Quantum universal variable-length source coding

Masahito Hayashi
Lab. for Mathematical Neuroscience
Brain Science Institute, RIKEN
Hirosawa, Wako, Saitama, 351-0198, Japan
e-mail: masahito@brain.riken.go.jp

Keiji Matsumoto
Quantum Computation and Information Project
ERATO, JST
5-28-3, Hongo, Bunkyo-ku,Tokyo, 113-0033, Japan
e-mail: keiji@qci.jst.go.jp

Abstract — **Independently of the input probability distribution p, we construct the** *quantum universal variable-length source code* **in which the average error concerning to Bures' distance tends to 0 and the probability that the coding rate is greater than the entropy rate $H(\overline{\rho}_p)$, tends to 0. If we can estimate the entropy $H(\overline{\rho}_p)$, we can compress the coding rate to the admissible rate $H(\overline{\rho}_p)$ with a probability close to 1. However, when we perform a naive measurement for the estimation of $H(\overline{\rho}_p)$, the input state is destroyed. Therefore, in our code, it is the main problem to treat the trade-off between the estimation of $H(\overline{\rho}_p)$ and the non-demolition of the input state.**

I. Review of quantum fixed-length source coding

Let $\mathcal{H}$ be a finite-dimensional Hilbert space that represents the physical system of interest and let $\mathcal{S}(\mathcal{H})$ be the set of density operators on $\mathcal{H}$. Consider a source which produces the pure state $\vec{\rho}_n := \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$ with the i.i.d. distribution p^n of the probability p on pure states. In *fixed-length source coding*, a sequence of states $\vec{\rho}_n$ is compressed to the state in a smaller Hilbert space $\mathcal{H}_n \subset \mathcal{H}^{\otimes n}$, whose dimension is e^{nR}. Here, the encoder and the decoder is a trace-preserving completely positive (TP-CP) map E^n and D^n, respectively. The average of the total error is given by

$$\epsilon_{n,p}(E^n, D^n) := \sum_{\vec{\rho}_n \in \mathcal{S}(\mathcal{H}^{\otimes n})} p^n(\vec{\rho}_n) b^2 \left(\vec{\rho}_n, D^n \circ E^n(\vec{\rho}_n)\right),$$

where Bures' distance is defined as $b(\rho,\sigma) := \sqrt{1 - \operatorname{Tr}\left|\sqrt{\rho}\sqrt{\sigma}\right|}$. In this setting, we focus on the infimum of the rate with which the average error goes to zero. The infimum is called the minimum admissible rate R_p of p, and is defined by

$$R_p := \inf \left\{ \limsup \frac{1}{n} \log \dim \mathcal{H}_n \,\middle|\, \begin{array}{l} \exists \{(\mathcal{H}_n, E^n, D^n)\}, \\ \epsilon_{n,p}(E^n, D^n) \to 0 \end{array} \right\}.$$

As was proven by Schumacher [1], and Jozsa and Schumacher [2], and Barnum et al. [9], the equation $R_p = H(\overline{\rho}_p) := -\operatorname{Tr} \overline{\rho}_p \log \overline{\rho}_p$ holds, where $\overline{\rho}_p := \sum_{\rho \in \mathcal{S}(\mathcal{H})} p(\rho)\rho$. Moreover Jozsa et al. [3] constructed the projections $P_{R,n}$ for a arbitarary rate R such that

$$\operatorname{rank} P_{R,n} \le e^{nR}, \quad \operatorname{Tr} P_{R,n} \rho^{\otimes n} \to 1, \tag{1}$$

for any density matrix ρ satisfynig $H(\rho) < R$, and proposed a *quantum universal fixed-length source code* depending only on the entropy rate.

II. Quantum universal variable-length source coding

In the classical system, depending on the input state, the encoder can determine the coding length. Such a code is called a variable-length code. Using this type code, we can compress any information without error. Lynch [4] and Davisson [5] proposed a variable-length code with no error, in which the coding rate is less than $H(p)$ except for a small enough probability under the distribution p. Such a code is called a universal variable-length source code. Today, their code can be regarded as the following two-stage code: at the first step, we send the empirical distribution which indicates a subset of data, and in the second step, we send information which indicates every sequence belonging to the subset.

This paper deals with quantum data compression in which the encoder determines the coding length, according to the input state. In order to make this decision, he must measure the input quantum system. Thus, we need describe a quantum measurement with state evolution, by using *an instrument* consisting of a decomposition $\mathbf{E}' = \{\mathbf{E}'_\omega\}_{\omega \in \Omega}$, by CP maps from $\mathcal{S}(\mathcal{H})$ to $\mathcal{S}(\mathcal{H})$ under the condition $\sum_{\omega \in \Omega} \operatorname{Tr} \mathbf{E}'_\omega(\rho) = 1, \quad \forall \rho \in \mathcal{S}(\mathcal{H})$. When we perform the instrument $\mathbf{E}' = \{\mathbf{E}'_\omega\}_{\omega \in \Omega}$ for an initial state ρ, we get the data ω and the final state $\frac{\mathbf{E}'_\omega(\rho)}{\operatorname{Tr} \mathbf{E}'_\omega(\rho)}$ with the probability $\operatorname{Tr} \mathbf{E}'_\omega(\rho)$. A quantum variable-length encoder $\mathbf{E}$ is given by a measurement process $\mathbf{E}'$ and encoding process E''_ω depending on the data ω, which is a TP-CP map from $\mathcal{S}(\mathcal{H})$ to $\mathcal{S}(\mathcal{H}_\omega)$, where the Hilbert space $\mathcal{H}_\omega$ depends on the data ω, as $\mathbf{E}_\omega = E''_\omega \circ \mathbf{E}'_\omega$. Therefore, any quantum variable-length encoder $\mathbf{E}$ consists of a decomposition $\mathbf{E} = \{\mathbf{E}_\omega\}_{\omega \in \Omega}$, by CP maps from $\mathcal{S}(\mathcal{H})$ to $\mathcal{S}(\mathcal{H}_\omega)$ under the condition $\sum_{\omega \in \Omega} \operatorname{Tr} \mathbf{E}_\omega(\rho) = 1, \quad \forall \rho \in \mathcal{S}(\mathcal{H})$. For a detail about instruments, see Ozawa [8].

The decoder is given by a set of TP-CP maps $\mathbf{D} = \{\mathbf{D}_\omega\}_{\omega \in \Omega}$, which presents the decoding process depending on the data ω. A pair of an encoder $\mathbf{E} = \{\mathbf{E}_\omega\}_{\omega \in \Omega}$ and a decoder $\mathbf{D} = \{\mathbf{D}_\omega\}_{\omega \in \Omega}$ is called a *quantum variable-length source code* on $\mathcal{H}$. The coding length is described by $\log |\Omega| + \log \dim \mathcal{H}_\omega$, which is a random variable obeying the probability $\mathrm{P}^{\mathbf{E}}_\rho(\omega) := \operatorname{Tr} \mathbf{E}_\omega(\rho)$ when the input state is ρ.

When the state $\vec{\rho}_n$ on $\mathcal{H}^{\otimes n}$ obeys the i.i.d. distribution p^n of the probability p on pure states, the error of decoding for a variable-length code $(\mathbf{E}^n, \mathbf{D}^n)$ on $\mathcal{H}^{\otimes n}$ is evaluated by Bures' distance as

$$\sum_{\omega_n \in \Omega_n} \operatorname{Tr} \mathbf{E}^n_{\omega_n}(\vec{\rho}_n) b^2 \left(\vec{\rho}_n, \mathbf{D}^n_{\omega_n} \left(\frac{\mathbf{E}^n_{\omega_n}(\vec{\rho}_n)}{\operatorname{Tr} \mathbf{E}^n_{\omega_n}(\vec{\rho}_n)} \right) \right),$$

and the average error is given by

$$\begin{aligned} \epsilon_{n,p}(\mathbf{E}^n, \mathbf{D}^n) &:= \sum_{\vec{\rho}_n \in \mathcal{S}(\mathcal{H}^{\otimes n})} p^n(\vec{\rho}_n) \\ &\times \sum_{\omega_n \in \Omega_n} \operatorname{Tr} \mathbf{E}^n_{\omega_n}(\vec{\rho}_n) b^2 \left(\vec{\rho}_n, \mathbf{D}^n_{\omega_n} \left(\frac{\mathbf{E}^n_{\omega_n}(\vec{\rho}_n)}{\operatorname{Tr} \mathbf{E}^n_{\omega_n}(\vec{\rho}_n)} \right) \right). \end{aligned}$$

In this case, the data ω_n obeys the probability:

$$\mathrm{P}^{\mathbf{E}^n}_{p^n}(\omega_n) := \sum_{\vec{\rho}_n \in \mathcal{S}(\mathcal{H}^{\otimes n})} p^n(\vec{\rho}_n) \operatorname{Tr} \mathbf{E}^n_{\omega_n}(\vec{\rho}_n) = \operatorname{Tr} \mathbf{E}^n_{\omega_n}(\overline{\rho}_p^{\otimes n}).$$

A sequence $\{(\mathbf{E}^n, \mathbf{D}^n)\}$ of quantum variable-length source code is called *universal* if $\epsilon_{n,p}(\mathbf{E}^n, \mathbf{D}^n) \to 0$ for any probability p on pure states.

As mentioned latter, there exists a quantum universal variable-length source code $\{(\mathbf{E}^n, \mathbf{D}^n)\}$ satisfying

$$\lim \mathrm{P}^{\mathbf{E}^n}_{p^n} \left\{ \frac{1}{n} (\log |\Omega_n| + \log \dim \mathcal{H}_{\omega_n}) \geq H(\overline{\rho}_p) + \epsilon \right\} = 0$$

for any $\epsilon > 0$. Conversely, if a quantum variable-length source code $\{(\mathbf{E}^n, \mathbf{D}^n)\}$ is universal and

$$\lim \mathrm{P}^{\mathbf{E}^n}_{p^n} \left\{ \frac{1}{n} (\log |\Omega_n| + \log \dim \mathcal{H}_{\omega_n}) \geq R \right\} = 0,$$

then $R \geq R_p = H(\overline{\rho}_p)$.

III. Construction of a quantum variable-length source code

First, for an intuitive explanation of our construction, we naively construct a good variable-length code. For this construction, we fixed a strictly increasing sequence $\vec{a} := \{a_i\}_{i=1}^{l+1}$ of real numbers such that $0 = a_1 < a_2 < \ldots < a_l < a_{l+1} = \log d$. We define the encoder $\mathbf{E}^{\vec{a},n}$ with the data set $\{1, \ldots, l\}$ by

$$\begin{aligned} P^{\vec{a},n}_i &:= P_{a_{i+1},n} - P_{a_i,n} \\ \mathbf{E}^{\vec{a},n}_i(\rho_n) &:= P^{\vec{a},n}_i \rho_n P^{\vec{a},n}_i, \quad \rho_n \in \mathcal{S}(\mathcal{H}^{\otimes n}), \end{aligned}$$

and define the decoder $\mathbf{D}^{\vec{a},n}_i$ as the embedding to $\mathcal{H}^{\otimes n}$. Assume that $H(\overline{\rho}_p)$ belongs to the interval $[a_i, a_{i+1})$. As is guaranteed by (1), if $H(\overline{\rho}_p)$ does not lie on the boundary on the interval $[a_i, a_{i+1})$, the probability $\operatorname{Tr} \overline{\rho}_p^{\otimes n} P^{\vec{a},n}_i$ tends to 1. Thus, we can prove $\epsilon_{n,p}(\mathbf{E}^{\vec{a},n}, \mathbf{D}^{\vec{a},n}) \to 0$. Of course, if we choose $a_{i+1} - a_i$ to be sufficiently small, the coding length is close to the entropy $H(\overline{\rho}_p)$ with almost probability 1. However, if $H(\overline{\rho}_p)$ lies on the boundary, the state is demolished, as is caused by the same reason of Lemma 2 in [7]. In this case, we can prove $\lim \epsilon_{n,p}(\mathbf{E}^{\vec{a},n}, \mathbf{D}^{\vec{a},n}) > 0$. Thus, it is not universal.

Next, we assume that the interval $a_{i+1} - a_i$ ($i = 2, \ldots, l-1$) is $\delta := \log d/(l-1)$ and that $a_2 - a_1, a_{l+1} - a_l < \delta$. Then, our code is uniquely defined by the choice of $a_2 \in (0, \delta)$. For the non-demolition of initial states, we construct a variable-length code, by choosing $a_2 \in \{\frac{k}{n} | \frac{k}{n} \in (0,\delta), k \in \mathbb{Z}\}$ at random. In this protocol, the set $\{\frac{k}{n} | \frac{k}{n} \in (0,\delta), k \in \mathbb{Z}\} \times \{1, 2, \ldots, l, l+1\}$ corresponds to the data set Ω_n, and we can expect that the average error tends to 0 for any probability p on pure states. In order to achieve the rate $H(\overline{\rho}_p)$, we need to choose the set Ω_n so that $\frac{1}{n} \log |\Omega_n| \to 0$. It is essential in our code to restrict a_2 to this lattice $\{\frac{k}{n} | k \in \mathbb{Z}\}$.

Moreover, when δ is large for a fixed number n, the demolition of initial state seems small and the coding length seems long. Therefore, roughly speaking, in this code for a finite number n, by choosing δ, we can treat the trade off between the coding length and the non-demolition of the input state. In Hayashi and Matsumoto [7], using the representation theory like as Keyl and Werner [6], we construct a code in which the average error uniformlly goes to 0 and the exponent of its overflow probability is optimal. In order to satisfy the universality and the optimality of the overflow exponent, we need choose δ depending on n, more carefully.

IV. Discussion

In our code, the nonzero number δ is essential. One may expect that the quantum variable-length source code $\{(\mathbf{E}^{0,n}, \mathbf{D}^{0,n})\}$ is universal. However, this code destroys the input state by a quantum measurement [7]. Moreover, it seems impossible to construct a universal code whose average error $\epsilon_{n,p}$ exponentially tends to 0 [7].

Acknowledgments

The authors wish to thank Professor H. Nagaoka and Dr. A. Winter for useful comments.

References

[1] B. Schumacher, Phys. Rev. A, **51**, 2738-2747, (1995).

[2] R. Jozsa and B. Schumacher, J. Mod. Optics, 41(12), 2343-2349, (1994).

[3] R. Jozsa, M. Horodecki, P. Horodecki and R. Horodecki, Phys. Rev. Lett. **81**, 1714 (1998); LANL eprint quant-ph/9805017.

[4] T. J. Lynch, Proc. IEEE, **54**, 1490-1491, (1966).

[5] L. D. Davisson, Proc. IEEE, **54**, 2010, (1966).

[6] M. Keyl and R. F. Werner, Phys. Rev. A **64**, 052311 (2001); LANL eprint quant-ph/0102027, (2001).

[7] M. Hayashi and K. Matsumoto, LANL eprint quant-ph/0202002, (2002); appear in Phys. Rev. A.

[8] M. Ozawa, J. Math. Phys. **25**, 79 (1984); M. Ozawa, Phys. Rev. A **62**, 062101 (2000).

[9] H. Barnum, C. A. Fuchs, R. Jozsa and B. Schumacher, Phys. Rev. A, **54**, 4707-4711 (1996).