Optimal quantum data compression using dynamical entropy

Physical Review A 100, 032301 (2019)

George Androulakis (joint work with Duncan Wright)

University of South Carolina, Department of Mathematics

giorgis@math.sc.edu — http://people.math.sc.edu/androula/

UNIVERSITY OF SOUTH CAROLINA

Quantum Data Compression for an i.i.d. quantum source

Notation

$H^{\oplus} := \oplus_{\ell=1}^{\infty} H^{\otimes \ell}$ is the **Free Fock space** associated with a Hilbert space H.

Given a **symbol set** $\mathcal{S}$ containing normalized vectors (pure states) spanning a Hilbert space $H_{\mathcal{S}}$, a **quantum source of symbols** produces finite strings (tensor products) of symbols $|x_1\rangle|x_2\rangle\cdots|x_n\rangle = |x_1x_2\cdots x_n\rangle \in \mathcal{S}^{\otimes n} \subseteq H_{\mathcal{S}}^{\otimes n}$ with $n \in \mathbb{N}$ and $|x_k\rangle \in \mathcal{S}$ for every k. For every $n \in \mathbb{N}$ the set $\mathcal{S}^{\otimes n}$ is equipped with a probability distribution $\mathbb{P}$ so that each string $|x_1\cdots x_n\rangle$ is produced with probability $\mathbb{P}(||x_1\cdots x_n\rangle)$. Let $(X_n)_{n\in\mathbb{N}}$ be the stochastic process describing the string of symbols produced. If $(X_n)_{n\in\mathbb{N}}$ is i.i.d. then we say that **the quantum source of symbols is i.i.d.** In that case, $\mathbb{P}(|x_1\cdots x_n\rangle) = \mathbb{P}(|x_1\rangle)\cdots\mathbb{P}(|x_n\rangle)$.
A **quantum code** is a linear isometry $U : H_{\mathcal{S}} \to (\mathbb{C}^2)^{\oplus}$. This induces a TPCP map on the set $\Sigma(H_{\mathcal{S}})$ of the states of $H_{\mathcal{S}}$ defined by $\rho \mapsto U\rho U^*$.
The **extended quantum code** $U^{\oplus} : H_{\mathcal{S}}^{\oplus} \to (\mathbb{C}^2)^{\oplus}$ is defined by "concatenation" (i.e. tensor products of the values of U).
The quantum code U is called **uniquely decodable** if $U^{\oplus}$ is an isometry.
The **average state produced** by the quantum source is equal to $\rho_{\mathcal{S}} = \sum_{|x\rangle\in\mathcal{S}} \mathbb{P}(|x\rangle)|x\rangle$.
The **von Neumann entropy** of a density matrix ρ is given by $S(\rho) = -\mathrm{tr}\,(\rho \log_2 \rho)$.

Theorem (Noiseless Quantum Coding Theorem) [6]

Assume that a quantum source is i.i.d. and let ρ be the average state produced. For any $\delta > 0$, $(S(\rho)+\delta)$ - many qubits per symbol are sufficient in order to encode strings of symbol states with probability of error tending to zero as the length of the strings tends to infinity.
Moreover, for any $R < S(\rho)$, if at most R qubits are used per symbol, then the probability of error tends to one as the length of the strings tends to infinity.

Interpetation

$S(\rho)$ is the minimum number of qubits per symbol that are necessary for asymptotically uniquely decodable (lossless) coding when the quantum source is i.i.d..

Indeterminate length and optimal codes

Definition [3, 7]

The **(indeterminate) length** of an element $\omega \in (\mathbb{C}^2)^{\oplus}$ is defined by length $(\omega) = \langle\omega|\Lambda|\omega\rangle = \mathrm{tr}\,(|\omega\rangle\langle\omega|\Lambda)$ where $\Lambda : (\mathbb{C}^2)^{\oplus} \to (\mathbb{C}^2)^{\oplus}$ is the **length observable** defined by $\Lambda = \sum_{\ell=1}^{\infty} \ell \Pi_\ell$ where $\Pi_\ell : (\mathbb{C}^2)^{\oplus} \to (\mathbb{C}^2)^{\otimes \ell}$ denotes the orthogonal projection. In general, if $\rho \in \Sigma((\mathbb{C}^2)\oplus)$ then define length $(\rho) = \mathrm{tr}\,(\rho\Lambda)$.

Let $U : H_{\mathcal{S}} \to (\mathbb{C}^2)^{\oplus}$ be a uniquely decodable quantum code. For $n \in \mathbb{N}$, the **average length of U over the symbols $\mathcal{S}^{\otimes n}$** is defined as

$$\begin{aligned} \mathrm{AL}_{\mathcal{S}^{\otimes n}}(U) &= \sum_{|x_1\cdots x_n\rangle\in\mathcal{S}^{\otimes n}} \mathbb{P}|x_1\cdots x_n\rangle)\mathrm{length}\,(U^{\otimes n}(|x_1\cdots x_n\rangle)) \\ &= \sum_{|x_1\cdots x_n\rangle\in\mathcal{S}^{\otimes n}} \mathbb{P}(|x_1\cdots x_n\rangle)\mathrm{tr}\,(U^{\otimes n}|x_1\cdots x_n\rangle\langle x_1\cdots x_n|U^{\dagger\otimes n}\Lambda). \end{aligned}$$

For $n \in \mathbb{N}$, a code $U_{\mathrm{opt},\mathcal{S}^{\otimes n}} : H_{\mathcal{S}} \to (\mathbb{C}^2)^{\oplus}$ is called **optimal for the set of symbols $\mathcal{S}^{\otimes n}$** if

$$U_{\mathrm{opt},\mathcal{S}^{\otimes n}} = \mathrm{argmin}_U\{\mathrm{AL}_{\mathcal{S}^{\otimes n}}(U) : U \text{ is a uniquely decodable code }\}.$$

The **optimal average length of $\mathcal{S}^{\otimes n}$** via lossless coding is defined as $\mathrm{OAL}(\mathcal{S}^{\otimes n}) = \mathrm{AL}_{\mathcal{S}^{\otimes n}}(U_{\mathrm{opt},\mathcal{S}^{\otimes n}})$.
The **optimal average length per symbol for the first n symbols** via lossless coding is defined as $\mathrm{OALPS}_n = \frac{\mathrm{OAL}(\mathcal{S}^{\otimes n})}{n}$. The **asymptotic optimal average length per symbol** via lossless coding, (i.e. **asymptotic optimal lossless compression rate**), is given by $R = \limsup_{n\to\infty} \mathrm{OALPS}_n$.

Theorem [3]

For every quantum source producing quantum states from a symbol set $\mathcal{S}$ we have

$$S(\rho_{\mathcal{S}}) \le \mathrm{OAL}(\mathcal{S}) < S(\rho_{\mathcal{S}}) + 1$$

where $\rho_{\mathcal{S}}$ is the average state produced by the quantum source.

Corollary

For every quantum source producing quantum states from a symbol set $\mathcal{S}$ and for every $n \in \mathbb{N}$ we have

$$\frac{1}{n}S(\rho_{\mathcal{S}^{\otimes n}}) \le \mathrm{OALPS}_n < \frac{1}{n}S(\rho_{\mathcal{S}^{\otimes n}}) + \frac{1}{n}$$

where $\rho_{\mathcal{S}^{\otimes n}}$ is the average state produced by the quantum source for the strings of symbols of length equal to n. Hence the asymptotic lossless optimal lossless compression rate is given by

$$R = \limsup_{n\to\infty} \frac{1}{n}S(\rho_{\mathcal{S}^{\otimes n}}).$$

How do we recognize the uniquely decodable codes? In the classical case one uses the Kraft-McMillan Inequality. There are a few versions of quantum Kraft-McMillan inequalities [3,5,7]. Our version generalizes [3] and has a converse statement, but in order to state it we need two definitions:
A quantum code $U : H_{\mathcal{S}} \to (\mathbb{C}^2)^{\oplus}$ is said to **have length eigenstates** if it has the form $U = \sum_i |\psi_i\rangle\langle e_i|$ where $(e_i)_i$ is an o.n. basis of $H_{\mathcal{S}}$ and there exist positive integers $(\ell_i)_i$, (called **length eigenstates**), such that $\psi_i \in (\mathbb{C}^2)^{\otimes \ell_i}$.
A uniquely decodable quantum code $U : H_{\mathcal{S}} \to (\mathbb{C}^2)^{\oplus}$ is called a **classical-quantum scheme**, if there exists a classical code $C : S \to \{0,1\}^+$ with $\#(S) = \dim\,(H_{\mathcal{S}}) = d$ and an o.n. basis $(e_i)_{i=1}^d$ of $H_{\mathcal{S}}$ such that $U = \sum_{i=1}^d |C(x_i)\rangle\langle e_i|$ where $S = \{x_i : i = 1,\ldots,d\}$.

Theorem (Quantum Kraft-McMillan Inequality) G. A. and D. Wright

If U is a uniquely decodable code with length eigenstates then $\mathrm{tr}\,(U^*2^{-\Lambda}U) \le 1$.
Conversely, if U is a linear isometry with length eigenstates satisfying the above inequality then there exists a classical-quantum scheme $\tilde{U}$ having the same number of ℓ eigenstates for every $\ell \in \mathbb{N}$.

Dynamical entropy of a quantum dynamical system

Definition

A **quantum dynamical system** is a triple $(\mathcal{A}, \Theta, \phi)$ where $\mathcal{A}$ is a von Neumann algebra, $\Theta : \mathcal{A} \to \mathcal{A}$ is a positive unital map, and ϕ is a state on $\mathcal{A}$.

Definition [1]

A **quantum Markov chain** is a tuple $(\phi, \mathcal{E})$ where ϕ is state on some von Neumann algebra $\mathcal{A}$, and $\mathcal{E}$ is a **transition expectation**, i.e. a linear bounded positive unital map $\mathcal{E} : M_d \otimes \mathcal{A} \to \mathcal{A}$ for some fixed positive integer d.

Definition [2, 4]

Let $(\mathcal{A}, \Theta, \phi)$ be a quantum dynamical system, and $\gamma = (\gamma_i^*\gamma_i)_{i=1}^d$ be a POVM (acting on the same Hilbert space as the von Neumann algebra $\mathcal{A}$).

Define a quantum Markov chain $(\phi, \mathcal{E})$ as follows: The state ϕ is the same as the one appearing in the quantum dynamical system. In order to define the transition expectation $\mathcal{E}$, first define an auxiliary transition expectation $\mathcal{E}_\gamma : M_d \otimes \mathcal{A} \to \mathcal{A}$ by

$$\mathcal{E}_\gamma([a_{i,j}]_{i,j=1}^d) = \sum_{i,j=1}^d \gamma_i^* a_{i,j} \gamma_j \quad \text{for all } [a_{i,j}]_{i,j=1}^d \in M_d \otimes \mathcal{A},$$

and then define $\mathcal{E} : M_d \otimes \mathcal{A} \to \mathcal{A}$ by $\mathcal{E} = \Theta \circ \mathcal{E}_\gamma$.

Define a **quantum Markov state** ψ on $M_d^{\mathbb{N}}$ of the quantum Markov chain $(\phi, \mathcal{E})$ by

$$\psi(m_1 \otimes \cdots \otimes m_n) = \phi(\mathcal{E}(m_1 \otimes \mathcal{E}(m_2 \otimes \mathcal{E}(\cdots \mathcal{E}(m_n \otimes \mathbb{1})\cdots)))) \quad \text{where } \mathbb{1} \text{ if the identity in } \mathcal{A}.$$

The **joint correlations of a quantum Markov chain** is a sequence $(\rho_n)_{n\in\mathbb{N}}$ where ρ_n is a density matrix in $M_d^{\otimes n}$ defined by $\psi(m_1 \otimes \cdots \otimes m_n) = \mathrm{tr}(\rho_n(m_1 \otimes \cdots m_n))$ for every $n \in \mathbb{N}$ and $m_1, \ldots, m_n \in M_d$.
The **dynamical entropy of the quantum Markov chain $(\phi, \mathcal{E})$** is defined as

$$\limsup_{n\to\infty} \frac{S(\rho_n)}{n},$$

and this is defined to be the **dynamical entropy of the dynamical system $(\mathcal{A}, \Theta, \phi)$ with respect to the POVM γ**.

Quantum data compression for a stationary Markov source

A stochastic process $(X_n)_{n\in\mathbb{N}}$ is called

- **stationary** if $(X_1, \ldots, X_n) \sim (X_{1+k}, \ldots, X_{n+k})$ for all n, k.
- **a Markov process** if $\mathbb{P}(X_{n+1}|X_1, X_2, \ldots, X_n) = \mathbb{P}(X_{n+1}|X_n)$.

Easy Fact

A stochastic process $(X_n)_{n\in\mathbb{N}}$ with values in a set $\mathcal{S}$ is a stationary Markov process if and only if the random variables X_n's are identically distributed and there exists a column stochastic matrix $(p_{i,j})_{i,j\in\mathcal{S}}$ and an invariant distribution $(p_s)_{s\in\mathcal{S}}$ such that for any $n \in \mathbb{N}$ and every $s_1, \ldots s_n \in \mathcal{S}$, $\mathbb{P}(X_1 = s_1, \ldots, X_n = s_n) = p_{s_1} \prod_{j=2}^n p_{s_j, s_{j-1}}$.

Let $(X_n)_{n\in\mathbb{N}}$ be a stochastic process describing the strings of quantum symbols that are produced by a quantum source. If $(X_n)_{n\in\mathbb{N}}$ is a stationary Markov process then we say that **the quantum source is a stationary Markov source**.

Thus every stationary quantum Markov source is characterized by:

- A symbol set $\mathcal{S} = \{s_1, \ldots, s_N\}$ of normalized vectors spanning some Hilbert space $H_{\mathcal{S}}$,
- a column stochastic matrix $(p_{i,j})_{i,j=1}^N$,
- an invariant distribution $(p_i)_{i=1}^N$ for the above column stochastic matrix.

Construction: A quantum dynamical system and a POVM associated with a quantum stationary Markov source

Let a stationary Markov quantum source associated with a symbol set $\mathcal{S} = \{s_1, \ldots, s_N\}$ spanning a d-dimensional Hilbert space $H_{\mathcal{S}}$, a column stochastic matrix $(p_{i,j})_{i,j=1}^N$ and a stationary distribution $(p_i)_{i=1}^N$ as above. We construct a quantum dynamical system $(\mathcal{A}, \Theta, \phi)$ and a POVM as follows:

- $\mathcal{A} = B(\mathbb{C}^N)$ (i.e. all bounded linear operators on $\mathbb{C}^N$),
- $\Theta : B(\mathbb{C}^N) \to B(\mathbb{C}^N)$ defined by $\Theta(|k\rangle\langle\ell|) = \delta_{k,\ell} \sum_{j=1}^N p_{k,j}|j\rangle\langle j|$,
- $\phi(\cdot) = \mathrm{tr}\,(\rho\cdot)$ where $\rho \in \Sigma(\mathbb{C}^N)$, (a state on $\mathbb{C}^N$) is defined by $\rho = \sum_{n=1}^N p_n|n\rangle\langle n|$,
- if $(e_i)_{i=1}^d$ is a fixed o.n. basis of $H_{\mathcal{S}}$ then define operators $(\gamma_i)_{i=1}^d$ on $\mathbb{C}^N$ by $\gamma_i = \sum_{n=1}^N \langle e_i|s_n\rangle|n\rangle\langle n|$. It is easy to verify that $\sum_{i=1}^d \gamma_i^*\gamma_i = \mathbb{1}_{\mathbb{C}^N}$ i.e. $(\gamma_i^*\gamma_i)_{i=1}^d$ is a POVM on $\mathbb{C}^N$.

Main Theorem (G. A. and D. Wright)

The asymptotic optimal lossless compression rate of a quantum stationary Markov source is equal to the dynamical entropy of the associated quantum dynamical system with respect to the POVM that were given in the above Construction.

Corollary

The asymptotic optimal lossless compression rate of an i.i.d. quantum source is equal to the von Neumann entropy of the average state produced.

References

[1] L. Accardi. The noncommutative Markovian property. *Func. Annal. Appl.*, 10:1, 1975.

[2] L. Accardi, M. Ohya, and N. Watanabe. Dynamical entropy through quantum Markov chains. *Open Syst. Inf. Dyn.*, 4:71, 1997.

[3] G. Bellomo, G.M. Bosyk, F. Holik, and S. Zozor. Loss- less quantum data compression with exponential penalization: An operational interpretation of the quantum Rényi entropy. *Sci. Rep.*, 7:14765, 2017.

[4] A. Kossakowski, M. Ohya, and N. Watanabe. Quantum dynamical entropy for completely positive maps. *Infin. Dimens. Anal. Quantum Prob. Relat. Top.*, 02:267, 1999.

[5] M. Müller and C. Rogers. Quantum bit strings and prefix-free Hilbert spaces. In *Proc. of the 2008 Intl. Conf. on Information Theory and Statistical Learning (Las Vegas)*. CSREA Press, 2008.

[6] B. Schumacher. Quantum coding. *Phys. Rev. A*, 51:2738, 1995.

[7] B. Schumacher and M.D. Westmoreland. Indeterminate length quantum coding. *Phys. Rev. A*, 51:042304, 2001.