

When is a Quantum Source Markov?

Emina Soljanin
Bell Labs, Lucent
Murray Hill NJ 07974, USA
Email: emina@lucent.com

Abstract—We consider a class of sources of quantum information simply defined based on classical information sources, and show that even for this limited class, certain basic information-theoretic questions are hard to answer.

I. INTRODUCTION

General definitions of quantum information sources involve the language of quantum statistical mechanics, C^* -algebras and dynamical systems, and are studied by the tools of these mathematical disciplines, see for example [1]–[3]. The literature on quantum sources is seldom accessible to information theorists (and perhaps even to quantum physicists). Moreover, although quantum counterparts of some important theorems (*e.g.*, Shannon-McMillan’s, Sanov’s) have been established, some other more basic questions (*e.g.*, computation of entropy, source memory) seem to be unaddressed. In this paper, we will consider a class of quantum sources that are in a straightforward manner defined based on classical sources, and show that even for this limited class, certain fundamental questions are hard to answer.

A discrete-time, finite-alphabet source of classical information produces a sequence of random variables taking values in a finite set called the *source alphabet*. We will be concerned with quantum systems which map classical source letters into *quantum states* for quantum transmission or storage. In the simplest case, quantum states correspond to unit length column vectors in a d -dimensional Hilbert space \mathcal{H}_d . Such quantum states are called *pure*. When $d = 2$, quantum states are called *qubits*. A column vector is denoted by $|\varphi\rangle$, its complex conjugate transpose by $\langle\varphi|$. A pure state is mathematically described by its *density matrix* equal to the outer product $|\varphi\rangle\langle\varphi|$. In a more complex case, all we know about a quantum state is that it is one of a finite number of possible pure states $|\varphi_i\rangle$ with probability p_i . Such quantum states are called *mixed*. A mixed state is also described by its density matrix which is equal to $\sum_i p_i |\varphi_i\rangle\langle\varphi_i|$. Note that a density matrix is a $d \times d$ Hermitian trace-one positive semidefinite

matrix. A classical analog to a mixed state can be a multi-faced coin which turns up as any of its faces with the corresponding probability. This paper will consider only sources of pure quantum states. We usually deal with source sequences rather than individual letters. The quantum state corresponding to a source sequence of length n has a $d^n \times d^n$ density matrix, equal to the tensor product of density matrices corresponding to the letters in the sequence.

Sources of pure quantum states whose underlying classical sources are memoryless are well understood [4],[5], but there are still some unanswered questions about their mixed-state counterparts that are a subject of current research [6]–[10]. This paper is concerned with sources whose underlying classical sources have memory. In Sec. II, we first review classical sources in a way that makes defining quantum sources more transparent, and then introduce quantum sources. For the rest of the paper, we are concerned with quantum sources whose underlying classical sources are Markov chains. We first consider the case when the quantum alphabet is composed of orthogonal states in Sec. III, and then move to the general case in Sec. IV. We discuss the nature of such quantum sources by looking into their n -th order density matrices, the n -th order entropies and the entropy rates, as well as the properties of classical sources induced by measurements of the quantum source outputs, which all indicate infinite memory when the quantum alphabet is composed of nonorthogonal states. This paper presents an extended outline of the general ideas presented in [13].

II. INFORMATION SOURCES

A. Classical Sources

A discrete-time, classical information source is modelled as a discrete-time, stochastic process, that is, a sequence $\{X_n\} = X_1, X_2, \dots, X_n, \dots$ of random variables. We are concerned with *finite alphabet* processes where each X_i takes values in a finite set \mathcal{A} . A subsequence $(X_m, X_{m+1}, \dots, X_k)$, $m < k$, will be

denoted by X_m^k . The process is characterized by the joint distributions

$$\Pr\{X_1^n = x_1^n\} = \mu_n(x_1^n), \quad \forall x_1^n \in \mathcal{A}^n \text{ and } n = 1, 2, \dots$$

where the n -th order joint distribution μ_n of the process is a probability distribution on \mathcal{A}^n . A sequence of probability distributions μ_n on \mathcal{A}^n , $n \geq 1$, constitutes the sequence of joint distributions of a process iff the following *consistency* condition holds for each $n \geq 1$:

$$\mu_n(a_1^n) = \sum_{a_{n+1} \in \mathcal{A}} \mu_{n+1}(a_1^{n+1}) \quad a_1^n \in \mathcal{A}^n \quad (1)$$

We shall consider only *stationary* processes, that is, processes whose joint distributions do not depend on the choice of time origin:

$$\Pr\{X_1^n = x_1^n\} = \Pr\{X_{1+\ell}^{n+\ell} = x_1^n\}$$

for every shift ℓ and for all $x_1^n \in \mathcal{A}^n$.

A process is *memoryless* if the X_i are independent random variables, *i.e.*,

$$\Pr\{X_n = x_n | X_1^{n-1} = x_1^{n-1}\} = \Pr\{X_n = x_n\}$$

The simplest examples of processes with memory are *Markov chains*. A sequence of random variables $\{X_n\}$ is a Markov chain if

$$\Pr\{X_n = x_n | X_1^{n-1} = x_1^{n-1}\} = \Pr\{X_n = x_n | X_{n-1} = x_{n-1}\}$$

holds for all $n \geq 2$ and all $x_1^n \in \mathcal{A}^n$. For a stationary Markov chain, *transition probabilities* $\Pr\{X_n = x_n | X_{n-1} = x_{n-1}\}$ do not depend on n . Thus, such process can be completely described by an $|\mathcal{A}| \times |\mathcal{A}|$ matrix P defined by

$$P_{ab} = P_{b|a} = \Pr\{X_n = a | X_{n-1} = b\}, \quad a, b \in \mathcal{A},$$

known as the *transition matrix* of the chain.

Example 1: Let $\mathcal{A} = \{0, 1\}$. A Markov chain with two states is illustrated in Fig. 1. The stationary distribution of the chain is given by $[\pi_0 \ \pi_1]$.

We now define one more type of process, known under various names (*e.g.*, finite-state process [11], hidden

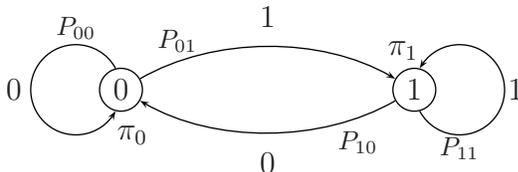


Fig. 1. A Markov chain with two states.

bution of the chain is given by $[\pi_0 \ \pi_1]$.

We now define one more type of process, known under various names (*e.g.*, finite-state process [11], hidden

Markov process [12]), that will be of interest in this paper. Let \mathcal{A} and \mathcal{S} be finite sets. An \mathcal{A} -valued process $\{Y\}$ is said to be a *finite-state process with state space* \mathcal{S} if

$$\Pr(S_n = s, Y_n = a | S_1^{n-1}, Y_1^{n-1}) = \Pr(S_n = s, Y_n = a | S_{n-1})$$

Note that while $\{Y_n\}$ is not a Markov chain, the pair $\{(S_n, Y_n)\}$ is; hence the name.

The n -th order entropy of the classical source $\{X\}$ is the Shannon entropy of X_1^n , defined as

$$H(X_1^n) = - \sum_{x_1^n \in \mathcal{A}^n} \mu_n(x_1^n) \log \mu_n(x_1^n),$$

and the process entropy rate as

$$H(X) = \lim_n \frac{1}{n} H(X_1^n).$$

The existence of the above limit for stationary processes follows from the subadditivity of the Shannon entropy:

$$\begin{aligned} H(X_1^{n+m}) &\leq H(X_1^n) + H(X_{n+1}^{n+m}) \\ &= H(X_1^n) + H(X_1^m), \end{aligned}$$

and the following result, known as the subadditivity lemma (see for example [11]).

Lemma 1: If $\{b_n\}$ is a sequence of nonnegative numbers which is subadditive, that is, $b_{n+m} \leq b_n + b_m$, then $\lim_n b_n$ exists and equals $\inf_n b_n/n$.

B. Quantum Sources

We have seen in Sec. II-A that a classical information source is described by a consistent sequence of probability distributions $\{\mu\} = \mu_1, \mu_2, \dots, \mu_n, \dots$, and the particular space on which random variables X_n are defined does not play any role. By analogy, we will model a discrete-time, quantum information source as a sequence $\{\rho\} = \rho_1, \rho_2, \dots, \rho_n, \dots$ of density matrices, that is, quantum counterparts to the probability distributions. When the quantum source alphabet consists of unit-length vectors in the d -dimensional Hilbert space \mathcal{H} , then ρ_n must be a Hermitian trace-one positive semidefinite matrix on the n -fold product $\mathcal{H}^{\otimes n}$. To preserve the notion of time, we will express this n -fold product as $\mathcal{H}^{\otimes n} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_n$. The only remaining condition the sequence of density matrices has to satisfy is (as in the classical case) the consistency condition:

$$\rho_n = \text{Tr}_{\mathcal{H}_{n+1}} \rho_{n+1}, \quad \forall n \geq 1, \quad (2)$$

where $\text{Tr}_{\mathcal{H}_{n+1}}$ denotes the partial trace over the space \mathcal{H}_n , and the operation is the quantum counterpart of the probability marginalization. It is interesting to compare the above expression with its classical counterpart (1).

The n -th order entropy of the quantum source $\{\rho\}$ is the von Neumann entropy of ρ_1^n , defined as

$$S(\rho^n) = -\text{Tr}(\rho_n \log \rho_n),$$

and the process entropy rate can be defined by analogy to the classical case as

$$S(\rho) = \lim_n \frac{1}{n} S(\rho^n).$$

The existence of the above limit for stationary processes follows from the subadditivity of the von Neumann entropy and the subadditivity lemma.

Suppose now that we have a classical information source described (as in Sec. II-A) by its set of consistent probability distributions $\{\mu_n | n \geq 1\}$, but have to use quantum pure states for transmission or storage. Thus each time the source outputs letter $a \in \mathcal{A}$, the quantum state $|\psi_a\rangle$ is produced. Each $|\psi_a\rangle$ is a unit-length vector in the d -dimensional Hilbert space \mathcal{H} , where we assume $d \geq |\mathcal{A}|$. At time n , the density matrix of the source ρ_n is given by

$$\rho_n = \sum_{x_1^n \in \mathcal{A}^n} \mu_n(x_1^n) |\psi_{x_1}\rangle\langle\psi_{x_1}| \otimes \cdots \otimes |\psi_{x_n}\rangle\langle\psi_{x_n}|.$$

It is easy to check that when μ_n are consistent probability distributions, then the corresponding ρ_n are consistent density matrices, and thus we effectively have a valid quantum source.

III. CHAINS WITH ORTHOGONAL STATES

For the rest of the paper, we will be concerned with the quantum source whose underlying source of classical information is the binary source described in Fig. 1. Each time the classical source outputs a 0, quantum state $|\psi_0\rangle$ is produced, and each time the source outputs a 1, quantum state $|\psi_1\rangle$ is produced. States $|\psi_0\rangle$ and $|\psi_1\rangle$ are unit-length vectors in the d -dimensional Hilbert space \mathcal{H} . Formally, the corresponding source of quantum information can be represented as shown in Fig. 2. Generalization

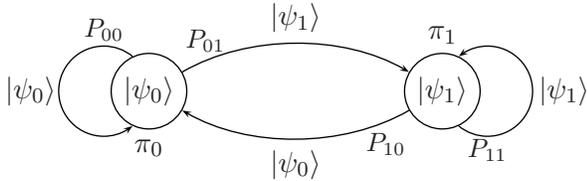


Fig. 2. A two-state source of quantum information.

to a non-binary case is straightforward.

We first consider the case when the states $|\psi_0\rangle$ and $|\psi_1\rangle$ are orthogonal. This source is essentially classical, and its von Neumann entropy rate is equal to the Shannon entropy rate of the underlying classical chain. We find an expression for the source density matrix of the form which allows easy comparison with the general nonorthogonal case.

A. The Density Matrix

We introduce the following diagonal matrix T to describe transitions in the chain:

$$T = P_{00}|\psi_0\rangle\langle\psi_0| \otimes |\psi_0\rangle\langle\psi_0| + P_{01}|\psi_0\rangle\langle\psi_0| \otimes |\psi_1\rangle\langle\psi_1| + P_{10}|\psi_1\rangle\langle\psi_1| \otimes |\psi_0\rangle\langle\psi_0| + P_{11}|\psi_1\rangle\langle\psi_1| \otimes |\psi_1\rangle\langle\psi_1|$$

Let I_1 be the identity matrix on the d -dimensional Hilbert space \mathcal{H} , and I_k the identity matrix on the k -fold product $\mathcal{H}^{\otimes k}$. It is easy to see that

$$\begin{aligned} \rho_1 &= \pi_0|\psi_0\rangle\langle\psi_0| + \pi_1|\psi_1\rangle\langle\psi_1| \\ \rho_2 &= (\rho_1 \otimes I_1)T \\ \rho_n &= (\rho_{n-1} \otimes I_1)(I_{n-2} \otimes T) \text{ for } n > 2. \end{aligned}$$

B. The Entropy

The n -th order von Neumann entropy of this quantum source is equal to the n -th order Shannon entropy of the underlying Markov chain. It is enough to only consider the third order entropy to indicate the difference with the nonorthogonal case considered in the next section:

$$\begin{aligned} S(\rho_3) &= H(X_1, X_2, X_3) \\ &= H(X_1, X_2) + H(X_2, X_3) - H(X_3) \end{aligned} \quad (3)$$

In general, when the random variables X_1, X_2, X_3 do not form a Markov chain, the equality in (3) has to be replaced with inequality, *i.e.*, $H(X_1, X_2, X_3) \leq H(X_1, X_2) + H(X_2, X_3) - H(X_3)$, which is known as the *strong subadditivity* of the Shannon entropy.

IV. CHAINS WITH NONORTHOGONAL STATES

We now consider the general case when the states $|\psi_0\rangle$ and $|\psi_1\rangle$ are not necessarily orthogonal.

A. The Density Matrix

When the classical source is in state 0, the density matrix of the quantum source at time $n = 1$ is

$$\rho_1^c(0) = P_{00}|\psi_0\rangle\langle\psi_0| + P_{01}|\psi_1\rangle\langle\psi_1|,$$

and when the classical source is in state 1, the density matrix of the quantum source at time $n = 1$ is

$$\rho_1^c(1) = P_{10}|\psi_0\rangle\langle\psi_0| + P_{11}|\psi_1\rangle\langle\psi_1|.$$

Similarly, the density matrices $\rho_n^c(0)$ and $\rho_n^c(1)$ at time n can be found. Note that matrix $\rho_1^c(i)$, $i = 1, 2$, (and similarly $\rho_n^c(i)$) consists of two terms, one corresponding to the sequences starting from i and ending in 0 and the other corresponding to the sequences starting from i and ending in 1.

We use the following matrix T to describe transitions in the chain:

$$T = \begin{bmatrix} P_{00}|\psi_0\rangle\langle\psi_0| & P_{01}|\psi_1\rangle\langle\psi_1| \\ P_{10}|\psi_0\rangle\langle\psi_0| & P_{11}|\psi_1\rangle\langle\psi_1| \end{bmatrix}.$$

Note that the i -th row of T corresponds to the two terms in $\rho_1^c(i)$. We define a family of matrices A_n to describe sequences produced by the source by the time n :

$$A_1 = T \\ A_n = (A_{n-1} \otimes I_1)(I_{n-1} \otimes T)$$

Note that the i -th row of A_n corresponds to the two terms in $\rho_n^c(i)$, $i = 1, 2$.

We can now express the source matrix ρ_n at time n in terms of matrix A_n :

$$\rho_n = ([\pi_0 \ \pi_1] \otimes I_n) A_n \begin{bmatrix} I_n \\ I_n \end{bmatrix}$$

It is interesting to note that in the nonorthogonal case, ρ_n cannot be expressed in terms of ρ_{n-1} , which indicates memory in the process.

B. The Entropy

Consider the third order von Neumann entropy of this quantum process $S(\rho_3)$. We introduce the notion of three quantum subsystems by denoting ρ_3 by ρ_{abc} . By the strong subadditivity of the van Neumann entropy,

$$S(\rho_{abc}) \leq S(\rho_{ab}) + S(\rho_{bc}) - S(\rho_b). \quad (4)$$

Only for quantum states with certain structure is the inequality (4) saturated [14], [15]. The quantum states ρ_n as defined above with nonorthogonal states $|\psi_0\rangle$ and $|\psi_1\rangle$ do not have that structure in general.

C. Measurement Induced Classical Processes

The simplest model of quantum measurement is known as the von Neumann's measurement. Mathematically, this type of measurement is defined by a set of pairwise orthogonal projection operators $\{\Pi_i\}$ which form a complete resolution of the identity, that is, $\sum_i \Pi_i = I$. For input $|\psi_j\rangle$, the classical output $\Pi_i|\psi_j\rangle$ happens with probability $|\langle\psi_j|\Pi_i|\psi_j\rangle|^2$. In a more general case, the pairwise orthogonal projection operators $\{\Pi_i\}$ are replaced by any positive-semidefinite operators $\{E_i\}$

which form a complete resolution of the identity. If each quantum state coming out of the quantum source $\{\rho\}$ (as described in Sec. II-B) is individually measured by employing a quantum measurement defined by operators $\{E_i\}$, the output is a sequence of random variables, *i.e.*, a classical source. We will say that this source is *induced by measurement* $\{E_i\}$ of quantum source $\{\rho\}$.

It is straightforward to show that the following holds:

Proposition 1: Let $\{\rho\}$ be a quantum process whose underlying classical process is a Markov chain $\{S\}$ over alphabet \mathcal{S} . Let $\{Y\}$ be a classical process induced by measurement $\{E_i\}$ of quantum source $\{\rho\}$. Then $\{Y\}$ is a Markov chain iff the quantum alphabet of $\{\rho\}$ consists of orthogonal states, otherwise $\{Y\}$ is a finite-state process with state space \mathcal{S} . The joint probability distributions of $\{Y\}$ are completely determined by the joint probability distributions of $\{S\}$, quantum alphabet of $\{\rho\}$, and measurement operators $\{E_i\}$.

REFERENCES

- [1] C. King and A. Leśniewski, "Quantum sources and a quantum coding theorem," *J. Math. Phys.* vol. 39, 1998, pp. 88–101.
- [2] D. Petz and M. Mosonyi, "Stationary quantum source coding," *arXiv:quant-ph/9912103*.
- [3] I. Bjelaković, T. Krüger, R. Siegmund-Schultze, and A. Szkola, "The Shannon-McMillan theorem for ergodic quantum lattice systems," *Invent. Math.*, vol. 155, pp. 203–222, June 2002.
- [4] B. W. Schumacher, "Quantum coding," *Physical Review A*, vol. 64, 2001.
- [5] B. W. Schumacher and R. Jozsa, "A new proof of the quantum noiseless coding theorem," *J. Mod. Opt.*, vol. 41, no. 12, pp. 2343–2349, 1994.
- [6] M. Horodecki, "Limits for compression of quantum information carried by ensembles of mixed states," *Phys. Rev. A*, vol. 61, 052309, 2001.
- [7] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, "On quantum coding ensembles of mixed states," *J. Phys. A: Math. Gen.*, vol. 34, pp. 6767–6785, 2001.
- [8] W. Dür, G. Vidal, and J. I. Cirac, "Visible compression of commuting mixed states," *Phys. Rev. A*, vol. 64, 022308, 2001.
- [9] E. Soljanin, "Compressing quantum mixed-state sources by sending classical information," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2263–2275, Aug. 2002.
- [10] M. Koashi and N. Imoto, "Compressibility of Mixed-State Signals," *arXiv:quant-ph/0103128*.
- [11] P. C. Shields, *The Ergodic Theory of Discrete Sample Paths*, (Graduate Studies in Mathematics, V. 13) American Mathematical Society, July 1996.
- [12] Y. Ephraim and N. Merhav, "Hidden Markov Processes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1518–1569, June 2002.
- [13] E. Soljanin, "On classical information sources with quantum outputs," *in preparation*.
- [14] M. B. Ruskai, "Inequalities for quantum entropy: a review with conditions for equality," *J. Math. Phys.*, vol. 43, pp. 4358–4375, 2002.
- [15] P. Hayden, R. Jozsa, D. Petz, and A. Winter, "Structure of states which satisfy strong subadditivity of quantum entropy with equality," *Commun. Math. Phys.*, pp. 359–374, 2004.