Quantum Lossless Data Compression Theorem for Ergodic Information Sources

Igor Bjelaković[1] * Arleta Szkoła[1] †

[1] *Technische Universität Berlin,*
Institut für Mathematik MA 7-2,
Strasse des 17. Juni 136, D-10623 Berlin, Germany.

Abstract. We extend Schumacher's noisless coding theorem for quantum i.i.d. sources to the case of ergodic quantum information sources. We formulate and prove the quantum data compression theorem using both the entanglement fidelity and the ensemble fidelity as measures of reliability of compression schemes. Moreover, we provide an asymptotically optimal compression scheme which is based on the concept of high probability subspaces. The rate of this compression scheme is equal to the von Neumann entropy rate, which is just the asymptotic rate of the logarithmic dimension of high probability subspaces as was shown in our previous work on the quantum Shannon-McMillan theorem.

Keywords: quantum data compression, asymptotically lossless quantum coding, high probability subspaces, fidelity notions

In classical information theory the Shannon entropy rate h of discrete stochastic processes modelling *information sources* (IS) gives the average information carried by individual signals: Operationally it means that any ergodic information source can be compressed by means of block coding using asymptotically not more than h bits per signal in a way that there exist decompression algorithms with asymptotically vanishing probability of error. Using an exponentially smaller number of bits the compression/decompression algorithms will fail to be asymptotically error-free.
In quantum information theory the corresponding quantity is the von Neumann entropy rate s of *quantum information sources* (QIS) modelled by quantum dynamical systems $(\mathcal{A}^\infty, \Psi, T)$, where $\mathcal{A}^\infty$ is a quasilocal C^*-algebra over $\mathbb{Z}$ constructed from a finite dimensional C^*-algebra $\mathcal{A}$, Ψ is a state and T the shift on $\mathcal{A}^\infty$. If Ψ is an extremal point in the compact convex set of stationary (T-invariant) states on $\mathcal{A}^\infty$, then the QIS is said to be *ergodic*. For simplicity in the sequel we restrict to the case $\mathcal{A} = \mathcal{B}(\mathcal{H})$, where $\mathcal{B}(\mathcal{H})$ is the algebra of linear operators on a finite dimensional Hilbert space $\mathcal{H}$. Notice that a stationary state Ψ correspondes one-to-one to a consistent family of density operators $\{\rho^{(n)}\}_{n\in\mathbb{N}}$ on $\mathcal{H}^{\otimes n}$, respectively.
A *compression scheme* $(\mathcal{C}, \mathcal{D})$ for stationary QIS is a sequence $\{(\mathcal{C}^{(n)}, \mathcal{D}^{(n)})\}_{n\in\mathbb{N}}$ of pairs of trace preserving quantum operations

$$\begin{aligned} \mathcal{C}^{(n)} : \mathcal{S}(\mathcal{H}^{\otimes n}) &\longrightarrow \mathcal{S}(\mathcal{H}^{(n)}), \\ \mathcal{D}^{(n)} : \mathcal{S}(\mathcal{H}^{(n)}) &\longrightarrow \mathcal{S}(\mathcal{H}^{\otimes n}) \end{aligned}$$

where $\mathcal{H}^{(n)} \subseteq \mathcal{H}^{\otimes n}$ for all $n \in \mathbb{N}$ and $\mathcal{S}(\mathcal{H}^{\otimes n})$ resp. $\mathcal{S}(\mathcal{H}^{(n)})$ denote the sets of density operators on $\mathcal{H}^{\otimes n}$ resp. $\mathcal{H}^{(n)}$. We refer to $\mathcal{C}^{(n)}$ resp. $\mathcal{D}^{(n)}$ as (n-block) compression resp. decompression map. Recall that each trace preserving quantum operation $\mathcal{E}$ on $\mathcal{S}(\mathcal{H})$ posseses the following representation known as Kraus or sum representation (cf. [4], [7])

$$\mathcal{E}(\rho) = \sum_i E_i \rho E_i^*,$$

where $E_i \in \mathcal{B}(\mathcal{H})$ and $\sum_i E_i^* E_i = \mathbf{1}$.
The *rate* $R(\mathcal{C})$ of a compression scheme $(\mathcal{C}, \mathcal{D})$ is defined by

$$R(\mathcal{C}) := \limsup_{n\to\infty} \frac{\log_2 \dim \mathcal{H}^{(n)}}{n},$$

and can be interpreted as the asymptotic number of qubits used per signal to represent the QIS.
We can show by construction that for any ergodic QIS there exists an asymptotically reliable compression scheme with rate R equal to the von Neumann entropy rate s.
Designing compression algorithms the goal is to achieve low rates. It turns out that s is the optimal rate in the sense that a rate $R \geq s$ is a necessary condition on asymptotical reliability of compression schemes and even more the fidelity of any compression scheme with rate $R < s$ vanishes asymptotically. Of course, this result depends on the underlying fidelity notion for the compression/decompression operations on the QIS. There are different definitions of fidelity suited for different applications. We use the entanglement fidelity F_e and the ensemble fidelity $\bar{F}$. Consider the fidelity F between two density operators ρ and σ on a Hilbert space $\mathcal{H}$:

$$F(\rho, \sigma) := \mathrm{tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}.$$

Then F_e and $\bar{F}$ are defined by

$$\begin{aligned} F_e(\rho, \mathcal{E}) &:= \left(F(|\Psi\rangle\langle\Psi|, (\mathbf{1}\otimes\mathcal{E})(|\Psi\rangle\langle\Psi|))\right)^2, \\ \bar{F}(\{(p_i, \rho_i)\}_{i=1}^k, \mathcal{E}) &:= \sum_{i=1}^k p_i (F(\rho_i, \mathcal{E}(\rho_i)))^2, \end{aligned}$$

where $\mathcal{E}$ is a quantum operation, ρ is a density operator on $\mathcal{H}$ with a purification $|\Psi\rangle \in \mathcal{H}^{\otimes 2}$ and $\{(p_i, \rho_i)\}_{i=1}^k$ is a weighted ρ-ensemble of k quantum states ρ_i. If the ρ_i

*igor@math.tu-berlin.de
†szkola@math.tu-berlin.de

are all pure states then we call the ensemble a pure one. Our result holds for $\bar{F}$ as well as for F_e, [3]. Define

$$F_s(\rho, \mathcal{E}) := \sup\{\bar{F}(\{(p_i, P_i)\}_{i=1}^k, \mathcal{E})|\ \{(p_i, P_i)\}_{i=1}^k \text{ pure } \rho\text{-ensemble}\}. \quad (1)$$

Using the relation

$$0 \leq F_e(\rho, \mathcal{E}) \leq \bar{F}(\{(p_i, \rho_i)\}_{i=1}^k) \leq 1,$$

which holds for arbitrary ρ-ensembles $\{(p_i, \rho_i)\}_{i=1}^k$, the data compression theorem reads as follows:

Theorem 1 (Data Compression Theorem) *Let $(\mathcal{A}^\infty, \Psi, T)$ be an ergodic quantum information source with the entropy rate s.*

1) *Each compression scheme $(\mathcal{C}, \mathcal{D})$ satisfying*

$$\lim_{n\to\infty} \bar{F}(\{(p_i^{(n)}, P_i^{(n)})\}_{i=1}^{k_n}, \mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}) = 1$$

for some sequence $\{\{(p_i^{(n)}, P_i^{(n)})\}_{i=1}^{k_n}\}_{n\in\mathbb{N}}$ of pure $\rho^{(n)}$-ensembles, respectively, fulfills

$$R(\mathcal{C}) \geq s.$$

2) *There exists a compression scheme $(\mathcal{C}, \mathcal{D})$ with $R(\mathcal{C}) = s$ such that*

$$\lim_{n\to\infty} F_e(\rho^{(n)}, \mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}) = 1.$$

3) *Any compression scheme $(\mathcal{C}, \mathcal{D})$ with $R(\mathcal{C}) < s$ satisfies*

$$\lim_{n\to\infty} F_s(\rho^{(n)}, \mathcal{D}^{(n)} \circ \mathcal{C}^{(n)}) = 0,$$

where F_s is defined by (1).

The main tool to construct compression schemes achieving the optimal rate s are high probability subspaces. A subspace $\mathcal{P}_\varepsilon^{(n)} \subseteq \mathcal{H}^{\otimes n}$ is called a high probability subspace to the level $\varepsilon \in (0,1)$ if

$$\begin{aligned} \mathrm{tr}\rho^{(n)} P_{\mathcal{P}_\varepsilon^{(n)}} &\geq 1-\varepsilon \\ \text{and} \quad \log\dim \mathcal{P}_\varepsilon^{(n)} &\leq \log\dim \mathcal{P}^{(n)} \end{aligned}$$

for any $\mathcal{P}^{(n)} \subseteq \mathcal{H}^{\otimes n}$ satisfying $\mathrm{tr}\rho^{(n)} P_{\mathcal{P}^{(n)}} \geq 1-\varepsilon$, where $P_{\mathcal{P}}$ denotes the projection onto a subspace $\mathcal{P}$.
Block compression maps that are essentially projections onto high probability subspaces, as for example

$$\mathcal{C}^{(n)}(\rho^{(n)}) := P_{\mathcal{P}_\varepsilon^{(n)}} \rho^{(n)} P_{\mathcal{P}_\varepsilon^{(n)}} + \sum_{e\in S^{(n)}} |0\rangle\langle e|\rho^{(n)}|e\rangle\langle 0|,$$

where $|0\rangle \in \mathcal{P}_\varepsilon^{(n)}$ and $S^{(n)}$ is an orthonormal system in $(\mathcal{P}_\varepsilon^{(n)})^\perp$, together with decompression maps $\mathcal{D}^{(n)}$, which are just the canonical embedding of $\mathcal{S}(\mathcal{P}_\varepsilon^{(n)})$ into $\mathcal{S}(\mathcal{H}^{\otimes n})$ provide a solution to the problem of optimal data compression. This is due to a basic result concerning high probability subspaces of ergodic QIS, which was proved in [2]. It asserts the convergence of the logarithmic dimension rate of high probability subspaces $\mathcal{P}_\varepsilon^{(n)}$ to the von Neumann entropy rate s independent of the level ε. This convergence was already conjectured by Petz/Hiai in [5].

Proposition 2 *Let $(\mathcal{A}^\infty, \Psi, T)$ be an ergodic QIS with the entropy rate s. Then for every $\varepsilon \in (0,1)$*

$$\lim_{n\to\infty} \frac{1}{n} \min\{\log_2(\dim \mathcal{P}^{(n)})|\ \mathcal{P}^{(n)} \subseteq \mathcal{H}^{\otimes n}, tr\rho^{(n)} P_{\mathcal{P}^{(n)}} \geq 1-\varepsilon\} = s.$$

The concept of high probability subspaces is crucial in the work of Petz/Mosonyi [8], where they prove a coding theorem for the class of *completely* ergodic QIS, that means for QIS, which are ergodic with respect to T^n, for all $n \in \mathbb{N}$. Using projections onto high probability subspaces they show that completely ergodic QIS can be compressed with any rate $R \geq s$ in such a way that the ensemble fidelity $\bar{F}$ is asymptotically equal to 1. On the other hand $\bar{F}$ cannot achieve 1 asymptotically if the rate satisfies $R < s$. The reason why they cannot conclude that in fact for $R < s$ the asymptotical fidelity $\bar{F}$ is equal to 0 is that they use the result of Hiai/Petz [5] which provides bounds on limit superior and limit inferior and not the limit of the minimal logarithmic dimension rate of the high probabilty subspaces. The result of Petz/Mosonyi represents an extension of the coding theorem formulated in [6] by Jozsa and Schumacher for the smaller class of independent identically distributed (i.i.d.) QIS and proved in [6] and [1]. An analogous result for i.i.d. QIS using the entanglement fidelity F_e as a criterion for the reliability of compression schemes is presented by Nielsen and Chuang in [7].

References

[1] H. Barnum, C. Fuchs, R. Jozsa, B. Schumacher, General Fidelity Limit for Quantum Channels, Phys. Rev. A 54, No 6, 4707-4711 (1996)

[2] I. Bjelaković, T. Krüger, Ra. Siegmund-Schultze, A. Szkoła, The Shannon-McMillan Theorem for Ergodic Quantum Lattice Systems, to appear in Inventiones Mathematicae 2003, arXiv: math.DS/0207121

[3] I. Bjelaković, A. Szkoła, The Data Compression Theorem for Ergodic Quantum Information Sources, to appear in Quantum Information Processing 2003, arXiv: quant-ph/0301043

[4] K.-E. Hellwig, K.Kraus, Operations and Measurements II, Commun. Math. Phys. 16, 142-147 (1970)

[5] F. Hiai, D. Petz, The Proper Formula for Relative Entropy and its Asymptotics in Quantum Probability, Commun. Math. Phys. 143, 99-114 (1991)

[6] R. Jozsa, B. Schumacher, A New Proof of the Quantum Noisless Coding theorem, Journal of Modern Optics, Vol.41, No.12, 2343-2349, (1994)

[7] M. A. Nielsen, I. L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, Cambridge 2000

[8] D. Petz, M. Mosonyi, Stationary Quantum Source Coding, J. Math. Phys. 42, 4857-4864 (2001)