Communication over Quantum Channels with Parameter Estimation

Uzi Pereg
Department of Electrical and Computer Engineering
Technical University of Munich
Munich 80333, Germany.
Email: uzi.pereg@tum.de

Abstract—Communication over a random-parameter quantum channel when the decoder reconstructs the parameter sequence is considered in different scenarios. Regularized formulas are derived for the capacity-distortion regions with strictly-causal, causal, or non-causal channel side information (CSI) available at the encoder, and also without CSI. Single-letter characterizations are established in special cases. In particular, a single-letter formula is given for entanglement-breaking channels when CSI is not available. As a consequence, we obtain regularized formulas for the capacity of random-parameter quantum channels with CSI, generalizing previous results on classical-quantum channels.

Index Terms—Quantum information, Shannon theory, state estimation, rate-and-state channel, state information.

I. INTRODUCTION

A fundamental task in classical information theory is to determine the ultimate transmission rate of communication. Various scenarios are modeled by a channel that depends on a random parameter when there is strictly-causal, causal, or non-causal channel side information (CSI) available at the encoder [17]. The applications include cognitive radio in wireless systems, memory storage [18], and digital watermarking [5].

In the rate-and-state (RnS) model [26], the receiver is not only required to recover the message, but also to estimate the parameter sequence with distortion. For example, in digital multicast [26], the message represents digital control information that is multicast on top of an existing analog transmission, which is also estimated by the receiver. The capacity-distortion tradeoff region with strictly-causal CSI and with causal CSI was determined by Choudhuri et al. [7]. Inner and outer bounds on the tradeoff region with non-causal CSI were derived by Sutivong in [25], with a full characterization in the Gaussian case [26]. The RnS channel with feedback was recently considered by Bross and Lapidoth [4].

The field of quantum information is rapidly evolving in both practice and theory [9, 2, 30]. The Holevo-Schumacher-Westmoreland (HSW) Theorem provides a regularized (“multi-letter”) formula for the capacity of a quantum channel [15, 22]. Although calculation of such a formula is intractable in general, it provides computable lower bounds, and there are special cases where the capacity can be computed exactly. The reason for this difficulty is that the Holevo information is not necessarily additive [14]. Shor has demonstrated additivity for the class of entanglement-breaking channels [24], in which case the HSW theorem provides a single-letter computable capacity formula. This class includes both classical-quantum channels and measurement channels.

Boche et al. [3] addressed the classical-quantum channel with CSI at the encoder. The capacity was determined given causal CSI, and a regularized formula was provided given non-causal CSI. Warsi and Coon [27] used an information-spectrum approach to derive regularized bounds for rate-limited side information. The entanglement-assisted capacity of a quantum channel with non-causal CSI was determined by Dupuis [10], and with causal CSI by the author [21]. One-shot communication with CSI is considered in [1]. Luo and Devetak [19] considered channel simulation with source side information (SSI) at the decoder (see also [8, 6]). Parameter estimation of quantum channels is also studied from the algorithmic point of view in different settings [12, 31].

In this paper, we consider a random-parameter quantum channel when the decoder reconstructs the parameter sequence with limited distortion. We study settings where either strictly-causal, causal, or non-causal channel side information (CSI) is available at the encoder, and also when CSI is not available. This model can be viewed as the quantum analog of the classical RnS channel. We derive regularized formulas for the capacity-distortion tradeoff regions. Single-letter characterizations are established in several special cases, and in particular, for entanglement-breaking channels when CSI is not available. As a consequence, we obtain regularized formulas for the capacity of random-parameter quantum channels with strictly-causal, causal, or non-causal CSI, generalizing previous results by Boche et al. [3] on classical-quantum channels.

To prove achievability with strictly-causal CSI, we extend the classical block Markov coding method in [7] to the quantum setting, and then apply the quantum packing lemma [16] for decoding the message, and the classical covering lemma for the reconstruction of the parameter sequence. The gentle measurement lemma (GML) [29] alleviates the proof, as it guarantees that multiple decoding measurements can be performed with negligible disturbance. Achievability with causal CSI is proved by introducing a quantum “Shannon-strategy” encoding operation [23]. As for non-causal CSI, we extend the classical binning technique [13] to the quantum setting. Considering entanglement-breaking channels without CSI, we use a different approach from that of Shor [24]. As
opposed to Shor [24], we do not show additivity of the capacity formula, but rather apply and generalize inequalities that are borrowed from Shor’s derivation [24] [28, Chapter 13] to prove the converse part in a more direct manner. We refer to the resulting inequality as the generalized Shor inequality. The full manuscript with proofs can be found in [20].

II. DEFINITIONS AND RELATED WORK

We begin with basic definitions.

A. Notation, States, and Information Measures

The state of a quantum system $A$ is given by a density operator $\rho$ on the Hilbert space $\mathcal{H}_A$. A density operator is an Hermitian, positive semidefinite operator, with unit trace, i.e., $\rho^\dagger = \rho$, $\rho \succeq 0$, and $\text{Tr}(\rho) = 1$. A measurement of a quantum system is any set of operators $\{\Lambda_j\}$ that forms a positive-operator-valued measure (POVM), i.e., the operators are positive semi-definite and $\sum_j \Lambda_j = 1$, where $1$ is the identity operator. According to the Born rule, if the system is in state $\rho$, then the probability of the measurement outcome $j$ is given by $P_A(j) = \text{Tr}(\Lambda_j \rho)$. Define the quantum entropy by $H(\rho) = -\sum_i p_i \log(p_i)$ where $p_i = \text{Tr}(\Lambda_i \rho)$. We assume that both the random parameter state and the quantum system is any set of operators $\{\Lambda_j\}$ corresponding to a quantum physical evolution. The channel $\mathcal{N}_{SA\rightarrow B}$ is a linear, completely positive, trace preserving map $\mathcal{N}_{SA\rightarrow B}$. A pair $(\mathcal{E} , \Lambda)$ is called achievable if for every state $\rho_{AE}$ with arbitrary $E$, the output state $(1 \otimes \mathcal{N}(s)) (\rho_{AE})$ is separable for all $s \in \mathcal{S}$.

A measurement channel (or, q-c channel) $\mathcal{M}_{A\rightarrow Y}$ has the form $\mathcal{M}(\rho_A) = \sum_{y \in Y} \text{Tr}(\Lambda_y \rho_A) |y\rangle \langle y|$ for some POVM $\{\Lambda_y\}$ and orthonormal $\{|y\rangle\}$. We denote the random-parameter measurement channel by $\mathcal{M}_{SA\rightarrow Y}$ to distinguish it from the general channel $\mathcal{N}_{SA\rightarrow B}$. Both classical-quantum and measurement channels are entanglement breaking.

C. Coding

We define a code to transmit classical information. With strictly-causal CSI, Alice has the sequence of past random parameters, $S_1, \ldots, S_i$, at time $i \in [1 : n]$. Let $d : \mathcal{S} \times \hat{\mathcal{S}} \rightarrow [0, \infty)$ be a bounded distortion function. Denote the average distortion between a parameter sequence $s^n$ and a reconstruction sequence $\hat{s}^n$ by

$$d^n(s^n, \hat{s}^n) \triangleq \frac{1}{n} \sum_{i=1}^n d(s_i, \hat{s}_i).$$

Definition 1. A $(2^nR, n)$ code with strictly-causal CSI at the encoder consists of the following: a message set $[1 : 2^nR]$, where $2^nR$ is assumed to be an integer, a sequence of encoding maps $\mathcal{E}_{M,S^{i-1} \rightarrow A^i}$, for $i \in [1 : n]$, and a decoding POVM $\{\Lambda_{B^n}\}_{m \in [1 : 2^nR], s^n \in \mathcal{S}^n}$. The encoding maps must be consistent in the sense that the states $\rho_{A^n,s^{n-1}} = \mathcal{E}(i)(m, s_i)$ satisfy $\text{Tr}_{A^n}(\rho_{A^n,s^{n-1}}) = \rho_{A^n,s^{n-1}}$. Denote the code by $(\mathcal{E}, \Lambda)$.

The communication scheme is depicted in Figure 1. Alice chooses a classical message $m \in [1 : 2^nR]$. At time $i \in [1 : n]$, given the sequence $s^{i-1} \in \mathcal{S}^{i-1}$, she prepares the state $\rho_{A^n,i} = F_{S^n}(m, s^n)$ and transmits the system $A^n$ over the channel $\mathcal{N}_{SA\rightarrow B}$.

Bob receives the channel output systems $B^n$ and performs the POVM $\{\Lambda_{B^n}\}$. The conditional probability of error, given that the message $m$ was sent, is given by

$$P_{e|m}(\mathcal{E}, \Lambda) = \text{Tr} \left[ (1 - \sum_{s^n \in \mathcal{S}^n} \Lambda_{B^n}^{m,s^n}) \sum_{s^n \in \mathcal{S}^n} q^n(s^n) \mathcal{N}_{A^n \rightarrow B^n}(\rho_{A^n,s^{n-1}}) \right].$$

The average distortion for the code $(\mathcal{E}, \Lambda)$ is

$$\Delta^{(n)}(\mathcal{E}, \Lambda) \triangleq \sum_{s^n \in \mathcal{S}^n} \sum_{\hat{s}^n \in \mathcal{S}^n} d^n(s^n, \hat{s}^n) q^n(s^n) \cdot \frac{1}{2^nR} \times 2^nR 2^nR \sum_{m=1}^{2^nR} \sum_{\hat{m}=1}^{2^nR} \text{Tr} \left[ \Lambda_{B^n}^{m,s^n} \mathcal{N}_{A^n \rightarrow B^n}(\rho_{A^n,s^{n-1}}) \right].$$

A $(2^nR, n, \varepsilon, D)$ rate-distortion code satisfies $P_{e|m}(\rho_{A^n}, \Lambda_{B^n}) \leq \varepsilon$ for all $m \in [1 : 2^nR]$, and $\Delta^{(n)}(\rho_{A^n}, \Lambda_{B^n}) \leq D$. A pair $(R, D)$ is called achievable if for every $\varepsilon > 0$ and sufficiently large $n$, there exists a $(2^nR, n, \varepsilon, D)$ code. The capacity-distortion region $\mathcal{C}_{\varepsilon}(\mathcal{N})$ is defined as the set of achievable pairs $(R, D)$. 

1819
Alternatively, one may fix $D > 0$ and define the capacity-distortion function $C_{c,c}(N, D)$ as the supremum of achievable rates. Note that $C_{c,c}(N, d_{\text{max}})$ is the standard capacity of a quantum channel, without parameter estimation.

We also address the causal and the non-causal setting. In the causal setting, Alice has the present parameter value $S_i$ as well, and sends $\rho_{A_i}^{m_i,s_i} = \hat{\mathcal{E}}^{(i)}_{M_S \rightarrow A_i}(m_i, s_i)$. Whereas, in the non-causal setting, Alice has the sequence $S^n$ a priori, and can thus send any sequence $\rho_{A^n}^{m^n,s^n}$. We use the subscripts ‘s-c’, ‘causal’, or ‘n-c’ to indicate whether CSI is available at the encoder in a strictly-causal, causal, or non-causal manner, respectively. The notation is summarized in the table below.

| $C(N)$ | $C_{c,c}(N)$ | $C_{\text{caus}}(N)$ | $C_{\text{n-c}}(N)$ |
|--------|-------------|----------------------|-------------------|
| $C(N)$ | $C_{c,c}(N)$ | $C_{\text{caus}}(N)$ | $C_{\text{n-c}}(N)$ |
| $C(N, D)$ | $C_{c,c}(N, D)$ | $C_{\text{caus}}(N, D)$ | $C_{\text{n-c}}(N, D)$ |

Fig. 2. Notation of channel capacity-distortion regions and functions with and without CSI. The notation of the capacity-distortion regions is given in the first row, and of the capacity-distortion functions in the second row.

D. Related Work

We briefly review known results for a quantum channel that does not depend on a random parameter and has no distortion constraint, i.e. $N_{A^n \rightarrow B}^{(s)} = \mathcal{P}_{A^n \rightarrow B} \forall s$ and $D = d_{\text{max}}$. Define

$$C(\mathcal{P}, d_{\text{max}}) = \max_{\mathcal{I}(X; B)} \frac{I(X; B)}{d_{\text{max}}}$$

with $\rho_{X^n} = \sum_{x \in X} p_X(x) \ketbra{x} \otimes \mathcal{P}(\ketbra{x}^A \otimes \ketbra{x}^B)$. Next, we cite the HSW Theorem, which provides a regularized capacity formula for a quantum channel without parameters or distortion requirements.

**Theorem 1** (see [15, 22, 24]). The capacity of a quantum channel $\mathcal{P}_{A^n \rightarrow B}$ without parameters is given by

$$C(\mathcal{P}, d_{\text{max}}) = \lim_{k \to \infty} \frac{1}{k} \mathcal{C}(\mathcal{P}^\otimes k, d_{\text{max}}).$$

Furthermore, if $\mathcal{P}_{A^n \rightarrow B}$ is entanglement-breaking, then $C(\mathcal{P}, d_{\text{max}}) = C(\mathcal{P}, d_{\text{max}})$.

**Remark 1.** We note that the setting of a random-parameter quantum channel $N_{A^n \rightarrow B}$ without side information and with $D = d_{\text{max}}$ is equivalent to that of a channel that does not depend on a random parameter, with $\mathcal{P}_{A^n \rightarrow B} = \sum_{s \in S} q(s)N^{(s)}_{A^n \rightarrow B}$.

III. MAIN RESULTS

We state our results on the random-parameter quantum channel $N_{A^n \rightarrow B}$ with and without CSI at the encoder.

A. In the Absence of Side Information

Consider the case where Alice does not have access to the parameter sequence. The achievability proof of a regularized formula in this case is straightforward. We also obtain a single letter formula for entanglement-breaking channels. While our proof is based on observations from [24], our approach is different. We do not show additivity, but rather apply and generalize inequalities from [24], which we refer to as Shor inequalities, in order to prove the converse part.

**Lemma 2** (Generalized Shor Inequality). Let $N_{A^n \rightarrow B}$ be an entanglement-breaking random-parameter channel, and let $n \geq 2$. Consider the classical-quantum states $\rho_{X^n} = \sum_{x \in X} p_X(x) \ketbra{x} \otimes \rho_{X^n A^n}^{(s)}$, and $\rho_{X^n A^n}^{(s)} \equiv \sum_{s \in S^n} \rho_{X^n}^{(s)} \mathbb{I} \otimes N^{(s)}_{A^n \rightarrow B}$. Then, there exists a classical-classical-quantum extension $\rho_{X^n Y^{n-1} B}$ such that

$$I(X; B^n) \leq \sum_{i=1}^n I(X, Y^{i-1}; B_i) \rho.$$

We prove Lemma 2 in [20]. Observe that given the generalized Shor inequality, a single-letter converse for Theorem 1 is straightforward. Indeed, by Fano’s inequality, we have $R - \varepsilon_n \leq \frac{1}{n} I(M; B^n) \rho$, where $\varepsilon_n$ tends to zero as $n \to \infty$. Based on our generalized Shor inequality, if the channel is entanglement breaking, then $R - \varepsilon_n \leq \frac{1}{n} \sum_{i=1}^n I(M, Y^{i-1}; B_i)$. Since the sequence $Y^{n-1}$ is classical, the rate is bounded by the Holevo information.

Next, we give our capacity-distortion theorem. Define

$$\mathcal{R}(N) \triangleq \left\{ \left( R, D \right) : R \leq I(X; B) \rho, D \geq \sum_{s \in S^n} q(s) p_X(x) \text{Tr}(\Gamma^s_{B|x} \rho_{X^n}) d(s, \tilde{s}) \right\}$$

where the union is over the set of all distributions $p_X(x)$, state collection $\{\phi_A^s\}$, and set of POVMs $\{\Gamma^s_{B|x}\}$, with $\rho_{X^n}^{(s)} = N^{(s)}_{A^n \rightarrow B}(\phi_A^s \otimes \phi_A^s)$. We prove Theorem 3 in [20].

**Theorem 3**. The capacity-distortion region of a random-parameter quantum channel $N_{A^n \rightarrow B}$ without CSI is given by

$$\mathcal{C}(N) = \bigcup_{k=1}^\infty \frac{1}{k} \mathcal{R}(N^\otimes k).$$

Furthermore, if $N_{A^n \rightarrow B}$ is entanglement-breaking, then $\mathcal{C}(N) = \mathcal{R}(N)$.

The proof of Theorem 3 is given in [20].

B. Strictly-Causal Side Information

We continue to our main result on the random-parameter quantum channel with strictly-causal CSI. Define the rate-distortion region

$$\mathcal{R}_{c,c}(N) \triangleq \left\{ \left( R, D \right) : R \leq I(Z, X; B) \rho - I(Z; S|X) \right\} \cup \left\{ D \geq \sum_{s \in S^n, x, z} q(s) p_X(x) p_{Z|X,S}(z|x, s) \text{Tr}(\Gamma^s_{B|x} N^{(s)}_{A^n \rightarrow B}(\phi_A^s \otimes \phi_A^s)) d(s, \tilde{s}) \right\}$$

where the union is over the set of all distributions $p_X(x)p_{Z|X,S}(z|x, s)$, state collection $\{\phi_A^s\}$, and set of POVMs $\{\Gamma^s_{B|x}\}$, with $\rho_{S+n Z+X B} = \sum_{s, x, z} q(s) p_X(x) p_{Z|X,S}(z|x, s) \ket{a} \otimes \ket{a} \otimes \ket{a} \otimes \ket{a} \otimes a \otimes z \otimes z | x \rangle \langle a | \otimes | a \rangle \langle a | \otimes N^{(s)}_{A^n \rightarrow B}(\phi_A^s \otimes \phi_A^s).$
Our main result is given below.

Theorem 4. The capacity-distortion region of a random-parameter quantum channel $\mathcal{N}_{SA \rightarrow B}$ with strictly-causal CSI at the encoder is given by

$$C_{n-c}(\mathcal{N}) = \bigcup_{k=1}^{\infty} \frac{1}{k} R_{n-c}(\mathcal{N}^\otimes k).$$

Furthermore, for a random-parameter measurement channel, we have $C_{c}(\mathcal{N}) = R_{n-c}(\mathcal{M})$.

The outline of the achievability proof is given in Section IV. The full proof for both the direct and converse parts can be found in [20]. To prove achievability, we extend the classical block Markov coding technique to the quantum setting, and then apply the quantum packing lemma for decoding the message and the classical covering lemma for the reconstruction of the parameter sequence. The GML [29] alleviates the proof, as it guarantees that multiple decoding measurements can be performed without “destroying” the quantum state, i.e. such that the output state after each measurement is almost the same. A formula for the capacity-distortion function $C(n-c)(\mathcal{N})$ immediately follows as the regularization of $C_{n-c}(\mathcal{N})$ immediately follows from Theorem 6 as the distortion constraint is inactive for $D = d_{\text{max}}$.

Theorem 6. The capacity-distortion region of the random-parameter quantum channel $\mathcal{N}_{SA \rightarrow B}$ with non-causal CSI at the encoder is given by

$$C_{n-c}(\mathcal{N}) = \bigcup_{k=1}^{\infty} \frac{1}{k} R_{n-c}(\mathcal{N}^\otimes k).$$

The proof of Theorem 6 is given in [20]. To prove achievability, we use an extension of the classical binning technique [13] to the quantum setting, and then apply the quantum packing lemma and the classical covering lemma. We note that even for a classical channel, a single-letter characterization for non-causal CSI is an open problem [4].

A formula for the capacity-distortion function $C_{n-c}(\mathcal{N}, D)$ immediately follows from the regularization of $C_{n-c}(\mathcal{N})$ as $\max[ I( X; B )_\rho - I( X; S | X ) ]$ (see [20]).

E. Without Parameter Estimation

As mentioned, the case $D = d_{\text{max}}$ yields the standard capacity of a random-parameter channel. The following results are direct consequences of the theorems above and they generalize previous results by Boche et al. [3].

Corollary 7. The capacity of a random-parameter quantum channel $\mathcal{N}_{SA \rightarrow B}$ with strictly-causal CSI at the encoder is the same as without CSI.

The direct part is immediate, since the encoder can ignore the CSI, while the converse part holds by convexity arguments.

Next, we consider causal CSI.

Corollary 8. The capacity of a random-parameter quantum channel $\mathcal{N}_{SA \rightarrow B}$ with causal CSI at the encoder is given by

$$C_{\text{caus}}(\mathcal{N}, d_{\text{max}}) = \lim_{k \to \infty} \frac{1}{k} \sup \{ I( X_k; B_k^k )_\rho \}$$

with $\rho_{S^kX^kB^k} = \sum_{s^k \in \mathcal{S}^k} p_{s^k} ( s^k ) p_{X^k}( x^k ) | s^k \rangle \langle s^k | \otimes | x^k \rangle \langle x^k |$ and

$$N_{A^k \rightarrow B^k} = \left( \mathcal{N}_{s, c}(\mathcal{N}^s_{s, c})^{A^k} (\phi_{A^k}^{s, c}) \right).$$

The direct part follows by taking $Z = \emptyset$, and the converse part holds by convexity arguments.

We move to non-causal CSI.

Corollary 9. The capacity of a random-parameter quantum channel $\mathcal{N}_{SA \rightarrow B}$ with non-causal CSI at the encoder is given by

$$C_{\text{n-c}}(\mathcal{N}, d_{\text{max}}) = \lim_{k \to \infty} \frac{1}{k} \sup \{ I( X_k; B_k^k )_\rho - I( X_k; S_k^k ) \}$$

with $\rho_{S^kX^kB^k} = \sum_{s^k \in \mathcal{S}^k} p_{s^k} ( s^k ) p_{X^k}( x^k ) | s^k \rangle \langle s^k | \otimes | x^k \rangle \langle x^k |$ and

$$N_{A^k \rightarrow B^k} = \left( \mathcal{N}_{s, c}(\mathcal{N}^s_{s, c})^{A^k} (\phi_{A^k}^{s, c}) \right).$$

The statement above immediately follows from Theorem 6 as the distortion constraint is inactive for $D = d_{\text{max}}$. 

1821
IV. Proof Outline of Theorem 4

We only give an outline for the achievability proof. The details can be found in [20]. We show that for every \( \ell_0, \varepsilon_1, D + \delta_0 > 0 \), there exists a \( (2^{2n(R_n - \delta_0)}, n, \varepsilon_1, D + \delta_0) \) code provided that \((R, D) \in \mathcal{R}_{NC}(N)\). Let \( \{p_X(x)p_Z|X(z|x), \phi^Z_A\} \) be a given ensemble, and define

\[
\rho^*_A = \sum_{z \in Z} p_Z(z|x) \phi^Z_A, \quad \rho^*_B = \sum_{s \in S} q(s) \mathcal{N}^*(s)(\rho^*_A) . \tag{16}
\]

We use \( T \) transmission blocks, each consists of \( n \) input systems. Given strictly-causal CSI, Alice knows the parameter sequences in the previous blocks. In effect, the \( j \)th transmission block encodes a message \( m_j \) and a compression of the parameter sequence \( s^n_j-1 \) from the previous block. The code construction and coding operations are described below.

1) Classical Code Construction: Let \( 0 < R_s < R_d \). For every \( j \in [2 : T] \), select \( 2^{nR(R_n + R_s)} \) independent sequences \( z^n_j(k_j|m_j, \ell_j-1) \), \( m_j \in [1 : 2^{nR_d}], \ell_j-1 \in [1 : 2^{nR_s}] \), at random according to \( \prod_{j=1}^n p_X(z_i) \). For every \( m_j \) and \( \ell_j-1 \), select \( 2^{nR_s} \) conditionally independent sequences \( z^n_j(k_j|m_j, \ell_j-1) \), \( k_j \in [1 : 2^{nR_d}] \), at random according to \( \prod_{j=1}^n p_Z(z_j|x_j,m_j,\ell_j-1) \). For \( j = 1 \), set \( \ell_0 \equiv 1 \) and select codewords in the same manner. Partition the index set \([1 : 2^{nR_d}]\) into bins \( \mathcal{K}(\ell_j) \) of equal size \( 2^{nR_s} \).

2) Encoding and Decoding: To send the messages \( (m_j) \), given \((s^n_1, \ldots, s^n_{j-1})\), Alice performs the following.

(i) At the end of block \( j \), find an index \( k_j \in [1 : 2^{nR_s}] \) such that \((s^n_j, z^n_j(k_j|m_j, \ell_j-1), x^n_j(m_j, \ell_j-1)) \in \mathcal{A}^n(p_{S,X,Z}) \), where \( p_{S,X,Z}(s,x,z) = q(s)p_X(x)p_Z(z|x,s) \). Set \( \ell_j \) to be the bin index of \( k_j \), i.e., such that \( k_j \in \mathcal{K}(\ell_j) \).

(ii) In block \( j + 1 \), send \( \rho_A^n_{j+1} = \bigotimes_{i=1}^n \rho_A^{x_{j+1},i}(m_{j+1},\ell_j) \). Bob receives the systems \( B^n_1, \ldots, B^n_T \) at state

\[
\rho_{BT^n} = \bigotimes_{j=1}^T \bigotimes_{i=1}^n \rho_{B_j}^{x_{j+1},i}(m_{j+1},\ell_j) \tag{17}
\]

and decodes as follows.

(i) At the end of block \( j + 1 \), decode \((\hat{m}_{j+1}, \hat{\ell}_j)\) by applying a POVM \( \{\Lambda^1_{m_{j+1},\ell_j}\}_{(m_{j+1},\ell_j) \in [1:2^n] \times [1:2^{nR_d}]} \).

(ii) Decode \( \hat{k}_j \) by applying a second POVM \( \{\Lambda^2_{k_j|x^n(m_{j+1},\ell_j)}\}_{k_j \in \mathcal{K}(\ell_j)} \).

(iii) Reconstruct the parameter sequence by applying the POVM \( \Gamma_{y^n_j}^{x_{j+1},i}(m_{j+1},\ell_j-1),z^n_j(k_j|m_{j+1},\ell_j-1) \) for \( j \in [1 : T] \), \( i \in [1 : n] \).

3) Analysis of Probability of Error and Distortion: By symmetry, we may assume w.l.o.g. that \( M_j = L_j = L_j-1 = 1 \). Consider the following events,

\[
\begin{align*}
\mathcal{E}_{1,j} &= \{ (S^n, X^n(1,1), Z^n(k_j|1,1)) \notin \mathcal{A}^n(p_{S,X,Z}) \forall k_j \} \\
\mathcal{E}_{2,j} &= \{ (\hat{M}_j, \hat{L}_j-1) \neq (1,1) \}, \quad \mathcal{E}_{3,j} = \{ \hat{K}_j \neq K_j \} \\
\mathcal{E}_{4,j} &= \{ d^n(S^n_j, \hat{S}^n_j) > D + \frac{\delta_0}{2} \}. \tag{18}
\end{align*}
\]

By the union of events bound, the probability of error is bounded by

\[
P_e^{(T_n)}(\varepsilon_1) = \sum_{j=1}^T \Pr(\mathcal{E}_{1,j}) + \sum_{j=0}^{T-1} \Pr(\mathcal{E}_{2,j+1} | \mathcal{E}_{1,j} \cap \mathcal{E}_{2,j}^c) + \sum_{j=0}^{T-1} \Pr(\mathcal{E}_{3,j+1} | \mathcal{E}_{1,j} \cap \mathcal{E}_{2,j}^c) + \sum_{j=0}^{T-1} \Pr(\mathcal{E}_{4,j+1} | \mathcal{E}_{1,j} \cap \mathcal{E}_{2,j}^c \cap \mathcal{E}_{3,j}^c \cap \mathcal{E}_{4,j}^c) \tag{19}
\]

Based on the classical covering lemma, \( \Pr(\mathcal{E}_{1,j}) \) tends to zero as \( n \to \infty \) for

\[
\tilde{R}_s > I(X;Z) + \varepsilon_1 = I(Z;S|X) + \varepsilon_1 \tag{20}
\]

where \( \varepsilon_1 \) is arbitrarily small. By applying the quantum packing lemma with appropriate typical space projectors, we show that there exists a POVM \( \Lambda^1_{m_{j+1},\ell_j} \) such that \( \Pr(\mathcal{E}_{2,j+1} | \mathcal{E}_{1,j} \cap \mathcal{E}_{2,j}^c) \leq 2^{-n(I(X;B)_{\rho} - (R_s - \frac{\delta_0}{2}) - \varepsilon_2)} \), which tends to zero for

\[
R < I(X;B)_{\rho} - R_s - \varepsilon_2 . \tag{21}
\]

Once more, the GML along with implications of the quantum packing lemma, we claim that given \( \mathcal{E}_{2,j+1} \), the post-measurement state \( \rho_{BT_j}^n \) is arbitrarily close to the original \( \rho_{BT_j}^n \) for large \( n \) and rates as in (21). Therefore, the distribution of measurement outcomes when \( \rho_{BT_j}^n \) is measured is roughly the same as if the POVM \( \Lambda^1_{m_{j+1},\ell_j} \) was never performed. Using the quantum packing lemma again with conditional typical space projectors, we show that there exists a POVM \( \Lambda^2_{k_j|x^n} \) such that the third sum tends to zero if

\[
R_s > \tilde{R}_s - I(Z;S|X) + \varepsilon_4 . \tag{22}
\]

ACKNOWLEDGMENTS

The author thanks Roberto Ferrara (Technical University of Munich) for useful discussions.

This work was supported by Bundesministerium für Bildung und Forschung (BMBF) through Grant 16KIS0856 and the Viterbi scholarship of the Technion.

REFERENCES

[1] A. Anshu, R. Jain, and N. A. Warsi. “On the near-optimality of one-shot classical communication over quantum channels”. J. Math. Phys. 60.1 (2019), p. 012204.
C. H. Bennett and G. Brassard. “Quantum cryptography: public key distribution and coin tossing.” Theor. Comput. Sci. 560.12 (2014), pp. 7–11.

H. Boche, N. Cai, and J. Nötzel. “The classical-quantum channel with random state parameters known to the sender”. J. Physics A: Math. and Theor. 49.19 (Apr. 2016), p. 195302.

S. I. Bross and A. Lapidoth. “The Rate-and-State Capacity with Feedback”. IEEE Trans. Inf. Theory 64.3 (Mar. 2018), pp. 1893–1918.

B. Chen and G. W. Wornell. “Quantization index modulation: A class of provably good methods for digital watermarking and information embedding”. IEEE Trans. Inf. Theory 47.4 (May 2001), pp. 1423–1443.

H. C. Cheng, E. P. Hanson, N. Datta, and M. H. Hsieh. “Duality between source coding with quantum side information and cq channel coding”. Proc. IEEE Int. Symp. Inf. Theory (ISIT’2019). Paris, France, July 2019, pp. 1142–1146.

C. Choudhuri, Y. H. Kim, and U. Mitra. “Causal state communication”. IEEE Trans. Inf. Theory 59.6 (June 2013), pp. 3709–3719.

I. Devetak and A. Winter. “Classical data compression with quantum side information”. Phys. Rev. A 68 (4 Oct. 2003), p. 042301.

J. P. Dowling and G. J. Milburn. “Quantum technology: the second quantum revolution”. Philos. Trans. Royal Soc. London. Series A: Math., Phys. and Eng. Sciences 361.1809 (2003), pp. 1655–1674.

F. Dupuis. “The capacity of quantum channels with side information at the transmitter”. Proc. IEEE Int. Symp. Inf. Theory (ISIT’2009). June 2009, pp. 948–952.

A. El Gamal and Y.H. Kim. Network Information Theory. Cambridge University Press, 2011.

A. Fujiwara. “Quantum channel identification problem”. Asympt. Theory Quantum Statist. Infer.: Selected Papers. World Scientific, 2005, pp. 487–493.

S. I. Gel’fand and M. S. Pinsker. “Coding for channel with random parameters”. Probl. Control Inform. Theory 9.1 (Jan. 1980), pp. 19–31.

M. B. Hastings. “Superadditivity of communication capacity using entangled inputs”. Nature Physics 5.4 (Mar. 2009), p. 255.

A. S. Holevo. “The capacity of the quantum channel with general signal states”. IEEE Trans. Inf. Theory 44.1 (Jan. 1998), pp. 269–273.

M. Hsieh, I. Devetak, and A. Winter. “Entanglement-assisted capacity of quantum multiple-access channels”. IEEE Trans. Inf. Theory 54.7 (July 2008), pp. 3078–3090.

G. Keshet, Y. Steinberg, and N. Merhav. “Channel coding in the presence of side information”. Foundations and Trends in Communications and Information Theory 4.6 (Jan. 2007), pp. 445–586.

A. V. Kuznetsov and B. S. Tsybakov. “Coding in a memory with defective cells”. Problemy peredachi informatsii 10.2 (1974), pp. 52–60.

Z. Luo and I. Devetak. “Channel Simulation With Quantum Side Information”. IEEE Trans. Inf. Theory 55.3 (Mar. 2009), pp. 1331–1342.

U. Peregr. “Communication over Quantum Channels with Parameter Estimation”, arXiv:2001.00836 (Jan. 2020). URL: https://arxiv.org/pdf/2001.00836.pdf.

U. Peregr. “Entanglement-assisted capacity of quantum channels with side information”. Int. Zürich Seminar Inf. Commun. (IZS’2020). Zürich, Switzerland, Feb. 2020.

B. Schumacher and M. D. Westmoreland. “Sending classical information via noisy quantum channels”. Phys. Rev. A 56.1 (July 1997), p. 131.

C. E. Shannon. “Channels with side information at the transmitter”. IBM J. Res. Dev. 2.4 (Oct. 1958), pp. 289–293.

P. W. Shor. “Additivity of the classical capacity of entanglement-breaking quantum channels”. J. Math. Phys. 43.9 (May 2002), pp. 4334–4340.

A. Sutivong. “Channel capacity and state estimation for state-dependent channel”. Ph.D. Dissertation, Palo Alto, CA, USA. Stanford University, 2003.

A. Sutivong, M. Chiang, T. M. Cover, and Y. H. Kim. “Channel capacity and state estimation for state-dependent Gaussian channels”. IEEE Trans. Inf. Theory 51.4 (Apr. 2005), pp. 1486–1495.

N. A. Warsi and J. P. Coon. “Coding for classical-quantum channels with rate limited side information at the encoder: information-spectrum approach”. IEEE Trans. Inf. Theory 63.5 (May 2017), pp. 3322–3331.

M. Wilde. Quantum information theory. 2nd ed. Cambridge University Press, 2017.

A. Winter. “Coding theorem and strong converse for quantum channels”. IEEE Trans. Inf. Theory 45.7 (Nov. 1999), pp. 2481–2485.

W. Zhang, D. S. Ding, Y. B. Sheng, L. Zhou, B. S. Shi, and G. C. Guo. “Quantum secure direct communication with quantum memory”. Phys. Rev. Lett. 118.22 (2017), p. 220501.

M. Zorzi, F. Ticozzi, and A. Ferrante. “On quantum channel estimation with minimal resources”. arXiv:1106.2105 (2011).