is exponential, that is, the number of local gates and random bits needed to generate a Haar random state grows exponentially in the number of degrees of freedom [16]. Nevertheless, a 2-design, which can be efficiently implemented [17–20], is sufficient to attain the Page-like property. Indeed, the conventional von Neumann entropy is insensitive to a lot of detailed information in the spectrum. The entanglement entropies given by generalized entropies that depend on higher powers of the reduced density operator are needed to distinguish the entanglement spectra of ensembles of different complexities.

This work mainly studies generalized (most importantly Rényi) entanglement entropies averaged over state and unitary designs, so as to obtain the strongest entanglement properties of pseudorandom ensembles (designs).

Key definitions. Designs can be defined in several equivalent ways. We directly use the following one by polynomials: Let $\mathrm{Hom}_{(t,t)}(\mathbb{C}^d)$ be the space of polynomials homogeneous of degree t both in the coordinates of vectors in $\mathbb{C}^d$ and in their complex conjugates. An ensemble ν of pure state vectors in dimension d is a (complex projective) *t-design* if $\mathbb{E}_\nu\, p(\psi) = \int p(\psi)\mathrm{d}\psi$ for all $p \in \mathrm{Hom}_{(t,t)}(\mathbb{C}^d)$, where the integral is taken with respect to the (normalized) uniform measure on the complex unit sphere in $\mathbb{C}^d$. Unitary designs are defined similarly, where $p \in \mathrm{Hom}_{(t,t)}(\mathrm{U}(d))$

and the integral is taken over the Haar measure on $\mathrm{U}(d)$.

The defining element of α-entropies of state ρ is the term $\mathrm{tr}\{\rho^\alpha\}$. We mostly focus on Rényi entropies. The Rényi-α entropy of a state ρ is given by $S_R^{(\alpha)}(\rho) = \frac{1}{1-\alpha}\log \mathrm{tr}\{\rho^\alpha\}$. The $\alpha \to \infty$ limit $S_{\min}(\rho) = -\log\|\rho\| = -\log\lambda_{\max}(\rho)$ is known as the min entropy.

Main results. We first find that Rényi-α entanglement entropy of a state or unitary sampled from an α-design is generically maximal. The results represent formal connections between the order of entanglement entropies and the order of designs. We mostly employ tools from random matrix theory, representation theory and Weingarten calculus to calculate the Haar integrals of generalized entanglement entropies, which provide lower bounds to the design-averaged values. Note that the entanglement properties of unitaries are studied via the Choi states. See [6] for motivations.

In the limit of large dimension, our results imply that the Rényi-α entanglement entropy between subsystems of equal size (for simplicity) averaged over α-designs is nearly maximal:

Theorem 1. *Consider bipartite systems on Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_A$ and $\mathcal{H}_B$ have dimensions d_A and d_B respectively. Let ν_α be a projective α-design. Consider equal partitions $d_A = d_B$. As $d_A \to \infty$,*

$$\mathbb{E}_{\nu_\alpha} S_R^{(\alpha)}(\rho_A) \geq \log d_A - \frac{\log \mathrm{Cat}_\alpha}{\alpha - 1} + O(d_A^{-2}) \geq \log d_A - O(1), \tag{1}$$

where Cat_α is the α-th Catalan number, satisfying $\frac{\log \mathrm{Cat}_\alpha}{\alpha-1} \leq 2$ for all $\alpha \geq 2$. Also consider a unitary $U = \sum_{ij} U_{ij}|i\rangle\langle j|$ on a d-dimensional Hilbert space. Let μ_α be a unitary α-design. The dual Choi state is given by $|U\rangle = \frac{1}{\sqrt{d}}\sum_{ji}|i\rangle_{in}|j\rangle_{out}$. Consider equal partitions of the input and output registers. As $d \to \infty$,

$$\mathbb{E}_{\mu_\alpha} S_R^{(\alpha)}(\rho_{AC}) \geq \log d - \frac{\log \mathrm{Cat}_\alpha}{\alpha - 1} + O(d^{-1}) \geq \log d - O(1). \tag{2}$$

That is, the average Rényi-α entanglement entropy of α-designs along all valid cuts is only smaller than the maximum value $\log d$ by at most a constant. It implies that a state/unitary drawn from an α-design is very likely to have almost maximal Rényi-α entanglement entropy.

We also provide explicit bounds in finite dimension:

Theorem 2. *Let ν_α be a projective α-design. Let $q := \alpha^3/(32d_B^2) < 1, h(q) := 1 + 2q/[3(1-q)]$. For all $d_A, d_B, 0 \leq \alpha \leq \infty$,*

$$\mathbb{E}_{\nu_\alpha} S_R^{(\alpha)}(\rho_A) \geq \log d_A - \frac{2\alpha - \frac{3}{2}\log\alpha + \log h(q) - \frac{1}{2}\log\pi}{\alpha - 1} \geq \log d_A - 2. \tag{3}$$

When $d_A < d_B$, the result can be improved as follows:

$$\mathbb{E}_{\nu_\alpha} S_R^{(\alpha)}(\rho_A) \geq \log d_A - 2\log\left(1 + \sqrt{\frac{d_A}{d_B}}\right) - \log c \geq \log d_A - 2\sqrt{\frac{d_A}{d_B}} - \log c, \tag{4}$$

where $c = 1$ if $\mathcal{H}$ is real and $c = 2$ if $\mathcal{H}$ is complex.

Let μ_α be a unitary α-design. Suppose $d > \sqrt{6}k^{7/4}, d_A \leq d_B$. Then

$$\mathbb{E}_{\mu_\alpha} S_R^{(\alpha)}(\rho_{AC}) \geq \log d - \frac{\log \mathrm{Cat}_\alpha}{\alpha - 1} - \frac{\log\left[\frac{a_\alpha h(q)}{8}\left(7 + \cosh\frac{2\alpha(\alpha-1)}{d}\right)\right]}{\alpha - 1}, \tag{5}$$

where $a_\alpha := \frac{1}{1 - \frac{6\alpha^{7/2}}{d^2}}$.

Now we focus on the extreme case of the Rényi family—the min entropy. It is the strongest entropy in the sense that, by definition, it is sensitive to any nonuniformity in the spectrum. The maximality of min entropy indicates that the spectrum is uniform everywhere (completely random). We find that, for states/unitaries in dimension d, designs of order $O(\log d)$ already exhibit nearly maximal average min entanglement entropy:

Theorem 3. *Let ν_α be a projective α-design, where $\alpha = \lceil (\log d_A)/a \rceil \leq (16 d_B^2)^{1/3}$ with $0 < a \leq 1$. Then*

$$\mathbb{E}_{\nu_\alpha} S_{\min}(\rho_A) \geq \log d_A - 2 - a. \tag{6}$$

In particular, $\mathbb{E}_{\nu_\alpha} S_{\min}(\rho_A) \geq \log d_A - 3$ if $\alpha = \lceil \log d_A \rceil$.

Let μ_α be a unitary α-design, where $1 \leq \alpha = \lceil \log d/a \rceil \leq \sqrt{d}/2$ and $a > 0$; then

$$\mathbb{E}_{\nu_\alpha} S_{\min}(\rho_{AC}) \geq \log d - 2 - a. \tag{7}$$

In particular, $\mathbb{E}_{\nu_\alpha} S_{\min}(\rho_{AC}) \geq \log d - 3$ if $\alpha \geq \lceil \log d \rceil$.

This result actually implies that, in terms of entanglement, designs of order only up to $O(\log d)$ can behave "pseudorandomly". Designs of higher orders have entanglement properties that are indistinguishable from Haar.

We also provide a result that separates different orders of Rényi entanglement entropies. We show that

Theorem 4. *There exists 2-designs such that the difference between average Rényi-α entanglement entropies from the maximum is unbounded for $\alpha > 2$.*

That is, the Rényi-2 entanglement entropy is nearly maximal by previous results, but the $\alpha > 2$ Rényi entanglement entropies are far from maximal, so we can distinguish some 2-designs from higher order designs by Rényi-2. We hope to extend this result to higher orders and unitary channels in the future.

Discussions. Our results reveal fundamental connections between the order of generalized entanglement entropies and the order of randomness. The results motivate a definition of "scrambling complexities" in terms of degree of randomness by Rényi entanglement entropies: if the Rényi-α entanglement entropy is nearly maximal between generic partitions, then the system behaves like α-designs in terms of entanglement, or α-designs can model the entanglement properties of this system. For example, the generic maximality of min entanglement entropy indicates that the system looks completely random (and the local information is completely lost) to any local observers,which we call "max-scrambling". Then by the log moment result, it is reasonable to conjecture that the minimum time for a physical system of n degrees of freedom to max-scramble scales as $\tilde{O}(n)$ (fast max-scrambling conjecture). It would be interesting to further study the dynamics of Rényi entanglement entropies in scrambling systems. It would also be interesting to consider the connections between scrambling complexities and computational power of certain physical systems.

[*] Corresponding author: zwliu@mit.edu

[1] P. Hayden and J. Preskill, Journal of High Energy Physics **2007**, 120 (2007).
[2] Y. Sekino and L. Susskind, Journal of High Energy Physics **10**, 065 (2008), arXiv:0808.2096 [hep-th].
[3] L. Susskind, ArXiv e-prints (2011), arXiv:1101.6048 [hep-th].
[4] K. Furuya, M. C. Nemes, and G. Q. Pellegrino, Phys. Rev. Lett. **80**, 5524 (1998).
[5] A. Lakshminarayan, Phys. Rev. E **64**, 036207 (2001).

[6] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, Journal of High Energy Physics **2016**, 1 (2016).
[7] R. Nandkishore and D. A. Huse, Annual Review of Condensed Matter Physics **6**, 15 (2015).
[8] B. M. Terhal, D. P. DiVincenzo, and D. W. Leung, Phys. Rev. Lett. **86**, 5807 (2001).
[9] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, IEEE Transactions on Information Theory **48**, 580 (2002).
[10] E. Lubkin, Journal of Mathematical Physics **19**, 1028 (1978).
[11] S. Lloyd and H. Pagels, Annals of Physics **188**, 186 (1988).
[12] D. N. Page, Phys. Rev. Lett. **71**, 1291 (1993).
[13] S. K. Foong and S. Kanno, Phys. Rev. Lett. **72**, 1148 (1994).
[14] J. Sánchez-Ruiz, Phys. Rev. E **52**, 5653 (1995).
[15] S. Sen, Phys. Rev. Lett. **77**, 1 (1996).
[16] E. Knill, eprint arXiv:quant-ph/9508006 (1995), quant-ph/9508006.
[17] F. G. S. L. Brandao, A. W. Harrow, and M. Horodecki, ArXiv e-prints (2012), arXiv:1208.0692 [quant-ph].
[18] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, Phys. Rev. Lett. **116**, 170502 (2016).
[19] G. Tóth and J. J. García-Ripoll, Phys. Rev. A **75**, 042311 (2007).
[20] Y. Nakata, C. Hirche, M. Koashi, and A. Winter, ArXiv e-prints (2016), arXiv:1609.07021 [quant-ph].

No-Hypersignaling Principle

Michele Dall'Arno[1] * Sarah Brandsen[1] Alessandro Tosini[2] Francesco Buscemi[3]
Vlatko Vedral[1 4]

[1] *Centre for Quantum Technologies, National University of Singapore, Singapore*
[2] *QUIT group, Physics Dept., Pavia University, and INFN Sezione di Pavia, Italy*
[3] *Graduate School of Informatics, Nagoya University, Nagoya, Japan*
[4] *Atomic and Laser Physics, Clarendon Laboratory, University of Oxford, Oxford, U.K.*

Abstract. A paramount topic in quantum foundations, rooted in the study of the EPR paradox and Bell inequalities, is that of characterizing quantum theory in terms of the space-like correlations it allows. Here we show that to focus only on space-like correlations is not enough: we explicitly construct a toy model theory that, while not contradicting classical and quantum theories at the level of space-like correlations, still displays an anomalous behavior in its time-like correlations. We call this anomaly, quantified in terms of a specific communication game, the "hypersignaling" phenomena. We hence conclude that the "principle of quantumness," if it exists, cannot be found in space-like correlations alone: nontrivial constraints need to be imposed also on time-like correlations, in order to exclude hypersignaling theories.

Keywords: no-hypersignaling principle

This presentation is based on Ref. [1].

One of the main tenets in modern physics is that if two space-like separated events are correlated, then such correlations must not carry any information [2]. This assumption, constituting the so-called *no-signaling principle*, was the starting point used by Bell [3] to quantify and compare space-like correlations of different theories on even grounds—an idea of vital importance for his argument about the EPR paradox [4] and the derivation of his famous inequality. Subsequently, due to seminal works by Tsirelson (Cirel'son) [5] and Popescu and Rohrlich [6], it became clear that the no-signaling principle alone is not enough to characterize "physical" space-like correlations: non-signaling space-like correlations allowed by quantum theory form a *strict* subset within the set of all non-signaling correlations [7].

A natural question is then to try to identify additional principles that, together with the no-signaling principle, may be able to rule out all super-quantum non-signaling correlations at once. Various ideas have been proposed, ranging from complexity theory, e.g. the collapse of the complexity tower [8] to information theory, e.g. the information causality principle [9]. However, none of these has been able to characterize the quantum/super-quantum boundary in full. In particular, an outstanding open question is whether quantum theory can be characterized in terms of the space-like correlations it allows [7].

In this presentation, we show that this cannot be done: any approach to characterize quantum theory based only on space-like correlations is necessarily incomplete unless it also takes into account time-like correlations as well. The characterization of time-like correlations is part of the program [10] of general probabilistic theories aimed at reconstructing operational features of quantum theory. Our approach, which is completely unrelated to the study of temporal correlations *à la* Leggett–Garg [11, 12, 13, 14], considers the elementary resource of noiseless communication and the input/output correlations that can be so established. By analogy with the no-signaling principle, we operationally introduce what we call the "no-hypersignaling principle," which roughly states that any input/output correlation that can be obtained by transmitting a composite system should also be obtainable by independently transmitting its constituents. As obvious as this may look (it is indeed so in classical and quantum theories), the fact that quantum theory obeys the no-hypersignaling principle (as we define it) is in fact a highly nontrivial consequence of a recent result by Frenkel and Weiner [15]. We also notice that the no-hypersignaling principle is not related with phenomena such as superadditivity of capacities of noisy quantum channels [16].

We then construct a toy model theory, that we refer to as the HS model, which violates the no-hypersignaling principle, but only possesses classical space-like correlations. As such, this theory (and other analogous theories) would go undetected in any test involving only space-like correlations, despite displaying the anomalous effect of hypersignaling. On the technical side, the HS model is closely related to the standard implementation [17, 18, 19] of Popescu–Rohrlich [6] super-quantum non-signaling space-like correlations (or "PR-boxes," for short). However, while the PR-box model theory relies on entangled states to outperform quantum *space*-like correlations, our HS model relies on *entangled measurements* to outperform quantum *time*-like correlations. Nonetheless, since in our model only separable states are available, no super-quantum space-like correlation can be obtained. Therefore, while the standard PR-box model theory can be ruled out on the basis of its super-quantum space-like correlations, the HS model can only be ruled out by the principle of no-hypersignaling.

It is now important to understand how hypersignaling is logically related with other possible "anomalies," such as the violation of local tomography or the violation of information causality. If any hypersignaling theory neces-

*cqtmda@nus.edu.sg

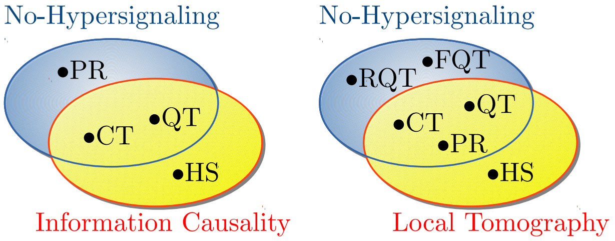


Figure 1: **No-Hypersignaling vs Information Causality and vs Local Tomography**. Left: the diagram compares theories satisfying information causality (yellow set) and the no-hypersignaling principle (blue set): CT (classical theory), QT (quantum theory), PR Model (the toy model theory for PR-boxes), and HS Model (the locally classical, hypersignaling theory constructed in this paper). Right: comparison between local tomography and no-hypersignaling as two features of general probabilistic theories. Examples of theories that are non-hypersignaling but violate local tomography are provided by real quantum theory (RQT) and fermionic quantum theory (FQT). The HS Model is locally tomographic but hypersignaling. Finally CT, QT, and the PR Model lie in the intersection, as they obey both local tomography and the no-hypersignaling principle.

sarily violates also other principles concerning space-like correlations, then one could rightly argue that the phenomenon of hypersignaling might be ruled out just by looking at space-like correlations. However, the point of this paper is to argue the opposite: that time-like correlations require a new *independent* principle.

The fact that hypersignaling and information causality are independent is easy to see. As a necessary condition for the violation of information causality is the presence of entangled states, and since the HS Model only contains separable states, then the HS Model necessarily obeys information causality, despite allowing hypersignaling. Vice versa, we know that the PR Model violates information causality but, since it only allows separable measurements, it cannot display any form of hypersignaling. The situation is depicted in Left Fig. 1.

We now turn to the condition of local tomography [20]. From the explicit expression of the pure states of the HS Model, it is possible to verify that the elementary system S has linear dimension $\ell(S) = 3$ and that the bipartite system $S \otimes S$ has linear dimension $\ell(S \otimes S) = 9 = \ell(S)^2$. Thus the HS Model is locally tomographic, despite being hypersignaling. Vice versa, there exist consistent theories that obey the no-hypersignaling principle and yet are not locally tomographic. As an example, let us consider restrictions (for example, superselections) of quantum theory, as introduced in Ref. [21]. Since such theories are restrictions of quantum theory, they cannot exhibit hypersignaling: if they did, then quantum theory would also exhibit hypersignaling, which is not true. For example, real quantum theory [20] and fermionic quantum theory [21] are two possible such restricted quantum theories. However, as proved in Refs. [21, 22, 20], both theories are not locally tomographic. The situation is summarized in Right Fig. 1.

We also notice that the no-hypersignaling principle can be violated by theories that do not show superadditivity of classical capacities. In Ref. [23] the authors show that a locally tomographic theory cannot feature superadditivity effects of classical capacities. Thus hypersignaling does not necessarily imply superadditivity of classical capacities, because the HS Model is locally tomographic.

One interesting question arises from noting that while the HS Model has classical space-like correlations and super-quantum time-like correlations, the PR Model has super-quantum space-like correlations and classical time-like correlations. Could it be that a theory can be super quantum only with respect to either space-like or time-like correlations, but not both? Could quantum theory have the unique distinction of “balancing” between these two extrema? It turns out that the answer is no, and follows from the example of the Hybrid Models derived in Ref. [1]. In order to obtain the hypersignaling correlation in Ref. [1] we need seven factorized states and seven effects among which only one is not factorized. Since such an entangled effect is exactly one of those admitted in the Hybrid Models, we know that the same hypersignaling correlation can be surely obtained in those models too. Moreover, since in the Hybrid Models two entangled states are also available, super-quantum spacelike correlations can also be created. Hence, the Hybrid Models have the ability to create both space-like and time-like super-quantum correlations.

Finally, we compare the no-hypersignaling principle with two recently proposed and related principles, that is, *dimension mismatch* [24] and *information content* [25]. Both such principles rule out superquantum theories on the basis of the correlations achievable by a single-partite system, in contrast with the no-hypersignaling principle which requires composite systems. However, they achieve this by considering a more complicated setup, where the choice of the information to be decoded is not fixed but depends on an additional input (a second question) to the receiver. Moreover, both the dimension mismatch principle and the information content principle rely on a certain degree of arbitrariness in the criteria chosen to benchmark operational theories: dimension mismatch is defined with respect to an arbitrarily chosen reference task, i.e. pairwise state discrimination, while information content is defined with respect to an arbitrarily chosen information measure, i.e. mutual information. This is in contrast with the no-hypersignaling principle proposed here, where the full set of input-output correlations is considered without the need to invoke any particular discrimination task or information measure. Finally, we notice that dimension mismatch, to be meaningful, requires the existence of perfectly distinguishable states, which is not guaranteed without the so-called “no-restriction assumption.”

References

[1] M. Dall’Arno, S. Brandsen, A. Tosini, F. Buscemi, V. Vedral, *No-Hypersignaling Principle*, Phys. Rev. Lett.

119, 020401, also available at arXiv:1609.09237.

[2] A. Einstein, *Zur Elektrodynamik bewegter Körper*, Annalen der Physik **17**, 891 (1905).

[3] J. Bell, *On the Einstein Podolsky Rosen Paradox*, Physics **1**, 195 (1964).

[4] A. Einstein, B. Podolsky, & N. Rosen, *Can quantum-mechanical description of physical reality be considered complete?*, Phys. Rev. **47**, 777 (1935).

[5] B.S. Tsirelson, *Quantum generalizations of Bells inequality*, Lett. Math. Phys. **4**, 93–100 (1980).

[6] S. Popescu, & D. Rohrlich, *Quantum Nonlocality as an Axiom*, Found. Phys. **24**, 379 (1994).

[7] S. Popescu, *Nonlocality beyond Quantum Mechanics*, Nature Physics **10**, 264–70 (2014).

[8] W. van Dam, *Implausible consequences of superstrong nonlocality*, Nat. Comput. **12**, 9 (2013).

[9] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, & M. Żukowski, *Information causality as a physical principle*, Nature **461**, 1101 (2009).

[10] G. Chiribella, & X. Yuan, *Bridging the gap between general probabilistic theories and the device-independent framework for nonlocality and contextuality*, Information and Computation, 250, 15-49 (2016).

[11] S. Brierley, A. Kosowski, M. Markiewicz, T. Paterek, & A. Przysiezna, *Nonclassicality of Temporal Correlations*, Phys. Rev. Lett. **115**, 120404 (2015).

[12] C. Budroni, & C. Emary, *Temporal Quantum Correlations and Leggett-Garg Inequalities in Multilevel Systems*, Phys. Rev. Lett. **113**, 050401 (2014).

[13] C. Budroni, T. Moroder, M. Kleinmann, & O. Gühne, *Bounding Temporal Quantum Correlations*, Phys. Rev. Lett. **111**, 020403 (2013).

[14] M. Markiewicz, P. Kurzyński, J. Thompson, S.-Y. Lee, A. Soeda, T. Paterek, & D. Kaszlikowski, *Unified approach to contextuality, nonlocality, and temporal correlations*, Phys. Rev. A **89**, 042109 (2014).

[15] P.E. Frenkel, & M. Weiner, *Classical Information Storage in an n-Level Quantum System*, Commun. Math. Phys. **340**, 563 (2015).

[16] M.B. Hastings, *Superadditivity of communication capacity using entangled inputs*, Nature Physics **5**, 255–257 (2009).

[17] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, & D. Roberts, *Nonlocal correlations as an information-theoretic resource*, Phys. Rev. A **71**, 022101 (2005).

[18] J. Barrett, *Information processing in generalized probabilistic theories* , Phys. Rev. A **75**, 032304 (2007).

[19] G.M. D'Ariano, & A. Tosini, *Testing axioms for quantum theory on probabilistic toy-theories*, Quant. Inf. Proc. **9**, 95 (2010).

[20] L. Hardy, & W.K. Wootters, *Limited Holism and Real-Vector-Space Quantum Theory*, Found. Phys. **42**, 454 (2012).

[21] G.M. D'Ariano, F. Manessi, P. Perinotti, & A. Tosini, *The Feynman problem and Fermionic entanglement: Fermionic theory versus qubit theory*, Int. J. Mod. Phys. A **29**, 1430025 (2014).

[22] G.M. D'Ariano, F. Manessi, P. Perinotti, & A. Tosini, *Fermionic computation is non-local tomographic and violates monogamy of entanglement*, Europhys. Lett. **107**, 20009 (2014).

[23] S. Massar, S. Pironio, & D. Pitalúa-García, *Hyperdense coding and superadditivity of classical capacities in hypersphere theories*, New J. Phys. **17**, 113002 (2015).

[24] N. Brunner, M. Kaplan, A. Leverrier, & P. Skrzypczyk, *Dimension of physical systems, information processing, and thermodynamics*, New J. Phys. **16**, 123050 (2014).

[25] L. Czekaj, M. Horodecki, P. Horodecki, & R. Horodecki, *Information Content of Systems as a Physical Principle*, Phys. Rev. A **95**, 022119 (2017).

A generalized quantum Slepian-Wolf

Anurag Anshu[1] * Rahul Jain[2] † Naqueeb Ahmad Warsi[3] ‡

[1] *Centre for Quantum Technologies, National University of Singapore, Singapore.*
[2] *Centre for Quantum Technologies, National University of Singapore and MajuLab, UMI 3654, Singapore.*
[3] *Centre for Quantum Technologies and SPMS, NTU, Singapore and IIITD, Delhi.*

Abstract. In this work we consider a quantum generalization of the task considered by Slepian and Wolf [1] regarding distributed source compression. In our task Alice, Bob, Charlie and Referee share a joint pure state. Alice and Bob wish to send a part of their respective systems to Charlie without collaborating with each other. We give achievability bounds for this task in the one-shot setting and provide asymptotic analysis in the case when there is no side information with Charlie. Our result implies the result of Abeyesinghe, Devetak, Hayden and Winter in [2] who studied a special case of this problem. As another special case wherein Bob holds trivial registers, we recover the result of Devetak and Yard [3] regarding quantum state redistribution.

Keywords: Coherent quantum protocols, Quantum Slepian Wolf, Quantum side information, Quantum information theory

In information theory, one of the most fundamental problems is the task of source-compression. The answer to this problem was given by Shannon in his celebrated work [4]. Slepian and Wolf, in their work [1], studied this task in the distributed network setting, which consists of three parties Alice $(X_1, X_2 \dots X_n\ )$, Bob $(Y_1, Y_2 \dots Y_n)$ and Charlie, where $(X_1, Y_1), (X_2, Y_2), \dots (X_n, Y_n)$ are pairs of independent and identically distributed correlated random variables. The goal here is that Alice needs to communicate $(X_1, X_2, \dots X_n)$ to Charlie and similarly, Bob needs to communicate $(Y_1, Y_2, \dots Y_n)$ to Charlie. Furthermore, Alice and Bob do not collaborate. From Shannon's result, one can easily see that the amount of total communication sufficient to accomplish this task is $nH(X)+nH(Y)$. However, the surprising feature of the result of Slepian and Wolf is that the amount of total communication only needs to be $nH(XY)$. Furthermore, their result implies that there is a trade-off on the amount of communication between (Alice, Charlie) and (Bob, Charlie).

The quantum version of this problem was studied by Abeyesinghe, Devetak, Hayden and Winter in [2]. In this setting, there are four parties, Alice (M), Bob (N), Charlie and Referee (R), where Referee serves as a purifying system for Alice and Bob. The goal is that Alice needs to communicate the register M to Charlie and Bob needs to communicate the register N to Charlie, such that the final quantum state between Referee and Charlie is close to the original pure state between Referee, Alice and Bob. The work [2] studied above task in the asymptotic and i.i.d setting. The authors introduced a protocol termed *Fully Quantum Slepian-Wolf* and combined it with Schumacher's compression [5] (using the notion of *time-sharing*) to obtain a rate pair.

The emerging framework of one-shot information theory is providing a new perspective on data compression and channel coding and is also relevant in the practical scenarios. This framework also provides insights into the conceptual details of information theoretic protocols, as the notational complications arising due to many copies of the state are no longer present (although we note that asymptotic and i.i.d setting also has its own conveniences). One-shot information theory also has found applications in both classical communication complexity [6, 7] and quantum communication complexity [8]. Many quantum tasks have been formulated in their one-shot setting, such as quantum state merging ([9, 10], originally introduced in [11]) and quantum state redistribution ([12, 13, 14], originally introduced in [3, 15]).

Given the importance of one-shot information theory, in this work we consider the one-shot version of the problem studied in [2]. To capture a more general scenario, along with the registers M, N we also allow Alice, Bob and Charlie to have additional registers A, B, C respectively. Thus, our setting is as follows, depicted in Figure 1.

Task: Alice (AM), Bob (BN), Charlie (C) and Referee (R) share a joint pure quantum state. The goal is that Alice needs to communicate the register M to Charlie and Bob needs to communicate the register N to Charlie, such that the final quantum state between Referee (R), Alice (A), Bob (B) and Charlie (CMN) is close to the original pure state between the parties. We allow preshared entanglement between (Alice, Charlie) and (Bob, Charlie) respectively.

This task is a natural generalization of the aforementioned task and also extends the well studied problem of quantum state redistribution [3, 15]. A special case when A is trivial was considered by [16] in which they studied the trade-off between amount of *entanglement consumed* between Alice and Charlie and communication between Bob and Charlie.

Our Results: Our one shot result is mentioned as Theorem 1 towards the end of this abstract. We emphasize on two main ingredients: first is that the rate region appears in terms of max-relative entropy and hypothesis testing

*a0109169@u.nus.edu
†rahul@comp.nus.edu.sg
‡warsi.naqueeb@gmail.com

relative entropy. Second ingredient is that the rate region is a union of a family of rate regions, each characterized by a quantum state that is close to original state Ψ and satisfies some max-relative entropy constraints.

Using this, we are able to obtain the following rate region in the asymptotic i.i.d setting when C is trivial:

$$\begin{aligned}
R_{A\to C} &\geq \frac{1}{2}\left(\mathrm{I}(RAB:M)-\mathrm{I}(A:M)\right),\\
R_{B\to C} &\geq \frac{1}{2}\left(\mathrm{I}(RAB:N)-\mathrm{I}(B:N)\right),\\
R_{A\to C}+R_{B\to C} &\geq \frac{1}{2}\Big(\mathrm{I}(RAB:M:N)\\
&\qquad -\mathrm{I}(A:M)-\mathrm{I}(B:N)\Big),
\end{aligned}$$

where $R_{A\to C}$ is the rate of quantum communication from Alice to Charlie, $R_{B\to C}$ is the rate of quantum communication from Bob to Charlie and all the information theoretic quantities calculated above are with respect to the state Ψ_{RABMN} shared between Alice, Bob and Referee. The quantity $\mathrm{I}\left(RAB:M:N\right)$ is the tripartite quantum mutual information, defined as $S(\Psi_{RAB})+S(\Psi_M)+S(\Psi_N)-S(\Psi_{RABMN})$, where $S(.)$ is the von-Neumann entropy.

An immediate consequence of the above result is the rate pair obtained for the task considered in [2], with registers A, B being trivial. Moreover, if registers B, N are trivial in the original task, then the task reduces to that of quantum state redistribution. In this case, the result of Theorem 1 also reproduces the bound given in [3, 15] for quantum state redistribution in asymptotic and i.i.d. setting.

Techniques: Along with the inherent challenges of one-shot information theory, an additional challenge for extending the result of [2] is the absence of the notion of time sharing in the one-shot case. The idea of time-sharing is as follows: given two rates $R=(R_1,R_2)$ and $R'=(R'_1,R'_2)$ at which Alice and Bob can communicate to Charlie, one can construct a protocol which achieves the rate $\alpha R+(1-\alpha)R'$ by using the first protocol for the first αn copies and using the second protocol for the last $(1-\alpha)n$ copies (see [17, Page 534]).

It is clear that this technique cannot extend to the one-shot setting which considers just one copy of input state. We overcome the obstacle of time sharing in the one-shot case by using the technique of *convex-split* [13] along with *position-based decoding* [18]. The convex-split technique allows one party to prepare a convex combination of states on the registers of other party, if the first party holds a purification of the registers of the second party. The concept of position-based decoding is essentially hypothesis testing on a global state.

Technical contribution of this work resides in two aspects. First is that we prove a new version of convex-split lemma [13, Page 3], which we refer to as *tri-partite* convex-split lemma, which requires Charlie to prepare a convex combination of quantum states shared between three parties Referee, Alice and Bob. We prove the sufficient conditions which allow Charlie to prepare such convex combination with small error. Second technical contribution is in our asymptotic analysis of the one-shot bounds. It can be seen that the time-sharing technique, along with the quantum state redistribution protocol of [3, 15], obtains the asymptotic achievability result mentioned above [1]. Since our one-shot result has no time-sharing involved, we provide an explicit analysis of our bound when there are many independent copies of the state Ψ shared between the parties, in the case where register C is absent. For this, we exploit several properties of the quantum information spectrum relative entropy (introduced in [19, 20]; the classical information spectrum approach originated in [21]) to show the existence of a quantum state that is close to the original state Ψ and satisfies several max-entropy constraints on the reduced systems. A special case of this analysis has also appeared in the context of quantum channel coding for broadcast channel in our work [18], suggesting a wide applicability of the techniques developed in the proof.

Following is our main result. Its proof can be found in the extended version of this work [22].

Theorem 1 *Fix $\varepsilon_1,\varepsilon_2,\delta>0$. Let Alice ($AM$), Bob ($BN$), Referee ($R$) and Charlie ($C$) share the pure state $|\Psi\rangle_{RAMBNC}$. There exists an entanglement assisted quantum protocol, with entanglement shared only between (Alice, Charlie) and (Bob, Charlie), such that at the end of the protocol, Alice (A), Bob (B), Referee (R) and Charlie (CMN) share the state Φ'_{RAMBNC} with the property that $\mathrm{P}(\Phi',|\Psi\rangle\langle\Psi|)\leq\varepsilon_1+5\varepsilon_2+2\sqrt{\delta}$. The number of qubits that Alice sends to Charlie is $R_{A\to C}$ and that Bob sends to Charlie is $R_{B\to C}$, where the pair $(R_{A\to C},R_{B\to C})$ lie in the union of the following rate region: for every $\Psi'_{RAMBNC}\in\mathcal{B}^{\varepsilon_1}(\Psi_{RAMBNC})$ such that $\Psi'_{RAB}\preceq 2^{\delta}\Psi_{RAB}$ and states σ_M,ω_N:*

$$\begin{aligned}
R_{A\to C} &\geq \frac{1}{2}\Big(\mathrm{D}_{\max}(\Psi'_{RABM}\|\Psi_{RAB}\otimes\sigma_M)\\
&\qquad -\mathrm{D}_{\mathrm{H}}^{\varepsilon_2^2}(\Psi_{AM}\|\Psi_A\otimes\sigma_M)+\log\frac{1}{\varepsilon_2^2\delta}\Big),\\
R_{B\to C} &\geq \frac{1}{2}\Big(\mathrm{D}_{\max}(\Psi'_{RABN}\|\Psi_{RAB}\otimes\omega_N)\\
&\qquad -\mathrm{D}_{\mathrm{H}}^{\varepsilon_2^2}(\Psi_{BN}\|\Psi_B\otimes\omega_N)+\log\frac{1}{\varepsilon_2^2\delta}\Big),\\
R_{A\to C}+R_{B\to C} &\geq \frac{1}{2}\Big(\mathrm{D}_{\max}(\Psi'_{RABMN}\|\Psi_{RAB}\otimes\sigma_M\otimes\omega_N)\\
&\qquad -\mathrm{D}_{\mathrm{H}}^{\varepsilon_2^2}(\Psi_{AM}\|\Psi_A\otimes\sigma_M)\\
&\qquad -\mathrm{D}_{\mathrm{H}}^{\varepsilon_2^2}(\Psi_{BN}\|\Psi_B\otimes\omega_N)+\log\frac{1}{\varepsilon_2^2\delta}\Big).
\end{aligned}$$

[1] The extremal points of the rate region are $(R_{A\to C},R_{B\to C})=\left(\frac{1}{2}\mathrm{I}\left(RB:M|NC\right),\frac{1}{2}\mathrm{I}\left(RAM:N|C\right)\right)$ and $(R_{A\to C},R_{B\to C})=\left(\frac{1}{2}\mathrm{I}(RBN:M|C),\frac{1}{2}\mathrm{I}(RA:N|MC)\right)$. The first can be achieved by Bob sending N to Charlie using quantum state redistribution, followed by Alice sending M to Charlie, again using quantum state redistribution. Second can be achieved in analogous fashion. Any rate pair can then be achieved by time sharing between these two protocols.

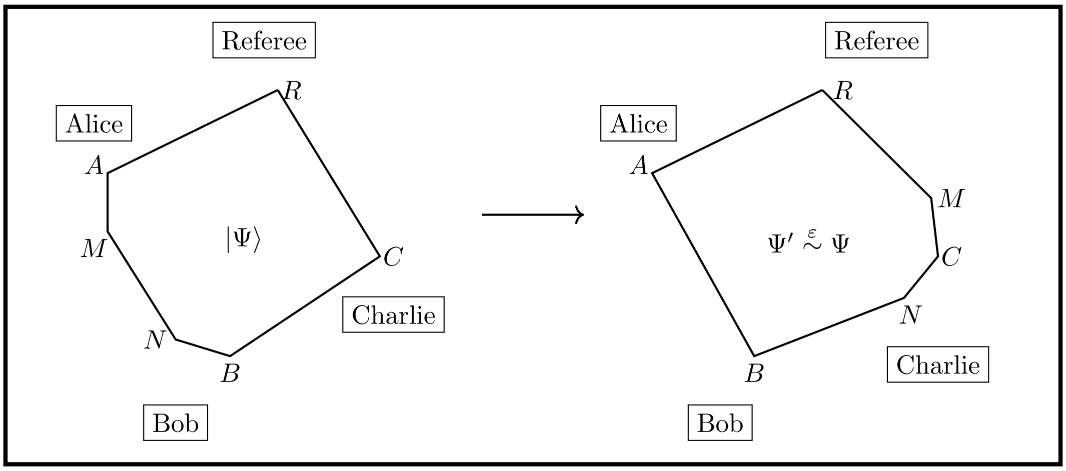


Figure 1: The task of a generalized quantum slepian wolf.

Above, $\mathrm{P}(\rho, \sigma)$ is the purified distance between quantum states ρ, σ, $\mathrm{D}_{\max}(\rho \| \sigma)$ is the max-relative entropy between quantum states ρ, σ and $\mathcal{B}^{\varepsilon}(\rho) := \{\sigma : \mathrm{P}(\rho, \sigma) \leq \varepsilon\}$.

References

[1] D. Slepian and J. Wolf, *IEEE Transactions on Information Theory*, vol. 19, pp. 471–480, Jul 1973.

[2] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 465, no. 2108, pp. 2537–2563, 2009.

[3] I. Devetak and J. Yard, *Phys. Rev. Lett.*, vol. 100, no. 230501, 2008.

[4] C. E. Shannon, *The Bell System Technical Journal*, vol. 27, pp. 379–423, July 1948.

[5] B. Schumacher, *Phys. Rev. A.*, vol. 51, pp. 2738–2747, 1995.

[6] P. Harsha, R. Jain, D. Mc.Allester, and J. Radhakrishnan, *IEEE Transcations on Information Theory*, vol. 56, pp. 438–449, 2010.

[7] M. Braverman and A. Rao, in *Proceedings of the 52nd Symposium on Foundations of Computer Science*, FOCS '11, (Washington, DC, USA), pp. 748–757, IEEE Computer Society, 2011.

[8] D. Touchette, *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, STOC '15, (New York, NY, USA), pp. 317–326, ACM, 2015.

[9] M. Berta, Master's thesis, ETH Zurich, http://arxiv.org/abs/0912.4495, 2009.

[10] M. Berta, M. Christandl, and R. Renner, *Commun. Math. Phys.*, vol. 306, no. 3, pp. 579–615, 2011.

[11] M. Horodecki, J. Oppenheim, and A. Winter, *Communications in Mathematical Physics*, vol. 269, no. 1, pp. 107–136, 2007.

[12] M. Berta, M. Christandl, and D. Touchette, http://arxiv.org/abs/1409.4338, 2014.

[13] A. Anshu, V. Devabathini, and R. Jain, http://arxiv.org/abs/1410.3031, 2014.

[14] A. Anshu, R. Jain, and N. Warsi, https://arxiv.org/abs/1702.02396, 2017.

[15] J. T. Yard and I. Devetak, *IEEE Transactions on Information Theory*, vol. 55, pp. 5339–5351, 2009.

[16] M. H. Hsieh and S. Watanabe, in *2015 IEEE Information Theory Workshop - Fall (ITW)*, pp. 307–311, Oct 2015.

[17] T. M. Cover and J. A. Thomas, *Elements of information theory.* Wiley Series in Telecommunications, New York, NY, USA: John Wiley & Sons, 1991.

[18] A. Anshu, R. Jain, and N. Warsi, https://arxiv.org/abs/1702.01940, 2017.

[19] M. Hayashi and H. Nagaoka, *IEEE Transactions on Information Theory*, vol. 49, pp. 1753–1768, July 2003.

[20] H. Nagaoka and M. Hayashi, *IEEE Transactions on Information Theory*, vol. 53, pp. 534–549, Feb 2007.

[21] T. S. Han and S. Verdu, *IEEE Transactions on Information Theory*, vol. 39, pp. 752–772, May 1993.

[22] A. Anshu, R. Jain, and N. Warsi, https://arxiv.org/abs/1703.09961, 2017.

Fundamental rate-loss trade-off for the quantum internet

Koji Azuma[1] * Akihiro Mizutani[2] † Hoi-Kwong Lo[3] ‡

[1] *NTT Basic Research Laboratories, NTT Corporation, 3-1, Morinosato-Wakamiya Atsugi-Shi, 243-0198, Japan*
[2] *Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan*
[3] *Department of Physics and Department of Electrical & Computer Engineering, University of Toronto, Toronto, Ontario, Canada*

Abstract. The quantum internet holds promise for achieving quantum communication freely between any clients all over the globe. The most primitive function of the quantum internet is to provide quantum entanglement or a secret key to two points efficiently, by using intermediate nodes connected by optical channels with each other. Here we derive a fundamental rate-loss trade-off for a quantum internet protocol, by generalizing the Takeoka-Guha-Wilde bound to be applicable to any network topology. Our result—putting a practical but general limitation on the quantum internet—enables us to grasp the potential of the future quantum internet.

Keywords: quantum internet, Takeoka-Guha-Wilde bound, squashed entanglement

1 Introduction

In the conventional Internet, if a client, Alice, wants to communicate with another client, Bob, an Internet protocol determines the path that the data follow to travel across multiple networks from Alice to Bob. Analogously, in the future, according to a request for performing quantum communication between Alice and Bob, a quantum internet [1] protocol will supply the resources—such as a secret key (secret bits) for the purpose of the unconditionally secure communication and quantum entanglement (ebits) for the purpose of the quantum teleportation—to Alice and Bob by utilizing proper intermediate nodes connected by optical channels—for instance, optical fibres—with each other.

In this presentation, we present [2] a general, fundamental and practical limitation on any two-party quantum communication over any quantum network composed of arbitrary quantum nodes and arbitrary quantum channels connecting the nodes. In particular, we derive an upper bound on obtainable ebits or secret bits between arbitrary two clients involved in the quantum network by using arbitrary combination of the quantum channels and local operations and classical communication (LOCC) among the nodes (see Fig. 1 for the detail). This achievement is notable in the sense that a priori working out upper bounds on secure key rates and entanglement generation rates for a general quantum internet topology is highly non-trivial because there are many intermediate nodes, various elements such as quantum memories and many different protocols such as entanglement generation, entanglement swapping, entanglement distillation and quantum error correction.

2 Main results

Here, we present the upper bound for the general quantum internet protocol. To obtain our bound, we need to define a general paradigm of two-party communication over the quantum internet (see Fig. 1a). In the quantum internet, there are a variety of quantum channels connecting nodes, for example, depending on the lengths of optical channels. This necessitates to generalize the paradigm [3, 4] of Takeoka *et al.* for the point-to-point communication, where it has been enough to treat only one optical channel between Alice and Bob. For instance, we need to allow the choice of which channel to use in the next round to depend on the outcomes of LOCC operations in previous rounds, in contrast to the paradigm of Takeoka *et al.*

To make this more precise, let us define the general protocol. We assume that any classical communication over the network is freely usable. Suppose that Alice (A) and Bob (B) call a quantum internet protocol to share a resource for quantum communication, unconditionally secure key or quantum entanglement, over the quantum network. Accordingly, the quantum internet protocol determines a subnetwork to supply the resource to Alice and Bob. The subnetwork is characterized by a directed graph $G = (V, E)$ with a set V of vertices and a set E of edges, where the vertices of G represent Alice's node, Bob's node and intermediate nodes $\{C^k\}_{k=1,2,\ldots,n}$ in the subnetwork, *i.e.*, $V = \{A, B, C^1, C^2, \ldots, C^n\}$, and an edge $\varepsilon = v_1 \to v_2 \in E$ of G for $v_1, v_2 \in V$ specifies a quantum channel $\mathcal{N}^{v_1 \to v_2}$ to send a quantum system from node v_1 to node v_2 in the subnetwork. Then, the most general protocol proceeds in an adaptive manner as follows [c.f. Fig. 1b which exemplifies a linear network with $n = 4$]. The protocol starts by preparing the whole system in a separable state $\hat{\rho}_1^{ABC^1C^2\ldots C^n}$ and then by using a quantum channel $\mathcal{N}^{e_1}$ with $e_1 \in E$. This is followed by arbitrary LOCC among all the nodes, which gives an outcome k_1 and a quantum state $\hat{\rho}_{k_1}^{ABC^1C^2\ldots C^n}$ with probability p_{k_1}. In the second round, depending on the outcome k_1, a node uses a quantum channel $\mathcal{N}^{e_{k_1}}$ with $e_{k_1} \in E$, followed by LOCC among all the nodes. This LOCC gives an outcome k_2 and a quantum state $\hat{\rho}_{k_2k_1}^{ABC^1C^2\ldots C^n}$ with probability $p_{k_2|k_1}$. Similarly, in the

*azuma.koji@lab.ntt.co.jp
†mizutani@qi.mp.es.osaka-u.ac.jp
‡hklo@ece.utoronto.ca

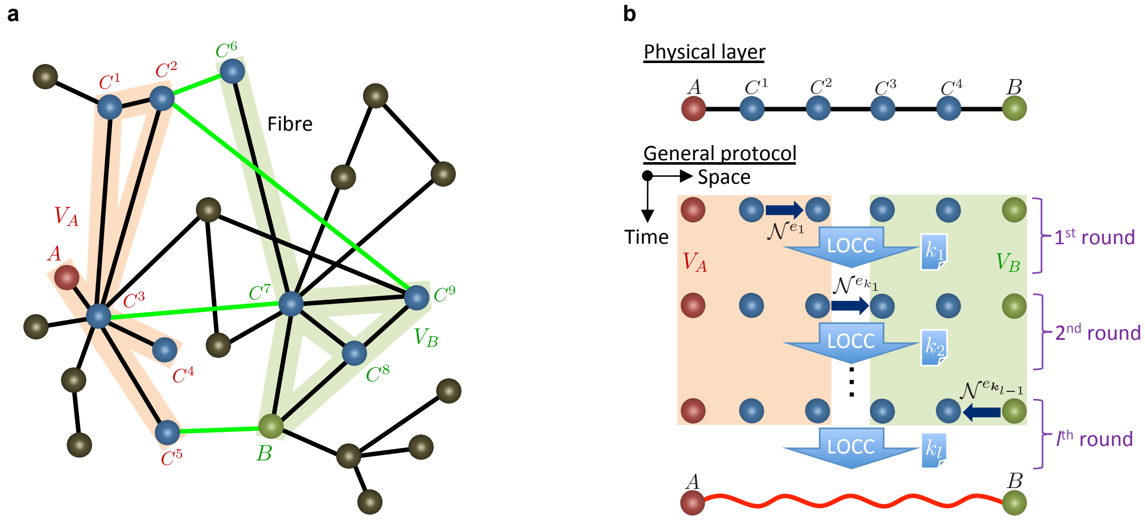


Figure 1: **Quantum internet (Reproduced from [2]).** Panel (a) depicts a general quantum internet where Alice (A) and Bob (B) request its internet protocol to supply them with the resources for quantum communication, such as a secret-key and quantum entanglement. Accordingly, the protocol chooses a quantum network G (which might be a quantum subnetwork) associated with a directed graph $G = (V, E)$. The set V of vertices is composed of the nodes as $V = \{A, B, C^1, C^2, \ldots, C^n\}$ ($n = 9$ in this panel) and the set E of edges specifies quantum channels $\{\mathcal{N}^e\}_{e\in E}$ in such a way that $\mathcal{N}^{v_1\to v_2}$ represents a quantum channel to send a quantum system from node $v_1 \in V$ to node $v_2 \in V$. The protocol can combine the quantum channels $\{\mathcal{N}^e\}_{e\in E}$ with LOCC arbitrarily, to provide the required resources for Alice and Bob. However, our bound suggests that the obtainable secret bits or ebits are upper bounded by a bound for the point-to-point communication between a single parity having nodes $V_A \subset V$ with A and another party having $V_B(= V \setminus V_A)$ with B. In panel (b), we describe the paradigm of the most general two-party communication protocols, by exemplifying a linear network with $n = 4$. In the i-th round ($i = 1, 2, \ldots, l$), according to the previous outcomes $\boldsymbol{k}_{i-1} = k_{i-1} \ldots k_2k_1$, the protocol may use a quantum channel $\mathcal{N}^{e_{\boldsymbol{k}_{i-1}}}$ with $e_{\boldsymbol{k}_{i-1}} \in E$, followed by LOCC providing a quantum state $\hat{\rho}^{ABC^1C^2\ldots C^n}_{\boldsymbol{k}_i}$ with a new outcome k_i. After an l-th round, Alice and Bob obtain a quantum state $\hat{\rho}^{ABC^1C^2\ldots C^n}_{\boldsymbol{k}_l}$, from which they can distill $\log_2 d_{\boldsymbol{k}_l}$ ebits or secret bits approximately.

i-th round, according to the previous outcomes $\boldsymbol{k}_{i-1} := k_{i-1} \ldots k_2k_1$ (with $\boldsymbol{k}_0 := 1$), the protocol uses a quantum channel $\mathcal{N}^{e_{\boldsymbol{k}_{i-1}}}$ with $e_{\boldsymbol{k}_{i-1}} \in E$, followed by LOCC providing a quantum state $\hat{\rho}^{ABC^1C^2\ldots C^n}_{\boldsymbol{k}_i}$ with a new outcome k_i with probability $p_{k_i|\boldsymbol{k}_{i-1}}$. After a finite number of rounds, say after an l-th round, the protocol must present $\hat{\rho}^{AB}_{\boldsymbol{k}_l} = \mathrm{Tr}_{C^1C^2\ldots C^n}(\hat{\rho}^{ABC^1C^2\ldots C^n}_{\boldsymbol{k}_l})$ close to a target state $\hat{\tau}^{AB}_{d_{\boldsymbol{k}_l}}$ with rank $d_{\boldsymbol{k}_l}$ in the sense of $||\hat{\rho}^{AB}_{\boldsymbol{k}_l} - \hat{\tau}^{AB}_{d_{\boldsymbol{k}_l}}||_1 \leq \epsilon$ for $\epsilon > 0$, from which Alice and Bob can distil $\log_2 d_{\boldsymbol{k}_l}$ secret bits for the purpose of the unconditionally secure communication or $\log_2 d_{\boldsymbol{k}_l}$ ebits for the purpose of the quantum teleportation. After all, the protocol results in presenting $\log_2 d_{\boldsymbol{k}_l}$ secret bits or ebits with probability $p_{\boldsymbol{k}_l}$ by using quantum channels $\{\mathcal{N}^{e_{\boldsymbol{k}_i}}\}_{i=0,1,\ldots,l-1}$, where $p_{\boldsymbol{k}_i} := p_{k_i|\boldsymbol{k}_{i-1}} \cdots p_{k_3|\boldsymbol{k}_2}p_{k_2|k_1}p_{k_1}$.

For this general adaptive protocol, our main result is described as follows. Let us divide set V into two disjoint sets, V_A including A and V_B including B, satisfying $V_A = V \setminus V_B$ and $V_B = V \setminus V_A$ [c.f. Fig. 1 for the examples]. If $\mathcal{N}^{e_{\boldsymbol{k}_i}}$ is a channel between a node in V_A and a node in V_B, we write $\boldsymbol{k}_i \in K_{V_A\leftrightarrow V_B}$. For example, $\boldsymbol{k}_1 \in K_{V_A\leftrightarrow V_B}$ in Fig. 1b. Then, for any choice of V_A and V_B, the most general protocol has a limitation described by

$$\sum_{\boldsymbol{k}_l} p_{\boldsymbol{k}_l} \log_2 d_{\boldsymbol{k}_l} \leq \sum_{i=0}^{l-1} \sum_{\boldsymbol{k}_i\in K_{V_A\leftrightarrow V_B}} p_{\boldsymbol{k}_i} E_{\mathrm{sq}}(\mathcal{N}^{e_{\boldsymbol{k}_i}}) + g(\epsilon), \tag{1}$$

where g is a is a continuous function [3, 5] with the property of $\lim_{\epsilon\to 0} g(\epsilon) = 0$ and $E_{\mathrm{sq}}(\mathcal{N})$ is the squashed entanglement of channel $\mathcal{N}$ [3, 4]. This bound is reduced to $\sum_{\boldsymbol{k}_l} p_{\boldsymbol{k}_l} \log_2 d_{\boldsymbol{k}_l} \leq \sum_{i=0}^{l-1} \sum_{\boldsymbol{k}_i\in K_{V_A\leftrightarrow V_B}} p_{\boldsymbol{k}_i} E_{\mathrm{sq}}(\mathcal{N}^{e_{\boldsymbol{k}_i}})$ for $\epsilon \to 0$. The bound (1) is obtained by regarding the general multi-party protocol as bipartite communication between V_A and V_B and by applying the Takeoka-Guha-Wilde bound (TGW) to the bipartite one (see Supplementary Note 1 in [2] for the proof). Since the bound holds for any choice of V_A, the bound shows that the average of the obtained secret bits or ebits is most tightly bounded by the choice of V_A that minimizes the right-hand side of Eq. (1).

The generality of our upper bound stems from that of the TGW bound applied to any quantum channel, in contrast to Pirandola's contemporary work [6] that instead uses the Pirandola-Laurenza-Ottaviani-Banchi (PLOB) bound [7] applied only to teleportation stretchable channels in order to obtain a good bound for multipath networks composed of lossy optical channels.

As an application of our bound in Eq. (1), we consider purely optical linear networks [like Fig. 1b] composed only of lossy optical channels, and we can conclude that existing intercity QKD schemes [8, 9, 10] and quantum repeater schemes [11, 12, 13, 14, 15, 16] have no scaling gap with our upper bound, implying that our upper bound is excellent enough to conclude that there is no further big improvement on the practical schemes. In addition, we apply our bound in Eq. (1) to the Duan-Lukin-Cirac-Zoller-type (DLCZ)-type quantum repeaters [17, 18, 19, 20] with time-dependent memory decay. To do this, we first show that DLCZ-type quantum repeaters with time-dependent memory decay can be regarded as a linear quantum network [like Fig. 1b] composed of lossy optical channels and noisy qubit channels (corresponding to the model of the decay of matter quantum memories) in the spacetime. Then, by applying our upper bound to the linear network, we obtain a nontrivial fact that the coherence time of matter quantum memories should be, at least, larger than 0.1 ms—which are comparable even with the up-to-date experimental result [21] with retaining the coupling efficiency with photons—to enjoy the blessing of arbitrary DLCZ-type quantum repeaters (see Supplementary Note 2 in [2] for the detail).

3 Acknowledgements

We thank S. Guha, S. Pirandola, M. Takeoka and M. M. Wilde for valuable discussions about their papers, and S. Azuma, S. Bäuml, G. Kato, R. Namiki and K. Tamaki for helpful discussions. K.A. thanks support from the Project UQCC by the National Institute of Information and Communications Technology and from the ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan). H.-K.L. acknowledges financial support from NSERC, CFI, ORF, US Office of Naval Research (ONR) and the National Research Council (NRC) of Canada. AM acknowledges support from Grant-in-Aid for JSPS Fellows (KAKENHI Grant No. JP17J04177).

References

[1] H. J. Kimble. *Nature* **453**, 1023-1030 (2008).

[2] K. Azuma, A. Mizutani & H.-K. Lo. *Nat. Commun.* **7**, 13523 (2016).

[3] M. Takeoka, S. Guha & M. M. Wilde. *Nat. Commun.* **5**, 5235 (2014).

[4] M. Takeoka, S. Guha & M. M. Wilde. *IEEE Trans. Inf. Theory* **60**(8), 4987-4998 (2014).

[5] M. M. Wilde. *Quantum Inf. Process.* **15**, 45634580 (2016).

[6] S. Pirandola. Preprint at http://arxiv.org/abs/1601.00966 (2016).

[7] S. Pirandola, R. Laurenza, C. Ottaviani & Leonardo Banchi. *Nat. Commun.* **8**, 15043 (2017).

[8] S. Abruzzo, H. Kampermann & D. Bruß. *Phys. Rev. A* **89**, 012301 (2014).

[9] C. Panayi, M. Razavi, X. Ma & N. Lütkenhaus. *New J. Phys.* **16**, 043005 (2014).

[10] K. Azuma, K. Tamaki & W. J. Munro. *Nat. Commun.* **6**, 10171 (2015).

[11] L. Jiang, *et al. Phys. Rev. A* **79**, 032325 (2009).

[12] A. G.Fowler, *et al. Phys. Rev. Lett.* **104**, 180503 (2010).

[13] W. J. Munro, K. A. Harrison, A. M. Stephens, S. J. Devitt & K. Nemoto. *Nat. Photon.* **4**, 792-796 (2010).

[14] W. J. Munro, A. M. Stephens, S. J. Devitt, K. A. Harrison & K. Nemoto. *Nat. Photon.* **6**, 777-781 (2012).

[15] P. Mazurek, *et al. Phys. Rev. A* **90**, 062311 (2014).

[16] K. Azuma, K. Tamaki & H.-K. Lo. *Nat. Commun.* **6**, 6787 (2015).

[17] L.-M. Duan, M. D. Lukin, J. I. Cirac & P. Zoller. *Nature* **414**, 413-418 (2001).

[18] P. Kok, C. P. Williams & J. P. Dowling. *Phys. Rev. A* **68**, 022301 (2003).

[19] K. Azuma, H. Takeda, M. Koashi & N. Imoto. *Phys. Rev. A* **85**, 062309 (2012).

[20] N. Sangouard, C. Simon, N. de Riedmatten & N. Gisin. *Rev. Mod. Phys.* **83**, 33-80 (2011).

[21] S.-J. Yang, X.-J. Wang, X.-H. Bao & J.-W. Pan. *Nat. Photon.* **10**, 381-384 (2016).

Approximate broadcasting of quantum correlations

Wei Xie[1] * Kun Fang[1] † Xin Wang[1] ‡ Runyao Duan[1] [2] § ¶

[1] *Centre for Quantum Software and Information,*
Faculty of Engineering and Information Technology,
University of Technology Sydney, NSW 2007, Australia
[2] *UTS-AMSS Joint Research Laboratory for Quantum Computation and Quantum Information Processing,*
Chinese Academy of Sciences, Beijing 100190, China

Abstract. Broadcasting quantum and classical information is a basic task in quantum information processing, and is also a useful model in the study of quantum correlations including quantum discord. We establish a full operational characterization of two-sided quantum discord in terms of bilocal broadcasting of quantum correlations. Moreover, we show that both the optimal fidelity of unilocal broadcasting of the correlations in an arbitrary bipartite quantum state and that of broadcasting an arbitrary set of quantum states can be formulated as semidefinite programs (SDPs), which are efficiently computable. We also analyze some properties of these SDPs and evaluate the broadcasting fidelities for some cases of interest.

Keywords: quantum correlation, quantum discord, broadcasting, semidefinite program

1 Introduction

Copying information is a rather simple task in the classical realm, but unfortunately not in the quantum realm. It is not allowed to create an identical copy of an arbitrary unknown pure quantum state due to the no-cloning theorem [1, 2]. One can clone a set of pure states if and only if they are orthogonal. The no-broadcasting theorem [3] generalizes this result to mixed states, saying that a set of quantum states can be broadcasted if and only if the states commute with each other.

These no-go theorems can be further extended to the setting of local broadcast for composite quantum systems. Given a bipartite quantum state ρ_{AB} shared by Alice and Bob, their objective is to perform local operations only (without communication) to produce a state $\widehat{\rho}_{A_1A_2B_1B_2} = (\Lambda_{A\to A_1A_2} \otimes \Gamma_{B\to B_1B_2})\rho_{AB}$ such that $\mathrm{Tr}_{A_1B_1}\widehat{\rho}_{A_1A_2B_1B_2} = \mathrm{Tr}_{A_2B_2}\widehat{\rho}_{A_1A_2B_1B_2} = \rho_{AB}$. (See Fig. 1 for bilocal broadcasting.) It is shown in Ref. [4] that this task can only be performed if and only if ρ_{AB} is classically correlated. Even if the task is relaxed to obtain two bipartite states with the same correlation as ρ_{AB} (measured by the mutual information), it is feasible to do the task if and only if the given state ρ_{AB} is classically correlated. This is called the no-local-broadcasting theorem [4]. Furthermore, when the local operations are only allowed for one party (e.g., Alice), the task can be done if and only if ρ_{AB} is classical on A [5, 6, 7]. (See Fig. 1 for unilocal broadcasting.)

When the task of perfect broadcasting cannot be accomplished, it is natural to ask whether the broadcasting can be performed in an approximate fashion, and how to design the optimal broadcasting operation. We shall study the approximate broadcasting of states and correlations by utilizing semidefinite programs (SDPs). In Ref. [8] the Bose-symmetric channel is considered as unilocal broadcasting operation and an SDP is derived for this problem.

Quantum discord, as an indispensable measure of quantum correlation beyond entanglement, is introduced in Refs. [9] and [10] independently. It is argued [11] that quantum discord is responsible for the quantum speed-up over classical algorithms. Quantum discord is a quite useful concept in many fields of quantum information processing [4, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26].

The local broadcasting paradigm can provide a natural operational interpretation to quantum discord. Remarkably, the minimum average loss of mutual information resulting from local operation $\Lambda_{A\to A_1\cdots A_n}$ on A for arbitrary quantum state ρ_{AB} approaches its quantum discord $D_A(\rho_{AB})$ as n goes to infinity [12, 27]. However, it remains open whether there is an analogous connection for the two-sided setting of redistributing correlations [26].

* wei.xie-4@student.uts.edu.au
† kun.fang-1@student.uts.edu.au
‡ xin.wang-8@student.uts.edu.au
§ runyao.duan@uts.edu.au
¶ This submission is based on arxiv: 1705.06071.

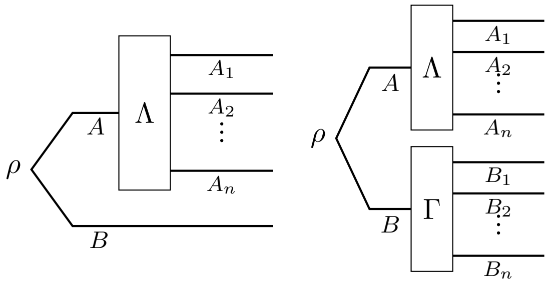


Figure 1: Unilocal (left) and bilocal (right) broadcasting of quantum correlations in initial state ρ_{AB}. The objective is for the quantum channels Λ, Γ to make the states on A_iB or A_iB_i as close to ρ_{AB} as possible measured in some way.

2 Overview of results

In this paper, we study the approximate broadcasting of quantum correlations in both asymptotic and non-asymptotic settings.

In the asymptotic regime, we rigorously prove the conjecture in Ref. [26] and provide an operational meaning of the two-sided quantum discord in terms of bilocal broadcasting of correlations.

Specifically, the two-sided quantum discord of bipartite state ρ_{AB} is given by

$$D_{AB}(\rho_{AB}) := \min_{\mathcal{E}_A, \mathcal{F}_B \in \mathrm{QC}} \left[I(A:B)_{\rho_{AB}} - I(A:B)_{(\mathcal{E}_A \otimes \mathcal{F}_B)\rho_{AB}} \right], \tag{1}$$

where QC denotes the set of all quantum-to-classical channels, $I(A:B)_{\rho_{AB}}$ denotes the quantum mutual information of state ρ_{AB}. Then we prove that for any bipartite state ρ_{AB},

$$D_{AB}(\rho_{AB}) = \lim_{n\to\infty} \min_{\substack{\Lambda_{A\to A_1\ldots A_n} \\ \Gamma_{B\to B_1\ldots B_n}}} \frac{1}{n} \sum_{j=1}^{n} \left[I(A:B)_{\rho_{AB}} - I(A_j:B_j)_{(\Lambda_j \otimes \Gamma_j)\rho_{AB}} \right], \tag{2}$$

where $\Lambda_{A\to A_1\ldots A_n}$ and $\Gamma_{B\to B_1\ldots B_n}$ be quantum channels and denote $\Lambda_j := \mathrm{Tr}_{\backslash A_j} \circ \Lambda$ and $\Gamma_j := \mathrm{Tr}_{\backslash B_j} \circ \Gamma$. This result shows that the asymptotic minimum average loss of correlation after optimal bilocal broadcasting is exactly the two-sided quantum discord of the initial state.

In the non-asymptotic regime, we give an SDP characterization of the optimal unilocal broadcasting fidelity and show that the universal quantum clone machine (UQCM) [28, 29] can also serve as the optimal universal unilocal broadcasting operation. In particular, it has the strongest power for universal unilocal broadcasting.

Specifically, given a bipartite state ρ_{AB}, the *optimal unilocal n-broadcasting fidelity* of ρ_{AB} on system A is defined as the following optimal average fidelity over all quantum channels $\Lambda_{A\to A_1\cdots A_n}$,

$$f_n(\rho_{AB}) := \sup_{\Lambda_{A\to A_1\cdots A_n}} \frac{1}{n} \sum_{j=1}^{n} F(\rho_{AB}, \mathrm{Tr}_{\backslash A_j B} \Lambda_{A\to A_1\ldots A_n}(\rho_{AB})). \tag{3}$$

We show that the quantity $f_n(\rho_{AB})$ can be characterized as SDP,

$$\begin{aligned} f_n(\rho_{AB}) = \max\ & \frac{1}{2} \mathrm{Tr}(X_{AB} + X_{AB}^\dagger) \\ \text{s.t.}\ & \begin{pmatrix} \rho_{AB} & X_{AB} \\ X_{AB}^\dagger & \mathrm{Tr}_{\backslash A_1 B}(J^{T_A} \rho_{AB}) \end{pmatrix} \geq 0, \\ & J_{AA_1\cdots A_n} \geq 0, \mathrm{Tr}_{\backslash A} J_{AA_1\cdots A_n} = \mathbb{1}_A, \\ & J_{AA_1\cdots A_n} = \frac{1}{n!} \sum_{\pi\in S_n} W_\pi J_{AA_1\cdots A_n} W_\pi^\dagger, \end{aligned} \tag{4}$$

where W_π is a unitary for each permutation π in symmetric group S_n, by the action $W_\pi |j_1, j_2, \ldots, j_n\rangle = |j_{\pi^{-1}(1)}, j_{\pi^{-1}(2)}, \ldots, j_{\pi^{-1}(n)}\rangle$ for any choice of $|j_1\rangle, |j_2\rangle, \ldots, |j_n\rangle$. A quantum channel $\Lambda_{A\to A_1\cdots A_n}$ is called a *symmetric broadcasting channel*, if $\Lambda(\rho) = W_\pi(\Lambda(\rho))W_\pi^\dagger$ for any input state ρ and permutation π.

The optimal unilocal 2-broadcasting fidelity of the maximally entangled state $|\Phi_d\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |jj\rangle$ is given by $f_2(\Phi_d) = \sqrt{\frac{d+1}{2d}}$. Interestingly, its optimal broadcasting channel is exactly the universal quantum clone machine, denoted as $\Upsilon^d_{A\to A_1A_2}$.

Moreover, the optimal unilocal 2-broadcasting fidelity for pure two-qubit states is analytically solved. For two-qubit pure state $\psi_\theta = |\psi_\theta\rangle\langle\psi_\theta|$ with $|\psi_\theta\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle$, $\theta \in (0, \pi/4]$, its optimal unilocal 2-broadcasting fidelity is given by

$$f_2(\psi_\theta) = \begin{cases} \cos^2\theta + (\sin^2\theta)/\sqrt{2}, & \theta \in (0, \arctan(2^{-1/4})] \\ (\frac{3}{2}(\cos^4\theta + \sin^4\theta))^{1/2}, & \theta \in (\arctan(2^{-1/4}), \pi/4] \end{cases}.$$

Furthermore, we introduce a worst-case quantifier for the performance of unilocal broadcasting of a symmetric channel. For any symmetric broadcasting channel $\Lambda_{A\to A_1\cdots A_n}$, we define its *unilocal broadcasting power* as

$$\mathcal{P}(\Lambda) := \inf_{\rho_{AB}\in\mathcal{S}(AB)} F(\rho_{AB}, \mathrm{Tr}_{\backslash A_1 B} \Lambda(\rho_{AB})). \tag{5}$$

The unilocal broadcasting power of a symmetric broadcasting channel gives a measure of the universal unilocal broadcasting ability for symmetric broadcasting channels. The universality means it is independent of the input state. The channel with a larger value of unilocal broadcasting power is more capable of unilocal broadcasting quantum states in a universal sense.

In particular, we prove that the optimal unilocal 2-broadcasting channel $\Upsilon^d_{A\to A_1A_2}$ for the maximally entangled state has the greatest power for unilocal 2-broadcasting, i.e.,

$$\max_{\Lambda_{A\to A_1A_2}} \mathcal{P}(\Lambda_{A\to A_1A_2}) = \mathcal{P}(\Upsilon^d_{A\to A_1A_2}), \tag{6}$$

where the maximum is taken over all symmetric broadcasting channels.

Finally, we discuss broadcasting of a set of quantum states. Assuming the given states ρ_i are on the system A, we study how to optimize the *n-broadcasting fidelity* $g_n(\eta)$ *of an ensemble* $\eta := \{p_i, \rho_i\}_{i=1}^m$, which is defined as

$$\begin{aligned} g_n(\eta) := \sup \ & \sum_{i=1}^m \sum_{j=1}^n \frac{1}{n} p_i F(\rho_i, \widehat{\rho}_{ij}) \\ \text{s.t. } & \widehat{\rho}_{ij} = \operatorname{Tr}_{\backslash A_j} \Lambda_{A\to A_1\cdots A_n}(\rho_i), \\ & \Lambda \text{ is a quantum channel.} \end{aligned} \tag{7}$$

Following the same line as the optimal unilocal broadcasting fidelity, we show that $g_n(\eta)$ can also be characterized as SDP,

$$\begin{aligned} g_n(\eta) := \max \ & \sum_{i=1}^m \frac{1}{2} p_i \operatorname{Tr}(X_i + X_i^\dagger) \\ \text{s.t. } & \begin{pmatrix} \rho_i & X_i \\ X_i^\dagger & \operatorname{Tr}_{\backslash A_1}(J_{AA_1\cdots A_n}\rho_i^T) \end{pmatrix} \geq 0, \forall i = 1, \cdots, m, \\ & J_{AA_1\cdots A_n} \geq 0, \operatorname{Tr}_{\backslash A} J_{AA_1\cdots A_n} = \mathbb{1}_A, \\ & J_{AA_1\cdots A_n} = \frac{1}{n!} \sum_{\pi\in S_n} W_\pi J_{AA_1\cdots A_n} W_\pi^\dagger. \end{aligned} \tag{8}$$

References

[1] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," *Nature*, vol. 299, no. 5886, pp. 802–803, 1982.

[2] D. Dieks, "Communication by epr devices," *Physics Letters A*, vol. 92, no. 6, pp. 271–272, 1982.

[3] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, "Noncommuting mixed states cannot be broadcast," *Physical Review Letters*, vol. 76, no. 15, p. 2818, 1996.

[4] M. Piani, P. Horodecki, and R. Horodecki, "No-local-broadcasting theorem for multipartite quantum correlations," *Physical Review Letters*, vol. 100, no. 9, p. 090502, 2008.

[5] S. Luo and W. Sun, "Decomposition of bipartite states with applications to quantum no-broadcasting theorems," *Physical Review A*, vol. 82, no. 1, p. 012338, 2010.

[6] S. Luo, "On quantum no-broadcasting," *Letters in Mathematical Physics*, vol. 92, no. 2, pp. 143–153, 2010.

[7] M. Piani, "Local broadcasting of quantum correlations," *arXiv preprint arXiv:1608.02650*, 2016.

[8] ——, "Hierarchy of efficiently computable and faithful lower bounds to quantum discord," *Physical Review Letters*, vol. 117, no. 8, p. 080401, 2016.

[9] H. Ollivier and W. H. Zurek, "Quantum discord: a measure of the quantumness of correlations," *Physical Review Letters*, vol. 88, no. 1, p. 017901, 2001.

[10] L. Henderson and V. Vedral, "Classical, quantum and total correlations," *Journal of Physics A: Mathematical and General*, vol. 34, no. 35, p. 6899, 2001.

[11] A. Datta, A. Shaji, and C. M. Caves, "Quantum discord and the power of one qubit," *Physical Review Letters*, vol. 100, no. 5, p. 050502, 2008.

[12] F. G. Brandão, M. Piani, and P. Horodecki, "Generic emergence of classical features in quantum darwinism," *Nature Communications*, vol. 6, 2015.

[13] E. Knill and R. Laflamme, "Power of one bit of quantum information," *Physical Review Letters*, vol. 81, no. 25, p. 5672, 1998.

[14] M. Piani, V. Narasimhachar, and J. Calsamiglia, "Quantumness of correlations, quantumness of ensembles and quantum data hiding," *New Journal of Physics*, vol. 16, no. 11, p. 113001, 2014.

[15] S. Boixo, L. Aolita, D. Cavalcanti, K. Modi, and A. Winter, "Quantum locking of classical correlations and quantum discord of classical-quantum states," *International Journal of Quantum Information*, vol. 9, no. 07n08, pp. 1643–1651, 2011.

[16] T. Chuan, J. Maillard, K. Modi, T. Paterek, M. Paternostro, and M. Piani, "Quantum discord bounds the amount of distributed entanglement," *Physical Review Letters*, vol. 109, no. 7, p. 070501, 2012.

[17] A. Streltsov, H. Kampermann, and D. Bruß, "Quantum cost for sending entanglement," *Physical Review Letters*, vol. 108, no. 25, p. 250501, 2012.

[18] I. Devetak and A. Winter, "Distilling common randomness from bipartite quantum states," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3183–3196, 2004.

[19] D. Cavalcanti, L. Aolita, S. Boixo, K. Modi, M. Piani, and A. Winter, "Operational interpretations of quantum discord," *Physical Review A*, vol. 83, no. 3, p. 032324, 2011.

[20] V. Madhok and A. Datta, "Interpreting quantum discord through quantum state merging," *Physical Review A*, vol. 83, no. 3, p. 032323, 2011.

[21] A. Streltsov, H. Kampermann, and D. Bruß, "Linking quantum discord to entanglement in a measurement," *Physical Review Letters*, vol. 106, no. 16, p. 160401, 2011.

[22] V. Madhok and A. Datta, "Quantum discord as a resource in quantum communication," *International Journal of Modern Physics B*, vol. 27, no. 01n03, p. 1345041, 2013.

[23] D. Girolami, A. M. Souza, V. Giovannetti, T. Tufarelli, J. G. Filgueiras, R. S. Sarthour, D. O. Soares-Pinto, I. S. Oliveira, and G. Adesso, "Quantum discord determines the interferometric power of quantum states," *Physical Review Letters*, vol. 112, no. 21, p. 210401, 2014.

[24] S. Pirandola, "Quantum discord as a resource for quantum cryptography," *arXiv preprint arXiv:1309.2446*, 2013.

[25] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, "The classical-quantum boundary for correlations: discord and related measures," *Reviews of Modern Physics*, vol. 84, no. 4, p. 1655, 2012.

[26] G. Adesso, T. R. Bromley, and M. Cianciaruso, "Measures and applications of quantum correlations," *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 47, p. 473001, 2016.

[27] A. Streltsov and W. H. Zurek, "Quantum discord cannot be shared," *Physical Review Letters*, vol. 111, no. 4, p. 040401, 2013.

[28] V. Bužek and M. Hillery, "Quantum copying: Beyond the no-cloning theorem," *Physical Review A*, vol. 54, no. 3, p. 1844, 1996.

[29] ——, "Universal optimal cloning of arbitrary quantum states: from qubits to quantum registers," *Physical review letters*, vol. 81, no. 22, p. 5003, 1998.

Universal extensions of restricted classes of quantum operations

Michał Oszmaniec[1,2] and Zoltán Zimborás[3,4]

[1]ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain
[2]National Quantum Information Centre of Gdańsk, 81-824 Sopot, Poland
[3]Wigner Research Centre for Physics, Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary
[4]Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany

INTRODUCTION

In many applications of quantum mechanics it is important to have full control over a quantum system used to perform a desired task or a quantum protocol. This amounts to being able to implement arbitrary unitary operation on the system in question. Perhaps the most well-known example is the circuit model of quantum computing, where the ability to implement arbitrary unitary gates on a system of many distinguishable particles (say, qubits) is a necessary ingredient for performing universal quantum computation [1, 2]. From the experimental perspective, it is typically very easy to implement single-qubit gates. This collection of gates, however, does not lead to universal quantum computation and to this aim has to be supplemented by an entangling gate [3]. Similar situations appears in other physical contexts. Typically, the set of *easily accessible* unitary gates acting on a given quantum system, does not ensure full controllability.

This work studies the *extension problem* for gate-sets appearing naturally in systems consisting of non-distinguishable particles: passive linear optics for (A) bosonic and (B) fermionic systems with fixed number of particles, as well as (C) active linear optics acting on fermionic system with fixed number of modes subject to the parity super-selection rule [4]. Specifically, for the aforementioned scenarios, we study what unitary transformations can be implemented if the restricted class of gates K is supplemented by additional unitary transformations - see Fig. 1. We investigate two variants of this problem:

(i) the gate-set K is supplemented with unitaries of the form $\exp(-\mathrm{i}tX)$ generated by the Hamiltonian X;

(ii) the gate-set K is supplemented by a *single* unitary transformation V.

We denote by $\langle K, X\rangle$ the set of unitaries that can be generated form the restricted gate-set K and unitaries of the form $\exp(-\mathrm{i}tX)$, where t is an arbitrary real number. Likewise, slightly abusing the notation, we denote by $\langle K, V\rangle$ the set of unitaries that can be generated by elements form K and an extra gate V. The merit of this work is to characterize sets $\langle K, X\rangle$ and $\langle K, V\rangle$ for different linear optical groups K (acting on the appropriate Hilbert spaces $\mathcal{H}$). If the resulting gate-set $\langle K, X\rangle$ (or $\langle K, V\rangle$) form the full unitary group $\mathrm{U}(\mathcal{H})$, we say that the Hamiltonian X (or the gate V) *promotes* the restricted collection of gates K to universality in $\mathcal{H}$.

CONTEXT AND MOTIVATION

Linear optical transformations are relevant in many contexts. Passive bosonic linear optics describes single-particle evolutions of a system of N identical bosons in d modes. Such transformations are natural for quantum optics, when quantum states of light pass through an optical network formed from beam-splitters and phase shifters [5]. Linear optics underpins the KLM scheme of quantum computing with photons [6] and the boson sampling strategy for demonstrating quantum supremacy with linear optical networks [7]. Moreover, this class of transformations is used to manipulate cold bosonic particles in optical traps [8, 9]. Similarly, passive fermionic linear optics describes single-particle evolutions of a system of N identical fermions in d modes [10, 11], which can be realized in integer quantum hall effect systems exhibiting edge channels [12]. Passive fermionic linear optics together with particle-number measurements yields a classically simulable model of quantum computation [11]. Finally, active fermionic linear optics describes free-fermion transformations that are not necessary particle preserving. These fermionic transformations are the basic ingredient of a classically simulable model of quantum computation [4, 10], which have been widely studied in the context of Matchgates [13–15].

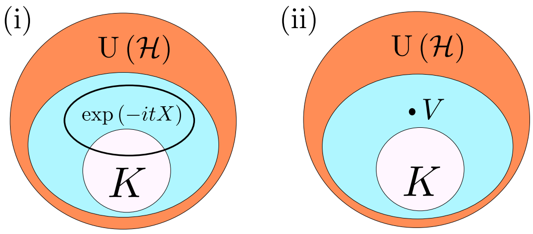


FIG. 1. A schematic presentation of the problems studied. (i) Given a family of gates K (white) and a one-parameter family of unitaries $\exp(-\mathrm{i}tX)$ (black loop), what class of gates (cyan) can be generated in the full unitary group $\mathrm{U}(\mathcal{H})$ (orange)? (ii) Given a family of gates K (white) and a single gate V (black dot), what class of gates (cyan) can be generated in the full unitary group $\mathrm{U}(\mathcal{H})$ (orange)?

SETTING

We denote the Hilbert space of N bosons in d modes by $\mathcal{H}_b$. We have $\mathcal{H}_b = \mathrm{Sym}^N(\mathbb{C}^d)$, i.e., in this case the bosonic Hilbert space can be identified with the totally symmetric subspace of the Hilbert space of N distinguishable qudits, $(\mathbb{C}^d)^{\otimes N}$. In this language the group of passive linear optical bosonic transformations, denoted by LO_b, can defined as the group of unitaries of the form $U^{\otimes N}$, with $U \in \mathrm{U}(d)$, restricted to the bosonic subspace $\mathcal{H}_b$.

The Hilbert space of N (spinless) fermionic particles in d modes ($d \geq N$) is $\mathcal{H}_f = \wedge^N(\mathbb{C}^d)$, i.e., the totally antisymmetric subspace of $(\mathbb{C}^d)^{\otimes N}$. Similarly to the bosonic case, the group of passive fermionic linear optics LO_f is defined as the group of unitaries $U^{\otimes N}$, with $U \in \mathrm{U}(d)$, restricted to the fermionic subspace $\mathcal{H}_f$. And in an analogous way, one can define the group of active parity-preserving fermionic linear optics (FLO) that acts on the positive parity fermionic Fock space, $\mathcal{H}^{+}_{\mathrm{Fock}}$.

RESULTS - THE MAIN IDEAS

In this work, we completely solve problems (i) and (ii) for the scenarios A-C. We characterize the unitary transformations that are implementable (maybe approximately) by linear optical gates supplemented with any additional Hamiltonian or a gate. Our characterization is given in terms of *explicit* algebraic conditions on the Hamiltonian X or the gate V that can be can be tested operationally. The resulting behavior is surprisingly rich and structurally depends on the number of modes and the number of particles. In particular, contrary to what intuition might suggest, it is not true that every nontrivial extra gate or Hamiltonian provides universality in scenarios A-C. Solution of problems (i) and (ii) gives the clear understanding of what resources are necessary to have full physical controllability in the contexts listed above. Moreover, our results can be viewed as a step towards a solution of the general problem of classification of invertible quantum circuits posed recently by Aaronson and co-workers [16, 17].

To give an idea about the structure of our findings (see [18] for the full version of the paper), we present the results concerning the gate addition for the simplest case of passive linear optics and for passive fermionic linear optics.

Theorem 1 (Extensions of Passive Bosonic Linear Optics with an additional gate)**.** *Let $V \notin \mathrm{LO}_b$ be a gate acting on the Hilbert space $\mathcal{H}_b$ of $N > 1$ bosons in d modes. Let $\langle \mathrm{LO}_b, V\rangle$ be the group of transformations generated by passive bosonic linear optics and V in $\mathcal{H}_b$. For $d = 2$ we define:*

$$\mathbb{L}_b = |\Psi_b\rangle\langle\Psi_b|,\ |\Psi_b\rangle = \sum_{k=0}^{N}(-1)^k|D_k\rangle|D_{N-k}\rangle \in \mathcal{H}_b\otimes\mathcal{H}_b\ , \tag{1}$$

where $|D_k\rangle$ denote the two-mode Dicke states with k-particles being in the first mode. We have the following possibilities:

(a) If $d > 2$, then $\langle \mathrm{LO}_b, V\rangle = \mathrm{U}\left(\mathcal{H}_b\right)$.

(b) If $d = 2$, $N \neq 6$ and $[V\otimes V, \mathbb{L}_b] = 0$, then

$$\langle \mathrm{LO}_b, V\rangle = G_b = \{\, U \in \mathrm{U}\left(\mathcal{H}_b\right) |\ [U\otimes U, \mathbb{L}_b] = 0\,\}.$$

(c) If $d = 2$ and $[V\otimes V, \mathbb{L}_b] \neq 0$, then $\langle \mathrm{LO}_b, V\rangle = \mathrm{U}\left(\mathcal{H}_b\right)$.

In the above theorem we have situations with $N = 1$ particles as for them $\mathrm{LO}_b = \mathrm{U}(\mathcal{H}_b)$. We see that for $d \neq 2$ any additional gate promotes LO_b to universality in the bosonic space $\mathcal{H}_b$. For $d = 2$ the resulting gate-set $\langle \mathrm{LO}_b, V\rangle$ depends only on the commutator $[V\otimes V, \mathbb{L}_b]$. If it is nonzero, then V again extends LO_b to universality; while if it vanishes (and $N \neq 6$) V extends LO_b to the "middle group" G_b. Up to a global phase the group G_b consists of unitaries that preserve the bilinear form defined by $B(|\psi\rangle, |\phi\rangle) = \langle\Psi_b|(|\psi\rangle\otimes|\phi\rangle)$. Here, by preservation we mean that $B_b(U|\psi\rangle, U|\phi\rangle) = B_b(|\psi\rangle, |\phi\rangle)$, for all vectors $|\phi\rangle$,$|\psi\rangle$. If the number of particles N is even then $|\Psi_b\rangle$ is a symmetric tensor and defines the real inner product. In this case we have $G_b = \langle \mathbb{T}(\mathcal{H}_b), \mathrm{SO}(\mathcal{H}_b)\rangle$, where $SO(\mathcal{H}_b)$ is the special orthogonal group on $\mathcal{H}_b$. When the number of particles N is odd, the vector $|\Psi_b\rangle$ is antisymmetric and defines the symplectic structure (i.e. non-degenerate and antisymmetric form) on $\mathcal{H}_b$. In this case we have $G_b = \langle \mathbb{T}(\mathcal{H}_b), \mathrm{USp}(\mathcal{H}_b)\rangle$, where $\mathrm{USp}(\mathcal{H}_b)$ is the unitary symplectic group. The two differ considerably as $\mathrm{USp}(\mathcal{H})$ acts transitively on the set of pure states on $\mathcal{H}$ [19–21]. On the other hand, $\mathrm{SO}(\mathcal{H})$ acts transitively only on "real" pure states. Thus, for odd number of particles, adding any additional gate gives either the full unitary controllability or the pure-state controllability. In the case of $d = 2$ modes and $N = 6$ particles if $[V\otimes V, \mathbb{L}_b] = 0$ the situation complicates due to the presence of additional group (related to the *exotic* group G_2) in between LO_b and G_b. We leave the description of this exceptional case as an interesting open problem.

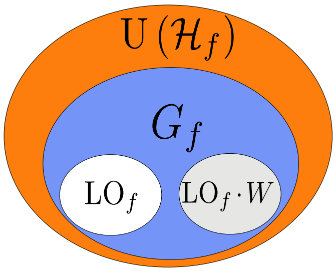


FIG. 2. A pictorial presentation of most complicated chain of group inclusions that can be realized for the problem (ii) in the considered scenarios. For passive fermionic linear optics LO_f in the half-filling case ($d = 2N$) we have $\mathrm{LO}_f \subset \mathrm{LO}_f \cup \mathrm{LO}_f \cdot W \subset G_f \subset \mathrm{U}\left(\mathcal{H}_f\right)$, where $W = \prod_{i=1}^{d}(f_i + f_i^{\dagger})$, and the "middle" group G_f is defined by the condition $[U\otimes U, \mathbb{L}_f] = 0$.

The analogous result of passive fermionic linear optics is the following

Theorem 2 (Extensions of Passive Fermionic Linear Optics with an additional gate)**.** *Let $V \notin \mathrm{LO}_f$ be a gate acting on Hilbert space of N fermions in d modes $\mathcal{H}_f$, where $N \notin \{0, 1, d-1, d\}$. Let $\langle \mathrm{LO}_f, V\rangle$ be the group of transformations generated by passive fermionic linear optics and V in $\mathcal{H}_f$. For $d = 2N$ (half-filling) we define:*

$$\mathbb{L}_f = |\Psi_f\rangle\langle\Psi_f|,\ \text{with}\ |\Psi_f\rangle = |1\rangle\wedge|2\rangle\wedge\ldots\wedge|2N\rangle \in \mathcal{H}_f\otimes\mathcal{H}_f \tag{2}$$

where $\wedge$ denotes the standard wedge product. We have the following possibilities:

(a) If $d \neq 2N$, then $\langle \mathrm{LO}_f, V\rangle = \mathrm{U}\left(\mathcal{H}_f\right)$.

(b) If $d = 2N$ and $V = Wk$, for $k \in \mathrm{LO}_f$ and $W = \prod_{i=1}^{d}(f_i + f_i^{\dagger})$, then $\langle \mathrm{LO}_f, V\rangle = \mathrm{LO}_f \cup \mathrm{LO}_f \cdot W$.

(c) If $d = 2N$, $V \neq gW$, for $g \in \mathrm{LO}_f$, and $[V\otimes V, \mathbb{L}_f] = 0$, then

$$\langle \mathrm{LO}_f, V\rangle = G_f = \{\, U \in \mathrm{U}\left(\mathcal{H}_f\right) |\ [V\otimes V, \mathbb{L}_f] = 0\,\}.$$

(d) If $d = 2N$ and $[V\otimes V, \mathbb{L}_f] \neq 0$, then $\langle \mathrm{LO}_f, V\rangle = \mathrm{U}\left(\mathcal{H}_f\right)$.

The structure of the above result is similar to the case of passive bosonic linear optics. In the formulation of the theorem we have excluded the non interesting cases $N \in \{0, 1, d-$

$1, d\}$ since for them LO_f equals the full unitary group on the respective Hilbert space. When $d \neq 2N$ every gate promotes LO_f to universality. However, in the physically relevant case of half-filling [22], a more interesting "onion" structure appears. In the case (b) addition of an extra gate of the form kW, where $k \in \mathrm{LO}_f$ and gate W (describing particle-hole transformation in $\mathcal{H}_f$) gives the gate-set $\mathrm{LO}_f \cup \mathrm{LO}_f \cdot W$ (it is crucial here that W commutes with $\mathbb{L}_f$ and that conjugation by W leaves LO_f invariant). The further possibilities are described, similarly to the bosonic case, by the commutation properties of $V \otimes V$ with LO_f. If d is not divisible by four we have $G_f = \langle \mathbb{T}(\mathcal{H}_f), \mathrm{SO}(\mathcal{H}_f) \rangle$. On the other hand, for d divisible by four, $G_f = \langle \mathbb{T}(\mathcal{H}_f), \mathrm{USp}(\mathcal{H}_f) \rangle$. The corresponding bilinear forms are defined by inner products with $|\Psi_f\rangle$.

EXAMPLES AND APPLICATIONS

We can apply our general results to many concrete physical examples. let us present first an exemplary application of our findings. In the reference [23], the authors were interested in extending bosonic linear optics to universality in $\mathcal{H}_b$ by adding an additional gate. This problem was motivated by the need to construct physically-accessible universal gate-set in $\mathcal{H}_b$, which can be used to generate, via construction based on random circuits [24], approximate bosonic t-designs. The example below proves that a singe gate based on the cross-Kerr nonlinearity suffices to promote bosonic linear optics to universality in $\mathcal{H}_b$. It should be mentioned that Kerr-like transformation have been previously used to obtain universal quantum computation in continuous-variable systems [25].

Example 1. *Consider a bosonic system with $d = 2$ modes and $N > 1$ particles, and a gate generated by the cross-Kerr interaction (acting on $\mathcal{H}_b$ for a fixed time t),*

$$V_t = \exp\left(-it\hat{n}_a\hat{n}_b\right) , \tag{3}$$

where $\hat{n}_{a,b}$ are the occupation number operators corresponding to modes a and b. Let $\langle \mathrm{LO}_b, V_t \rangle$ be the group of transformations generated by passive bosonic linear optics and V_t. Then, $\langle \mathrm{LO}_b, V_t \rangle = \mathrm{U}\left(\mathcal{H}_b\right)$ if and only if

$$\mathrm{e}^{2it[l(N-l)-k(N-k)]} \neq 1 \, , \tag{4}$$

for at least one pair (k, l) , where $k, l = 0, \ldots, N$. In particular, the gate $V_{\frac{\pi}{3}}$ promotes passive bosonic linear optics to universality in $\mathcal{H}_b$ for $d = 2$ modes.

Using the general results, we can also easily see that there might exist physical Hamiltonians that add different controllability properties to LO_b, depending on the number of particles N. The following example shows a particular case.

Example 2. *Consider the Hamiltonian $X_3 = \hat{n}_a^3 - \hat{n}_b^3$ acting on $\mathcal{H}_b$ for $d = 2$ modes and $N \neq 6$ particles. Deepening on the value N we get different types of gate-sets after supplementing passive bosonic optics with X_3*

(a) For even N: $\langle \mathrm{LO}_b, X_3 \rangle = \langle \mathbb{T}(\mathcal{H}_b), \mathrm{SO}(\mathcal{H}_b) \rangle$;

(b) For odd N: $\langle \mathrm{LO}_b, X_3 \rangle = \langle \mathbb{T}(\mathcal{H}_b), \mathrm{USp}(\mathcal{H}_b) \rangle$.

In, particular for odd N we have full controllability on the set of pure states on $\mathcal{H}_b$, whereas for even N this is not the case.

We can also give analogous results for passive and active fermionic linear optics, here we provide two:

Example 3. *For any non-quadratic Hamiltonian M containing only two-mode terms, the generated group $\langle \mathrm{LO}_f, M \rangle$ is the entire unitary group $\mathrm{U}\left(\mathcal{H}_f\right)$.*

Hamiltonians that are not composed of two-mode terms are also often studied. One typical family of these are the so-called correlated hopping Hamiltonians, where the hopping-term between two sites is multiplied with number operators belonging to other sites. For such Hamiltonians universality is not guaranteed:

Example 4. *Consider the correlated hopping Hamiltonian*

$$Y = \sum_{j=1}^{d/2-1} (\hat{n}_{2j} - \hat{n}_{2j+2})^2 (f_{2j-1}^\dagger f_{2j+1} + f_{2j+1}^\dagger f_{2j-1}) + (\hat{n}_{2j-1} - \hat{n}_{2j+1})^2 (f_{2j}^\dagger f_{2j+2} + f_{2j+2}^\dagger f_{2j}). \tag{5}$$

acting on $\mathcal{H}_f$ for the case of half filling ($d = 2N$). Then, we have the following situations

(a) for even N: $\langle \mathrm{LO}_f, Y \rangle = \langle \mathbb{T}(\mathcal{H}_f), \mathrm{SO}(\mathcal{H}_f) \rangle$;

(b) for odd N: $\langle \mathrm{LO}_f, Y \rangle = \langle \mathbb{T}(\mathcal{H}_f), \mathrm{USp}(\mathcal{H}_f) \rangle$.

For odd N the Hamiltonian Y together with LO_f ensures full controllability on the set of pure states on $\mathcal{H}_f$. However, for even N this is not the case. The above statements are even true for each term appearing in sum Eq. (5). The correlated hopping Hamiltonian Y often appears (in a relabeled form) in the literature on extended Hubbard models [26].

DISCUSSION AND OUTLOOK

In this extended abstract, we presented a comprehensive treatment of the extension problems for various classes of linear optical gates for bosons and fermions (see [18] for the full version of the paper, containing the discussion of the active linear optics for fermions and the proof of our results . The resulting behavior is surprisingly rich and critically depends on the number of modes and number of particles present in the system. However, there is a number of interesting problems we did not addressed here. First, it would be interesting to analyze which extra gates or Hamiltonians allow for the most efficient control [27] or the efficient approximation of gates from the appropriate unitary group [28]. Another important problem concerns the robustness of the extra gate or Hamiltonian to the noise that inevitably affects any quantum system. In future works we also plan to use our results to study (computational) universality of classically simulable models of computation supported on fermionic systems [10] and Machgates [13–15].

[1] A. Y. Kitaev, A. Shen, and M. N. Vyalyi, *Classical and quantum computation*, Vol. 47 (American Mathematical Society Providence, 2002).

[2] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge university press, 2010).

[3] J.-L. Brylinski and R. Brylinski, Mathematics of Quantum Computation **79** (2002).

[4] S. B. Bravyi and A. Y. Kitaev, Annals of Physics **298**, 210 (2002).

[5] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, *et al.*, Science **349**, 711 (2015).

[6] E. Knill, R. Laflamme, and G. J. Milburn, Nature **409**, 46 (2001).

[7] S. Aaronson and A. Arkhipov, in *Proceedings of the forty-third annual ACM symposium on Theory of computing* (ACM, 2011) pp. 333–342.

[8] G. Milburn, J. Corney, E. M. Wright, and D. Walls, Physical Review A **55**, 4318 (1997).

[9] M. Albiez, R. Gati, J. Fölling, S. Hunsmann, M. Cristiani, and M. K. Oberthaler, Phys. Rev. Lett. **95**, 010402 (2005).

[10] B. M. Terhal and D. P. DiVincenzo, Physical Review A **65**, 032325 (2002).

[11] D. P. DiVincenzo and B. M. Terhal, Foundations of Physics **35**, 1967 (2005).

[12] E. Bocquillon, V. Freulon, F. D. Parmentier, J.-M. Berroir, B. Plaçais, C. Wahl, J. Rech, T. Jonckheere, T. Martin, C. Grenier, *et al.*, Annalen der Physik **526**, 1 (2014).

[13] L. G. Valiant, SIAM Journal on Computing **31**, 1229 (2002).

[14] E. Knill, arXiv preprint quant-ph/0108033 (2001).

[15] R. Jozsa and A. Miyake, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 464 (The Royal Society, 2008) pp. 3089–3106.

[16] A. Bouland and S. Aaronson, Physical Review A **89**, 062316 (2014).

[17] S. Aaronson, D. Grier, and L. Schaeffer, arXiv preprint arXiv:1504.05155 (2015).

[18] M. Oszmaniec and Z. Zimborás, arXiv preprint quant-ph/1705.11188 (2017).

[19] Z. Zimborás, R. Zeier, M. Keyl, and T. Schulte-Herbrüggen, EPJ Quantum Technology **1**, 11 (2014).

[20] S. G. Schirmer, A. I. Solomon, and J. V. Leahy, J. Phys. A **35**, 4125 (2002).

[21] F. Albertini and D. D'Alessandro, IEEE Trans. Aut. Cont. **48**, 1399 (2003).

[22] M. Greiter, X.-G. Wen, and F. Wilczek, Phys. Rev. Lett. **66**, 3205 (1991).

[23] M. Oszmaniec, R. Augusiak, C. Gogolin, J. Kołodyński, A. Acín, and M. Lewenstein, Phys. Rev. X **6**, 041044 (2016).

[24] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, Phys. Rev. Lett. **116**, 170502 (2016).

[25] S. Lloyd and S. L. Braunstein, Phys. Rev. Lett. **82**, 1784 (1999).

[26] F. Dolcini and A. Montorsi, Phys. Rev. B **88**, 115115 (2013).

[27] M. A. Nielsen, M. R. Dowling, M. Gu, and A. C. Doherty, Science **311**, 1133 (2006).

[28] A. W. Harrow, B. Recht, and I. L. Chuang, J. Math. Phys. **43**, 4445 (2002).

Optimal quantum error correcting codes from absolutely maximally entangled states

Zahra Raissi[1 2 *] Christian Gogolin[1 †] Arnau Riera[1 ‡] Antonio Acín[1 3 §]

[1] *ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, Castelldefels (Barcelona), 08860, Spain*

[2] *Department of physics, Sharif University of Technology, Tehran, P.O. Box 111555-9161, Iran*

[3] *ICREA-Institució Catalana de Recerca i Estudis Avançats, Lluis Companys 23, Barcelona, 08010, Spain*

Abstract. Absolutely maximally entangled (AME) states are pure multi-partite generalizations of the bipartite maximally entangled states with the property that all reduced states of at most half the system size are in the maximally mixed state. AME states are of interest for multipartite teleportation and quantum secret sharing and have recently found new applications in the context of high-energy physics in toy models realizing the AdS/CFT-correspondence. We work out in detail the connection between AME states of minimal support and classical maximum distance separable (MDS) error correcting codes and, in particular, provide explicit closed form expressions for AME states of n parties with local dimension q a power of a prime for all $q \geq n1$. Building on this, we construct a generalization of the Bell-basis consisting of AME states and develop a stabilizer formalism for AME states. For every $q \geq n1$ prime we show how to construct QECCs that encode a logical qudit into a subspace spanned by AME states. Under a conjecture for which we provide numerical evidence, this construction produces a family of quantum error correcting codes $[\![n, 1, n/2]\!]_q$ for n even, with the highest code distance allowed by the quantum Singleton bound. We show that our conjecture is equivalent to the existence of an operator whose support cannot be decreased by multiplying it with stabilizer products and explicitly construct the codes up to $n = 8$.

Keywords: absolutely maximally entangled states, quantum error correcting codes, MDS codes, stabilizer, singleton bound

The maximally entangled states of two qubits, the so-called EPR states, are pure states of 2-qubits having reduced density matrices on each half of the system that are maximally mixed. A very intriguing question is whether systems made out of more than two parties can exhibit this property that all reduced states of at most half of the system size are maximally mixed. Such states are called Absolutely Maximally Entangled (AME) states and are pure multi-partite generalizations of the bipartite maximally entangled states.

AME states are known to play an important role in quantum information processing when dealing with many parties. They are useful for multipartite teleportation and in quantum secret sharing [1]. AME states have also deep connections with apparently unrelated areas of mathematics such as combinatorial designs and structures [2], classical error correcting codes [3], and quantum error correcting codes (QECC) [4]. Recently, they have gained new relevance as building blocks for holographic theories and in high energy physics. There they allow for the construction of tensor network states that realize discrete instances of the AdS/CFT correspondence and holography [5, 6, 7, 8].

At the same time it is still largely unknown for which n and q AME states exist and how they can be constructed. In the case of qubits for instance, it has been proven analytically that there are no AME states for $n = 4$ and $n \geq 7$. The non-existence in the cases $n = 4$ and $n \geq 8$ was proven by finding a contradiction in a linear program [9, 4]. Qubit AME states for $n = 2, 3$ were long known, a state for $n = 5$ was found in [10] and more recently such for $n = 5, 6$ were found numerically in [11, 12, 13]. The existence of such states was previously known in the context of quantum error correction [14]. Only very recently it was shown that there cannot be a qubit AME state for the case $n = 7$ [15].

We work out in detail the connection between AME states of minimal support and classical maximum distance separable (MDS) error correcting codes and, in particular, provide explicit closed form expressions for AME(n, q) states of n parties with local dimension q a power of a prime for all $n \leq q + 1$. Further, from a single AME state, we show how to produce an orthonormal basis of AME states. Based on our construction of minimal support AME states, we derive the generators of the stabilizer group of our AME states

Further we conjecture the existence of an infinite family of QECC whose code spaces are spanned by AME states and explicitly construct them for several cases. More precisely, we present a construction that we believe yields for every $q \geq n - 1$ a power of a prime a QECCs that encodes a logical q-level qudit into a subspace spanned by AME states of n such qudits. Under a conjecture for which we provide numerical evidence, this construction produces a family of quantum error correcting codes $[\![n, 1, n/2]\!]_q$ for n even achieve the highest code distance allowed the quantum Singleton bound. For $n \mod 4 = 3$ these codes can correct arbitrary errors on the same number of subsystems as a QMDS code with the same n and k. Our method provides QECC for smaller local dimension q than previously known QECC with otherwise identical parameters and

[*] zahra.raissi@icfo.eu

[†] christian.gogolin@icfo.eu

[‡] arnau.riera@icfo.eu

[§] antonio.acin@icfo.eu

we explicitly construct them for $n = 4, 6, 8$. Also, our proposal has a very clear physical motivation and nicely complements other constructions of non-binary QECC (see for example [16, 17] for an overview and [18] for tables of known codes with $q = 2$).

This extended abstract is based on [19].

References

[1] W. Helwig, W. Cui, J. I. Latorre, A. Riera, and H.-K. Lo, "Absolute maximal entanglement and quantum secret sharing," *Phys. Rev. A* **86** (Nov, 2012) 052335.

[2] D. Goyeneche, D. Alsina, J. I. Latorre, A. Riera, and K. Życzkowski, "Absolutely maximally entangled states, combinatorial designs, and multiunitary matrices," *Phys. Rev. A* **92** (Sep, 2015) 032316.

[3] W. Helwig and W. Cui, "Absolutely maximally entangled states: Existence and applications,".

[4] A. J. Scott, "Multipartite entanglement, quantum-error-correcting codes, and entangling power of quantum evolutions," *Phys. Rev. A* **69** (May, 2004) 052330.

[5] J. I. Latorre and G. Sierra, "Holographic codes,".

[6] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, "Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence," *Journal of High Energy Physics* **2015** no. 6, (2015) 149.

[7] A. Almheiri, X. Dong, and D. Harlow, "Bulk locality and quantum error correction in ads/cft," *Journal of High Energy Physics* **2015** no. 4, (2015) 163.

[8] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang, "Holographic duality from random tensor networks," *Journal of High Energy Physics* **2016** no. 11, (2016) 9.

[9] E. M. Rains, "Quantum shadow enumerators," *IEEE Transactions on Information Theory* **45** no. 6418557, (1999) 2361 – 2366.

[10] R. Laflamme, C. Miquel, J. P. Paz, and W. Zurek, "Perfect quantum error correcting code," *Physical Review Lett.* **77** no. 198, (1996) .

[11] A. Borras, A. R. Plastino, J. Batle, C. Zander, M. Casas, and A. Plastino, "Multiqubit systems: highly entangled states and entanglement distribution," *J. Phys. A: Math. Theor.* **40** (October, 2008) 13407.

[12] P. Facchi, G. Florio, G. Parisi, and S. Pascazio, "Maximally multipartite entangled states," *Phys. Rev. A* **77** (Jun, 2008) 060304.

[13] P. Facchi, G. Florio, U. Marzolino, G. Parisi, and S. Pascazio, "Classical statistical mechanics approach to multipartite entanglement," *Journal of Physics A: Mathematical and Theoretical* **43** no. 22, (2010) 225303.

[14] E. M. Rains, "Nonbinary quantum codes," *IEEE Transactions on Information Theory* **45** no. 6341997, (1999) 1827 – 1832.

[15] F. Huber, O. Gühne, and J. Siewert, "Absolutely maximally entangled states of seven qubits do not exist," *Phys. Rev. Lett* **118** (May, 2017) 200502.

[16] A. Ketkar, A. Klappenecker, S. Kumar, and P. K. Sarvepalli, "Nonbinary stabilizer codes over finite fields," *IEEE Transactions on Information Theory* **52** no. 11, (Nov, 2006) 4892–4914.

[17] M. Grassl and M. Rötteler, "Quantum mds codes over small fields," in *2015 IEEE International Symposium on Information Theory (ISIT)*, pp. 1104–1108. June, 2015.

[18] M. Grassl, "Bounds on the minimum distance of linear codes and quantum codes." `http://www.codetables.de`. Accessed on 2016-08-10.

[19] Z. Raissi, C. Gogolin, A. Riera, and A. Acín, "Constructing optimal quantum error correcting codes from absolute maximally entangled states," `arXiv:1701.03359`.

Efficient unitary designs with nearly time-independent Hamiltonian dynamics

Yoshifumi Nakata[1 2 *] Christoph Hirche[2 †] Masato Koashi[1 ‡] Andreas Winter[2 3 §]

[1] *Photon Science Center, Graduate School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo, Japan.*
[2] *Departament de Física: Grup d'Informació Quàntica, Universitat Autònoma de Barcelona, Bellaterra, Spain.*
[3] *ICREA–Institució Catalana de Recerca i Estudis Avançats, Psg. Lluis Companys, 23, Barcelona, Spain.*

Abstract. We provide new constructions of unitary designs on one qudit and those on N qubits, based on the schemes of repeating random unitaires diagonal in mutually unbiased bases. We first show that, if a pair of the bases satisfies a certain condition, the process on one qudit approximately forms a unitary t-design after $O(t)$ repetitions. We then construct quantum circuits on N qubits that achieve unitary t-designs for $t = o(N^{1/2})$ using $O(tN^2)$ gates, improving the best known result using $O(t^{10}N^2)$ gates in terms of t. Furthermore, we propose a *design Hamiltonian*, a random Hamiltonian of which dynamics always forms a unitary design after a threshold time, and present one with two-local spin-glass-type interactions. After changing the interactions only $O(t)$ times, the dynamics forms unitary t-designs. We also generalise the fast scrambling conjecture in terms of the design Hamiltonian.

Keywords: unitary design, random circuits, scrambling

1 Introduction and main results

Random processes play key roles in quantum information processing, as one of the important primitives in quantum Shannon theory and as a useful resource of quantum advantages, and also in fundamental physics of complex quantum systems, leading to new developments in quantum thermodynamics, the black hole information science and strongly correlated many-body physics. Traditionally, quantum random processes are represented by *Haar random unitaries.* However, they cannot be efficiently implemented by quantum circuits, which also implies that Haar random unitaries rarely appear in natural many-body systems. This fact has led to the study of finite-degree approximations of them, called *unitary t-designs* [1, 2]. Unitary 2-designs were intensely studied so far, but little is known about efficient implementations of t-designs for general t [3, 4], amongst which the best known result is to use *local random quantum circuits* [4]. They approximately form t-designs using $O(t^{10}N^2)$ gates,. This result can be interpreted in terms of Hamiltonian dynamics that the interactions should be changed uniformly at random $O(t^{10}N)$ times before unitary t-designs are generated, which is highly time-dependent and may not be physically feasible especially in a large system.

In this work based on Ref. [5], we show the following:

1. In a one-qudit system, $O(t)$ repetitions of random unitaries diagonal in two bases achieve unitary t-designs if a pair of the two bases satisfies the *Fourier-type* condition (Theorem 1).
2. In an N-qubit system, we provide a quantum circuit with $O(tN^2)$ gates forming a unitary t-design for $t = o(N^{1/2})$ (Theorem 2, see Table. 1 as well).
3. We introduce a *design Hamiltonian* and present one with spin-glass interactions (Corollary 1). The interactions should be changed only $O(t)$ times to achieve unitary t-designs.

Our results contribute both to quantum information science and to fundamental physics. In quantum information science, unitary t-designs for $t \geq 4$ are useful not only because they have direct applications [7, 8, 9, 10] but also because they provide better performance in a number of applications of 2-designs. There, it is often shown that one unitary chosen from a 2-design achieves the goal with high probability. If a higher-design is used instead, the probability gets much higher due to the large deviation bounds of designs [11]. Hence, our result improves *any* applications of unitary t-designs for $t \geq 2$.

On the other hand, our result about the design Hamiltonian contributes to developing a unified framework of studying random unitaries in isolated quantum systems, where Hamiltonians should be time-independent and fixed. The idea of design Hamiltonian opens for the first time the possibility to address the *fast scrambling conjecture* (even the generalised version) in terms of time-independent Hamiltonians. Since scrambling and its higher order generalisation are the key to understanding fundamental physics in complex quantum many-body systems, we believe that the design Hamiltonian provides the solid basis of these studies.

2 Brief description of theorems

An *ϵ-approximate unitary t-design* is a random unitary U such that $\|\mathcal{G}_U^{(t)} - \mathcal{G}_{U^H}^{(t)}\|_\diamond \leq \epsilon$, where $\mathcal{G}_U^{(t)} := \mathbb{E}_U[U^{\otimes t} X U^{\dagger \otimes t}]$, $\mathbb{E}_U$ is an average over U, and U^H is a Haar random unitary. A random unitary D^E diagonal in a fixed basis E is a diagonal unitary with random phases. Let $(E = \{|k\rangle\}_{k=0}^{d-1}, F = \{|\alpha\rangle\}_{\alpha=0}^{d-1})$ be a pair of mutually unbiased bases. When the phases $\theta_{k\alpha}$ of the inner product, i.e. $\langle k|\alpha\rangle = e^{i\theta_{k\alpha}}/\sqrt{d}$, satisfy $\theta_{k+l,\alpha} = \theta_{k\alpha} + \theta_{l\alpha}$

[*] nakata@qi.t.u-tokyo.ac.jp
[†] christoph.hirche@uab.cat
[‡] koashi@qi.t.u-tokyo.ac.jp
[§] andreas.winter@uab.cat

Table 1: A comparison of quantum circuits generating unitary t-designs on N qubits. The cTPE in the first row stands for the classical tensor product expander. The number of gates needed in the case of cTPE+Fourier transformation is not explicitly given in the original paper [3]. Using the result in Ref. [6] about the efficient sampling of cTPE, the number of gates necessary for the implementation turns out to be $O(t^3N^4)$. If $t \leq 2^{N/50}$, it can be improved to $O(t^2N^3)$. In the last column, we write what if the circuit is interpreted as a dynamics generated by a natural Hamiltonian, which is important to understanding fundamental physics in complex quantum many-body systems.

	Total number of gates	it works for	Natural Hamiltonian
cTPE + Fourier trans. [3]	$O(t^3N^4)$	$t = O(N/\log N))$	N.A.
Local random circuits [4]	$O(t^{10}N^2)$	$t =$poly(N)	Highly time-dependent
Our result (Theorem 2)	$O(tN^2)$	$t = o(N^{1/2})$	Nearly time-independent

$(\forall k, l, \alpha \in [0, d-1])$, where $+$ is an operation making $[0, d-1]$ an additive group, we call the pair *Fourier-type*.

Theorem 1 (Unitary designs on a qudit) *Let (E, F) be a Fourier-type pair of bases on a qudit and let $d = \Omega(t^2 t!)$. A random unitary $D[\ell] := (\prod_{i=\ell}^{1} D_i^E D_i^F) D_0^E$, where D_i^E and D_i^F are independent random diagonal unitaries in the basis of E and F, respectively, is an ϵ-approximate unitary t-design if $\ell \geq \frac{1}{\log d - 2\log t!}(t \log d + \log 1/\epsilon)$.*

On N qubits, we define a random quantum circuit $\mathcal{C}_Z^{(t)}$ diagonal in the Pauli-Z basis, where random diagonal two-qubit gates in the form of $(\mathrm{diag}_Z\{1, e^{i\varphi_1}\} \otimes \mathrm{diag}_Z\{1, e^{i\varphi_2}\})\mathrm{diag}_Z\{1, 1, 1, e^{i\vartheta}\}$ are applied onto all pairs of two qubits, where φ_1 and φ_2 are chosen independently from a set $\{2\pi m/(t+1) : m \in [0, t]\}$ uniformly at random, and ϑ from $\{2\pi m/(\lfloor t/2 \rfloor + 1) : m \in [0, \lfloor t/2 \rfloor]\}$.

Theorem 2 (Unitary designs on N qubits) *Let $t = o(N^{1/2})$. Then, $(\mathcal{C}_Z^{(t)} H_N)^{2\ell} \mathcal{C}_Z^{(t)}$, where H_N is the Hadamard transformation on N qubits, is an ϵ-approximate unitary t-design if $\ell \geq t + \frac{1}{N}\log_2 1/\epsilon$. The number of two-qubit gates and random bits are $O\big(N(tN + \log_2 1/\epsilon)\big)$ and $O\big((\log_2 t)N(tN + \log_2 1/\epsilon)\big)$, respectively.*

The third main result is about an *ϵ-approximate t-design Hamiltonian*, a random Hamiltonian of which dynamics forms an ϵ-approximate unitary t-design at any time after a *design time*.

Corollary 3 (Design Hamiltonians) *Let $t = o(N^{1/2})$ and $\mathfrak{H}_{XZ}^{(t)}$ be a set of 2-local time-dependent Hamiltonians in the form of*

$$H_{XZ}(T) = \begin{cases} H_Z^{(m)} & \textit{if} \;\; 2m\pi \leq T < (2m+1)\pi, \\ H_X^{(m)} & \textit{if} \;\; (2m+1)\pi \leq T < 2(m+1)\pi, \end{cases} \tag{1}$$

where T denotes time, $m = 0, 1, \cdots$, $H_W^{(m)} = \{-\sum_{j<k} J_{ik} W_j \otimes W_k - \sum_j B_j W_j\}_{J_{jk}, B_j \in \mathcal{P}_t}$ ($W = X, Z$), and $\mathcal{P}_t := \{j/(2\lfloor t/2 \rfloor + 1) : j \in [-\lfloor t/2 \rfloor, \lfloor t/2 \rfloor]\}$ ($\lfloor x \rfloor$ is the floor function). Then, a Hamiltonian drawn uniformly at random from $\mathfrak{H}_{XZ}^{(t)}$ is an ϵ-approximate t-design Hamiltonian and the design time is $(2t + 1 + \frac{2}{N}\log 1/\epsilon)\pi$.

The design Hamiltonian H_{XZ} generates a unitary design in a short time irrespective of the system size. For the Hamiltonians with local interactions, we conjecture that, as a generalisation of the fast scrambling conjecture, there exists a natural design Hamiltonian with time-independent local interactions that achieves unitary designs in $O(t \log N)$ time.

References

[1] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal. Quantum data hiding. *IEEE Trans. Inf. Theory*, 48:580, 2002.

[2] C. Dankert, R. Cleve, J. Emerson, and E. Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Phys. Rev. A*, 80:012304, 2009.

[3] A. W. Harrow and R. A. Low. Efficient Quantum Tensor Product Expanders and k-Designs. In *Proc. RANDOM'09.*

[4] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki. Local random quantum circuits are approximate polynomial-designs. *Commun. Math. Phys.*, 346(2):397–434, September 2016.

[5] Y. Nakata, C. Hirche, M. Koashi, and A. Winter. Efficient Quantum Pseudorandomness with Nearly Time-Independent Hamiltonian Dynamics. *Phys. Rev. X*, 7(2):021006, 2017. see also arXiv:1609.07021.

[6] S. Hoory and A. Brodsky. Simple Permutations Mix Even Better. *arXiv:math/0411098*, 2004.

[7] P. Sen. Random measurement bases, quantum state distinction and applications to the hidden subgroup problem. *arXiv:quant-ph/0512085*, 2005.

[8] F. G. S. L. Brandão and M. Horodecki. Exponential Quantum Speed-ups are Generic. *Q. Inf. Comp. 13, 0901*, 2013.

[9] R. Kueng, H. Rauhut, and U. Terstiege. Low rank matrix recovery from rank one measurements. *arXiv:1410.6913*, 2014.

[10] S. Kimmel and Y.-K. Liu. Quantum compressed sensing using 2-designs. arXiv:1510.08887, 2015.

[11] R. A. Low. Large deviation bounds for k-designs. *Proc. R. Soc. A*, 465(2111):3289, 2009.

A semi-device-independent framework based on natural physical assumptions and its application to random number generation

Stefano Pironio

Brussels, Belgium

Abstract: The semi-device-independent approach provides a framework for prepare-and-measure quantum protocols using devices whose behaviour does not need to be characterized or trusted, except for a single assumption on the dimension of the Hilbert space characterizing the quantum carriers. Here, we propose instead to constrain the quantum carriers through a bound on the mean value of a well chosen observable. This modified assumption is physically better motivated than a dimension bound and closer to the description of actual experiments. In particular, we consider quantum optical schemes where the source emits quantum states described in an infinite-dimensiona Fock space and model our assumption as an upper bound on the average photon number in the emitted states. We characterize the set of correlations that may be exhibited in the simplest possible scenario compatible with our new framework, based on two energy-constrained state preparations and a two-outcome measurement. Interestingly, we uncover the existence of quantum correlations exceeding the set of classical correlations that can be produced by devices behaving in a purely pre-determined fashion (possibly including shared randomness). This feature suggests immediate applications to certified randomness generation. Along this line, we analyze the achievable correlations in several prepare-and-measure optical schemes with a mean photon-number constraint and demonstrate that they allow for the generation of certified randomness. Our simplest optical scheme works by the on-off keying of an attenuated laser source followed by photocounting. It opens the path to more sophisticated energy-constrained semi-device-independent quantum cryptography protocols, such as quantum key distribution.