

Coding for classical-quantum channels with rate limited side information at the encoder: An information-spectrum approach

Naqeeb Ahmad Warsi, Justin P. Coon, *Senior Member, IEEE*

Abstract—We study the hybrid classical-quantum version of the channel coding problem for the famous Gel’fand-Pinsker channel. In the classical setting for this channel the conditional distribution of the channel output given the channel input is a function of a random parameter called the channel state. We study this problem when a rate limited version of the channel state is available at the encoder for the classical-quantum Gel’fand-Pinsker channel. We establish the capacity region for this problem in the information-spectrum setting. The capacity region is quantified in terms of spectral-sup classical mutual information rate and spectral-inf quantum mutual information rate.

I. INTRODUCTION

In traditional information theory literature it is common to study the underlying problems assuming that the channel characteristics do not change over multiple use. The proofs appeal to *typicality* of sequences or typical subspaces in the quantum setting [1], [2]: the empirical distribution of symbols in a long sequence of trials will with high probability be close to the true distribution. However, information theoretic arguments based on typicality or the related Asymptotic Equipartition Property (AEP) assume that both the source and channel are stationary and/or ergodic (memoryless), assumptions that are not always valid, for example, in [3] Gray analyzes the details of asymptotically mean stationary sources, which are neither stationary nor ergodic. To overcome such assumptions Verdú and Han pioneered the technique of information-spectrum methods in their seminal work [4]. In this work Verdú and Han define the notions of limit inferior and limit superior in probability. They then use these definitions to establish the capacity of general channels (channels that are not necessarily stationary and/or memoryless). Since this work of Verdú and Han there have been a considerable interest in generalizing the results of information theory in the information spectrum setting, see for example, [5], [6], [7], [8] and references therein.

This general technique of information-spectrum methods wherein no assumptions are made on the channels and/or

sources were extended to the quantum case by Hayashi, Nagaoka and Ogawa. Using this method they studied the problem of quantum hypothesis testing [9], [10], deriving the classical capacity formula of general quantum channels [11] and establishing general formula for the optimal rate of entanglement concentration [12]. Since the work of Hayashi, Nagaoka and Ogawa the study of various quantum information theoretic protocols in the information spectrum setting have been one of the most interesting areas of research in the theoretical quantum information science. In [13] Bowen and Datta further carried forward this approach to study various other quantum information theoretic problems. In [14] Datta and Renner showed that there is a close relationship with the information theoretic quantities that arise in the information-spectrum scenario and smooth Rényi entropies which play a crucial role in one-shot information theory. In [15] Radhakrishnan, Sen and Warsi proved one-shot version of the Marton inner bound for the classical-quantum broadcast channels. They then showed that their one-shot bounds yields the quantum information-spectrum generalization of the Marton inner bound in the asymptotic setting.

In this paper, we carry forward the subject of studying quantum information theoretic protocols in the information-spectrum setting. We study the problem of communication over the sequence of channels $\{\mathcal{X}^n, \mathcal{S}^n, \mathcal{N}^{\mathcal{X}^n \mathcal{S}^n \rightarrow B^n}(x^n, s^n) = \rho_{x^n, s^n}^{B^n}\}_{n=1}^{\infty}$ (also called as the classical-quantum Gel’fand-Pinsker channel), where $\mathcal{X}^n, \mathcal{S}^n$ are the input and state alphabets and $\rho_{x^n, s^n}^{B^n}$ is a positive operator with trace one acting on the Hilbert space $\mathcal{H}_B^{\otimes n}$. The classical version of this problem in the information spectrum setting was studied by Tan in [16].

The problem of transmitting messages successfully across the channels with states differs from the problem studied by Hayashi and Nagaoka in [11]. In [11] it is assumed that the channel statistics do not change over transmissions and are completely known to the senders and the receivers. In this work, we let go of this assumption. Figure I below models a communication settings where the channel statistics are not fully known, uncertainty about the channel statistics in such settings is captured by a state that may vary randomly over transmissions.

Naqeeb is with IIITD, Delhi and the Department of Engineering Science, University of Oxford, email: naqeeb@iiitd.ac.in. Justin is with the Department of Engineering Science, University of Oxford, email: justin.coon@eng.ox.ac.uk.

This paper was presented at ISIT 2016, Barcelona, Spain.

Copyright (c) 2014 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

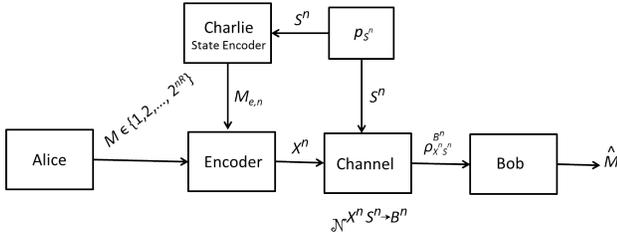


Fig. 1. Classical-Quantum Gel'fand-Pinsker channel communication scheme when rate limited channel state information is available at the encoder. The message M is independent of the state sequence S^n and is uniformly distributed over the set $[1 : 2^{nR}]$. The index $M_{e,n}$ is distributed over the set $[1 : 2^{nR_S}]$ and acts as a rate limited side information about S^n to the encoder.

II. DEFINITION

Definition 1. Let $(\mathbf{U}, \tilde{\mathbf{S}}) := \{U^n, \tilde{S}^n\}_{n=1}^\infty$ be a sequence of pair of random variables, where for every n $(U^n, \tilde{S}^n) \sim p_{U^n S^n}$ and take values over the set $(\mathcal{U}^n \times \tilde{\mathcal{S}}^n)$. The spectral-sup mutual information rate $\bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}]$ between \mathbf{U} and $\tilde{\mathbf{S}}$ is defined as follows:

$$\bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] := \inf \left\{ a : \lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{p_{U^n \tilde{S}^n}}{p_{U^n} p_{\tilde{S}^n}} > a \right\} = 0 \right\}, \quad (1)$$

where a is arbitrary and the probability above is calculated with respect to $p_{U^n \tilde{S}^n}$.

Definition 2. Let $\rho := \{\rho_n\}_{n=1}^\infty$ and $\sigma := \{\sigma_n\}_{n=1}^\infty$ be sequences of quantum states where for every n , ρ_n and σ_n are density matrices acting on the Hilbert space $\mathcal{H}_n := \mathcal{H}^{\otimes n}$. The spectral-inf mutual information rate $\underline{\mathbf{I}}[\rho; \sigma]$ between ρ and σ is defined as follows:

$$\underline{\mathbf{I}}[\rho; \sigma] := \sup \left\{ a : \lim_{n \rightarrow \infty} \text{Tr} [\{\rho_n \succ 2^{na} \sigma_n\} \rho_n] = 1 \right\}, \quad (2)$$

where $\{\rho_n \succ 2^{na} \sigma_n\}$ represents a projection operator onto the non-negative Eigen space of the operator $(\rho_n - 2^{na} \sigma_n)$.

Definition 3. An $(n, M_n, M_{e,n}, \varepsilon_n)$ code for the Gel'fand-Pinsker channel $\{\mathcal{X}^n, \mathcal{S}^n, \mathcal{N}^{X^n S^n \rightarrow B^n}(x^n, s^n) = \rho_{x^n, s^n}^{B^n}\}$ with coded side information available at the encoder consists of

- a state encoding $f_{e,n} : \mathcal{S}^n \rightarrow [1 : M_{e,n}]$
- an encoding function $f_n : [1 : M_n] \times [1 : M_{e,n}] \rightarrow \mathcal{X}^n$ (possibly randomized)
- A decoding POVM $\{\beta(m) : m \in [1 : M_n]\}$ such that

$$\frac{1}{M_n} \sum_{s^n} p_{S^n}(s^n) \text{Tr} [(\mathbb{I} - \mathcal{N}(f_n(m), f_{e,n}(s^n)), s^n)) \beta(m)] \leq \varepsilon_n.$$

Definition 4. A rate pair (R, R_S) is achievable if there exists a sequence of $(n, M_n, M_{e,n}, \varepsilon_n)$ codes such that

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log M_n &> R \\ \limsup_{n \rightarrow \infty} \varepsilon_n &= 0 \\ \limsup_{n \rightarrow \infty} \frac{1}{n} \log M_{e,n} &< R_S. \end{aligned}$$

The set of all achievable rate pairs is known as the capacity region.

III. PREVIOUS WORKS

The classical version of the problem described in Figure 1 was studied by Heggard and El Gamal (achievability) in [17] in the asymptotic iid setting. They proved the following:

Theorem 1. Fix a discrete memoryless channel with state characterized by $p(y | x, s)$. Let (R, R_S) be such that

$$\begin{aligned} R &< I[U; Y] - I[U; \tilde{S}] \\ R_S &> I[S; \tilde{S}], \end{aligned}$$

for some distribution $p(s, \tilde{s}, u, x, y) = p(s)p(\tilde{s} | s)p(u | \tilde{s})p(x | u, \tilde{s})p(y | x, s)$. Then, the rate pair (R, R_S) is achievable.

Furthermore, in [17] Heggard and El Gamal argued that Theorem 1 implies the result of Gel'fand and Pinsker [18] who showed the following:

Theorem 2. Fix a discrete memoryless channel with state characterized by $p(y | x, s)$. The capacity of this channel when the state information is directly available non-causally at the encoder is

$$C = \max_{p_{u|s}, g: \mathcal{U} \times \mathcal{S} \rightarrow \mathcal{X}} (I[U; Y] - I[U; S]),$$

where $|\mathcal{U}| \leq \min \{|\mathcal{X}|, |\mathcal{S}|, |\mathcal{Y}| + |\mathcal{S}| - 1\}$.

The above formula for the capacity can be derived from Theorem 1 by assuming S and \tilde{S} to be independent and subject to the same probability distribution.

OUR RESULT

We establish the capacity region of the classical-quantum Gel'fand-Pinsker channel in the information-spectrum setting when rate limited version of the channel state is available at the encoder. In the information-spectrum setting the channel output $\rho_{X^n, S^n}^{B^n}$, need not be a tensor product state. Furthermore, the channel state $S^n \sim p_{S^n}$, is a sequence of arbitrarily distributed random variables. This extremely general setting is the hallmark of information-spectrum approach. We prove the following:

Theorem 3. Let $\{\mathcal{X}^n, \mathcal{S}^n, \mathcal{N}^{X^n S^n \rightarrow B^n}(x^n, s^n) = \rho_{x^n, s^n}^{B^n}\}_{n=1}^\infty$ be a sequence of classical-quantum Gel'fand-Pinsker channels. The capacity region for this sequence of channels with rate

limited version of the channel state available only at the encoder is the set of rate pairs satisfying the following:

$$\begin{aligned} R &\leq \underline{\mathbf{I}}[\mathbf{U}; \mathbf{B}] - \bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}]; \\ R_S &\geq \bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}]. \end{aligned}$$

The information theoretic quantities are calculated with respect to the sequence of states $\Theta^{S^n \tilde{S}^n U^n X^n B^n} = \left\{ \Theta^{S^n \tilde{S}^n U^n X^n B^n} \right\}_{n=1}^{\infty}$, where for every n ,

$$\begin{aligned} &\Theta^{S^n \tilde{S}^n U^n X^n B^n} \\ &:= \sum_{s^n, \tilde{s}^n, u^n, x^n} p_{S^n}(s^n) p_{\tilde{S}^n | S^n}(\tilde{s}^n | s^n) p_{U^n | \tilde{S}^n}(u^n | \tilde{s}^n) \\ &\quad p_{X^n | U^n \tilde{S}^n}(x^n | u^n, \tilde{s}^n) |s^n\rangle \langle s^n|^{S^n} \otimes |s^n\rangle \langle s^n|^{S^n} \\ &\quad \otimes |u^n\rangle \langle u^n|^{U^n} \otimes |x^n\rangle \langle x^n|^{X^n} \otimes \rho_{x^n, s^n}^{B^n}. \end{aligned}$$

An immediate consequence of Theorem 3 is the following corollary:

Corollary 1. (a) (Hayashi and Nagaoka, [11]) The capacity of a sequence of classical-quantum channels $\{\mathcal{X}^n, \mathcal{N}^{X^n \rightarrow B^n}(x^n) = \rho_{x^n}^{B^n}\}_{n=1}^{\infty}$ is the following:

$$C = \sup_{\{X^n\}_{n=1}^{\infty}} \underline{\mathbf{I}}[\mathbf{X}; \mathbf{B}].$$

(b) The capacity of a sequence of classical-quantum Gel'fand-Pinsker channels $\{\mathcal{X}^n, \mathcal{S}^n, \mathcal{N}^{X^n S^n \rightarrow B^n}(x^n, s^n) = \rho_{x^n, s^n}^{B^n}\}_{n=1}^{\infty}$ with channel state directly available at the encoder is the following:

$$C = \sup_{\{\Theta_n\}_{n=1}^{\infty}} \underline{\mathbf{I}}[\mathbf{U}; \mathbf{B}] - \bar{\mathbf{I}}[\mathbf{U}; \mathbf{S}],$$

where for every n ,

$$\begin{aligned} \Theta_n &= \sum_{s^n, u^n, x^n} p_{S^n}(s^n) p_{U^n | S^n}(u^n | s^n) p_{X^n | U^n S^n}(x^n | u^n, s^n) \\ &\quad |s^n\rangle \langle s^n|^{S^n} \otimes |u^n\rangle \langle u^n|^{U^n} \otimes |x^n\rangle \langle x^n|^{X^n} \\ &\quad \otimes \rho_{x^n, s^n}^{B^n}. \end{aligned}$$

Proof. a) The proof follows by setting $\tilde{S}^n = S^n = \emptyset$ and $U^n = X^n$ in Theorem 3.

b) Let \tilde{S}^n and S^n be independent and subject to same probability distribution in Theorem 3. \square

Techniques: The proof of Theorem 3 involves both classical and quantum arguments. In the classical version of the problem (Theorem 1) in the asymptotic i.i.d. case the proof heavily relies on the covering lemma [Lemma 3.3, [19]], packing lemma [Lemma 3.1, [19]] and conditional typicality lemma [page 27, [19]]: conditioned on the input to the channel being typical the output will be typical with very high probability. However, in the non-i.i.d. case and with channel output being quantum we do not have such powerful results. We show in the subsequent sections below that it is indeed still possible to derive the quantum version of Theorem 1. In

the asymptotic i.i.d. case one of the key idea while encoding at the Charlie's and Alice's end is based on typicality: the empirical distribution of symbols in a long sequence of trials will with high probability be close to the true distribution. We overcome this typicality based encoding by using a *rejection sampling* based encoding approach. Also, to overcome the notion of conditional typicality in the non-i.i.d case we *blend* our rejection sampling technique with some carefully defined events which play a very crucial in bounding the decoding error probability at Bob's end. Further, it requires a very careful analysis of the probabilistic terms to derive the result of Theorem 3. Section IV below makes the discussions of this paragraph more precise.

IV. PROOF OF THEOREM 3

A. Achievability

Let,

$$\begin{aligned} &\rho_{u^n, \tilde{s}^n}^{B^n} \\ &= \sum_{s^n, x^n} p_{S^n | \tilde{S}^n}(s^n | \tilde{s}^n) p_{X^n | U^n \tilde{S}^n}(x^n | u^n, \tilde{s}^n) \rho_{x^n, s^n}^{B^n} \end{aligned} \quad (3)$$

$$\Theta^{U^n B^n} = \text{Tr}_{S^n \tilde{S}^n X^n} \left[\Theta^{S^n \tilde{S}^n U^n X^n B^n} \right] \quad (4)$$

$$\Theta^{U^n} = \text{Tr}_{B^n} \left[\Theta^{U^n B^n} \right] \quad (5)$$

$$\Theta^{B^n} = \text{Tr}_{U^n} \left[\Theta^{U^n B^n} \right]. \quad (6)$$

Fix $\gamma > 0$. Let,

$$\Pi^{U^n B^n} := \left\{ \Theta^{U^n B^n} \succeq 2^{n(\underline{\mathbf{I}}[\mathbf{U}; \mathbf{B}] - \gamma)} \Theta^{U^n} \otimes \Theta^{B^n} \right\}, \quad (7)$$

where $\underline{\mathbf{I}}[\mathbf{U}; \mathbf{B}]$ is calculated with respect to the sequence of states $\{\Theta^{U^n B^n}\}_{n=1}^{\infty}$ and $\{\Theta^{U^n} \otimes \Theta^{B^n}\}_{n=1}^{\infty}$. Further, for every $u^n \in \mathcal{U}^n$, let

$$\Lambda_{u^n} := \text{Tr}_{U^n} \left[\Pi^{U^n B^n} (|u^n\rangle \langle u^n| \otimes \mathbb{I}) \right]. \quad (8)$$

Define the following sets:

$$\begin{aligned} &\mathcal{T}_n(p_{S^n \tilde{S}^n}) \\ &:= \left\{ (s^n, \tilde{s}^n) : \frac{1}{n} \log \frac{p_{S^n \tilde{S}^n}(s^n, \tilde{s}^n)}{p_{S^n}(s^n) p_{\tilde{S}^n}(\tilde{s}^n)} \leq \bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma \right\}; \\ &\mathcal{T}_n(p_{U^n \tilde{S}^n}) \\ &:= \left\{ (u^n, \tilde{s}^n) : \frac{1}{n} \log \frac{p_{U^n \tilde{S}^n}(u^n, \tilde{s}^n)}{p_{U^n}(u^n) p_{\tilde{S}^n}(\tilde{s}^n)} \leq \bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] + \gamma \right\}. \end{aligned}$$

Furthermore, let $g_1 : \tilde{\mathcal{S}}^n \rightarrow [0, 1]$ and $g_2 : \tilde{\mathcal{S}}^n \rightarrow [0, 1]$ be defined as follows:

$$g_1(\tilde{s}^n) = \sum_{u^n : (u^n, \tilde{s}^n) \notin \mathcal{T}_n(p_{U^n \tilde{S}^n})} p_{U^n | \tilde{S}^n}(u^n | \tilde{s}^n); \quad (9)$$

$$g_2(\tilde{s}^n) = \sum_{u^n : \text{Tr}[\Lambda_{u^n} \rho_{u^n, \tilde{s}^n}^{B^n}] \leq 1 - \sqrt{\varepsilon}} p_{U^n | \tilde{S}^n}(u^n | \tilde{s}^n). \quad (10)$$

Remark: Note that $\mathcal{T}_n(p_{S^n \tilde{S}^n})$ and $\mathcal{T}_n(p_{U^n \tilde{S}^n})$ depend on γ . Further $g_2(\tilde{s}^n)$ depends on both γ and ε .

The codebook: Let $\Theta^{S^n \tilde{S}^n U^n X^n B^n}$ be as in the statement of the theorem. Let $R = \mathbf{I}[\mathbf{U}; \mathbf{B}] - \mathbf{I}[\mathbf{U}; \tilde{\mathbf{S}}] - 6\gamma$, $r = \mathbf{I}[\mathbf{U}; \tilde{\mathbf{S}}] + 2\gamma$ so that $R+r = \mathbf{I}[\mathbf{U}; \mathbf{B}] - 4\gamma$. Let $u^n[1], u^n[2], \dots, u^n[2^{n(R+r)}]$ be drawn independently according to the distribution p_{U^n} . We associate these samples with a row vector $\mathcal{C}_n^{(A)}$ having $2^{n(R+r)}$ entries. We then partition this row vector into 2^{nR} classes each containing 2^{nr} elements. Every message $m \in [1 : 2^{nR}]$ is uniquely assigned a class. We will denote the class corresponding to the message m by $\mathcal{C}_n^{(A)}(m)$.

Fix $R_S = \mathbf{I}[\mathbf{S}; \tilde{\mathbf{S}}] + 2\gamma$. Further, let $\tilde{s}^n[1], \tilde{s}^n[2], \dots, \tilde{s}^n[2^{nR_S}]$ be drawn independently according to the distribution $p_{\tilde{S}^n}$. We will denote this collection of sequences by $\mathcal{C}_n^{(C)}$. These collection of sequences present in $\mathcal{C}_n^{(C)}$ are made known to Alice as well.

Charlie's encoding strategy: For each $k \in [1 : 2^{nR_S}]$, let $Z(k)$ be independently and uniformly distributed over $[0, 1]$. For a given realisation of the state sequence s^n , let $\zeta(k)$ be an indicator random variable defined as follows:

$$\zeta(k) = \begin{cases} 1 & \text{if } Z(k) \leq \frac{p_{\tilde{S}^n \tilde{S}^n}(\tilde{s}^n[k], s^n)}{2^{n(\mathbf{I}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma)} p_{\tilde{S}^n}(\tilde{s}^n[k]) p_{S^n}(s^n)}; \\ 0 & \text{otherwise.} \end{cases} \quad (11)$$

Further, for a given realisation of the state sequence s^n , let $\mathbf{I}(k)$ be an indicator random variable defined as follows:

$$\mathbf{I}(k) = \begin{cases} 1 & \text{if } \zeta(k) = 1, g_1(\tilde{S}^n[k]) < \sqrt{\varepsilon} \text{ and} \\ & g_2(\tilde{S}^n[k]) < \varepsilon^{\frac{1}{4}}; \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

where $g_1(\tilde{s}^n)$ and $g_2(\tilde{s}^n)$ are defined in (9) and (10). Charlie on observing the state sequence s^n finds an index k such that $\mathbf{I}(k) = 1$. If there are more than one such indices then k is set as the smallest one among them. If there is none such index then $k = 1$. Charlie then sends this index k to Alice.

Alice's encoding strategy: For each pair $(k, \ell) \in [1 : 2^{nR_S}] \times [1 : 2^{n(R+r)}]$, let $\eta(k, \ell)$ be independently and uniformly distributed over $[0, 1]$ and let $g(k, \ell)$ be defined as follows:

$$g(k, \ell) := \text{Tr} \left[\Lambda_{u^n[\ell]} \rho_{u^n[\ell], \tilde{s}^n[k]}^{B^n} \right], \quad (13)$$

where $\rho_{u^n, \tilde{s}^n}^{B^n}$ is defined in (3) and Λ_{u^n} is defined in (8). Let $\mathbf{I}(k, \ell)$ be an indicator random variable such that

$$\mathbf{I}(k, \ell) = \begin{cases} 1 & \text{if } \eta(k, \ell) \\ & \leq \frac{p_{\tilde{S}^n U^n}(\tilde{s}^n[k], u^n[\ell])}{2^{n(\mathbf{I}[\mathbf{U}; \tilde{\mathbf{S}}] + \gamma)} p_{\tilde{S}^n}(\tilde{s}^n[k]) p_{U^n}(u^n[\ell])}; \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

Further, let $\mathbf{J}(k, \ell)$ be an indicator random variable defined as follows:

$$\mathbf{J}(k, \ell) = \begin{cases} 1 & \text{if } \mathbf{I}(k, \ell) = 1 \text{ and} \\ & g(k, \ell) > 1 - \sqrt{\varepsilon}; \\ 0 & \text{otherwise.} \end{cases} \quad (15)$$

To send a message $m \in [1 : 2^{nR}]$ and on receiving the index k from Charlie, Alice finds an index $\ell \in \mathcal{C}_n^{(A)}(m)$ such that $\mathbf{J}(k, \ell) = 1$. If there are more than such indices

then ℓ is set as the smallest one among them. If there is none such index then $\ell = 1$. Alice then randomly generates $x^n \sim p_{X^n|U^n[\ell], \tilde{S}^n[k]}$ and transmits it over the classical-quantum channel over n channel uses. In the discussions below we will use the notation $x^n(u^n[\ell], \tilde{s}^n[k])$ to highlight the dependence of x^n on $(u^n[\ell], \tilde{s}^n[k])$. A similar encoding technique was also used by Radhakrishnan, Sen and Warsi in [15].

Bob's decoding strategy: For each $\ell \in [1 : 2^{n(R+r)}]$, we have the operators $\Lambda_{u^n[\ell]}$ as defined in (8). Bob will normalize these operators to obtain a POVM. The POVM element corresponding to ℓ will be

$$\beta_n(\ell) := \left(\sum_{\ell' \in [1 : 2^{n(R+r)}]} \Lambda_{u^n[\ell']} \right)^{-\frac{1}{2}} \Lambda_{u^n[\ell]} \left(\sum_{\ell' \in [1 : 2^{n(R+r)}]} \Lambda_{u^n[\ell']} \right)^{-\frac{1}{2}}.$$

Bob on receiving the channel output measures it using these operators. If the measurement outcome is $\tilde{\ell}$ then he outputs \tilde{m} if $\tilde{\ell} \in \mathcal{C}_n^{(A)}(\tilde{m})$. Similar decoding POVM elements were also used by Hayashi and Nagaoka in [11] and by Wang and Renner in [20].

Probability of error analysis: Let a message $m \in [1 : 2^{nR}]$ be transmitted by Alice by using the protocol discussed above and suppose it is decoded as \tilde{m} by Charlie. We will now show that the probability $\tilde{m} \neq m$, averaged over the random choice of codebook, the state sequence S^n and X^n is arbitrary close to zero. By the symmetry of the code construction it is enough to prove the claim for $m = 1$. There are following sources of error:

- 1) Charlie on observing the state sequence S^n does not find a suitable $k \in \mathcal{C}_n^{(C)}$ such that $\mathbf{I}(k) = 1$.
- 2) Alice on receiving the index k from Charlie is not able to find a suitable $U^n[\ell] \in \mathcal{C}_n^{(A)}(1)$ such that $\mathbf{J}(k, \ell) = 1$.
- 3) Charlie finds a suitable k and Alice finds a suitable ℓ , but Bob's measurement is not able to determine the index ℓ correctly.

Let k^* and ℓ^* be the indices chosen by Charlie and Alice. Let us now upper bound the probability of error while decoding the transmitted message. Towards this we first define the following events:

$$\begin{aligned} \mathcal{E}_1 &:= \text{for all } k \in [1 : 2^{nR_S}] : \mathbf{I}(k) = 0; \\ \mathcal{E}_2 &:= \text{for all } \ell \in \mathcal{C}_n^{(A)}(1) : \mathbf{J}(k^*, \ell) = 0. \end{aligned}$$

We now have the following bound on the error probability:

$$\begin{aligned} & \Pr \{ \tilde{m} \neq 1 \} \\ & \leq \Pr \{ \tilde{\ell} \neq \ell^* \} \\ & \leq \Pr \{ \mathcal{E}_1 \cup \mathcal{E}_2 \} + \Pr \{ (\mathcal{E}_1 \cup \mathcal{E}_2)^c, \tilde{\ell} \neq \ell^* \} \\ & \leq \Pr \{ \mathcal{E}_1 \} + \Pr \{ \mathcal{E}_2 \} + \Pr \{ \mathcal{E}_1^c \cap \mathcal{E}_2^c, \tilde{\ell} \neq \ell^* \} \\ & \leq 2 \Pr \{ \mathcal{E}_1 \} + \Pr \{ \mathcal{E}_1^c \cap \mathcal{E}_2 \} \\ & \quad + \Pr \{ \mathcal{E}_1^c \cap \mathcal{E}_2^c, \tilde{\ell} \neq \ell^* \}, \end{aligned} \quad (16)$$

where the first inequality follows from the setting of the protocol discussed above and remaining all of the inequalities till (16) follow from the union bound. In what follows we will now show that for n large enough we have

$$2 \Pr \{ \mathcal{E}_1 \} + \Pr \{ \mathcal{E}_1^c \cap \mathcal{E}_2 \} + \Pr \left\{ \mathcal{E}_1^c \cap \mathcal{E}_2^c, \tilde{\ell} \neq \ell^* \right\} \leq 6\varepsilon \\ + 3\sqrt{\varepsilon} + 3\varepsilon^{\frac{1}{4}} + \frac{2\sqrt{\varepsilon}}{(1 - \varepsilon - \sqrt{\varepsilon} - \varepsilon^{\frac{1}{4}})} + 3 \exp(-2^{n\gamma}),$$

where $\varepsilon > 0$ is arbitrarily close to zero such that $(\varepsilon + \sqrt{\varepsilon} + \varepsilon^{\frac{1}{4}}) < 1$.

Consider $\Pr \{ \mathcal{E}_1 \}$:

$$\Pr \{ \mathcal{E}_1 \} \\ \stackrel{a}{=} \sum_{s^n \in \mathcal{S}^n} p_{S^n}(s^n) (1 - \Pr \{ \mathbf{I}(k) = 1 \mid S^n = s^n \})^{2^{nR_S}} \\ \leq \sum_{s^n \in \mathcal{S}^n} p_{S^n}(s^n) \left(1 - \sum_{\substack{\tilde{s}^n: (\tilde{s}^n, s^n) \in \mathcal{T}_n(p_{S^n \tilde{S}^n}) \\ g_1(\tilde{s}^n) < \sqrt{\varepsilon} \\ g_2(\tilde{s}^n) < \varepsilon^{\frac{1}{4}}}} p_{\tilde{S}^n}(\tilde{s}^n) \right)^{2^{nR_S}} \\ \Pr \left\{ \mathbf{I}(k) = 1 \mid S^n = s^n, \tilde{S}^n[k] = \tilde{s}^n \right\} \\ \stackrel{b}{=} \sum_{s^n \in \mathcal{S}^n} p_{S^n}(s^n) \left(1 - 2^{-n(\bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma)} \right. \\ \left. \sum_{\substack{\tilde{s}^n: (\tilde{s}^n, s^n) \in \mathcal{T}_n(p_{S^n \tilde{S}^n}) \\ g_1(\tilde{s}^n) < \sqrt{\varepsilon} \\ g_2(\tilde{s}^n) < \varepsilon^{\frac{1}{4}}}} p_{\tilde{S}^n}(\tilde{s}^n) \frac{p_{S^n \tilde{S}^n}(s^n, \tilde{s}^n)}{p_{S^n}(s^n) p_{\tilde{S}^n}(\tilde{s}^n)} \right)^{2^{nR_S}} \\ \stackrel{c}{\leq} \sum_{s^n \in \mathcal{S}^n} p_{S^n}(s^n) \exp \left(-2^{n(R_S - (\bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma))} \right. \\ \left. \sum_{\substack{\tilde{s}^n: (\tilde{s}^n, s^n) \in \mathcal{T}_n(p_{S^n \tilde{S}^n}) \\ g_1(\tilde{s}^n) < \sqrt{\varepsilon} \\ g_2(\tilde{s}^n) < \varepsilon^{\frac{1}{4}}}} p_{\tilde{S}^n | S^n}(\tilde{s}^n \mid s^n) \right) \\ \stackrel{d}{=} \sum_{s^n \in \mathcal{S}^n} p_{S^n}(s^n) \exp \left(-2^{n\gamma} \sum_{\substack{\tilde{s}^n: (\tilde{s}^n, s^n) \in \mathcal{T}_n(p_{S^n \tilde{S}^n}) \\ g_1(\tilde{s}^n) < \sqrt{\varepsilon} \\ g_2(\tilde{s}^n) < \varepsilon^{\frac{1}{4}}}} p_{\tilde{S}^n | S^n}(\tilde{s}^n \mid s^n) \right) \\ \stackrel{e}{\leq} \Pr \{ \mathcal{T}_n^c(p_{S^n \tilde{S}^n}) \} + \Pr \left\{ g_1(\tilde{S}^n) \geq \sqrt{\varepsilon} \right\} \\ + \Pr \left\{ g_2(\tilde{S}^n) \geq \varepsilon^{\frac{1}{4}} \right\} + \exp(-2^{n\gamma}) \\ \stackrel{f}{\leq} \varepsilon + \exp(-2^{n\gamma}) + \Pr \left\{ g_1(\tilde{S}^n) \geq \sqrt{\varepsilon} \right\} \\ + \Pr \left\{ g_2(\tilde{S}^n) \geq \varepsilon^{\frac{1}{4}} \right\}, \quad (17)$$

where a follows because $\mathbf{I}(1), \dots, \mathbf{I}(2^{nR_S})$ are independent and identically distributed and $\tilde{S}^n[1], \dots, \tilde{S}^n[2^{nR_S}]$ are independent and identically distributed according to the distribution $p_{\tilde{S}^n}$, b follows from the definition of $\mathbf{I}(k)$, c follows from the inequality $(1 - x)^y \leq e^{-xy}$ ($0 \leq x \leq 1, y \geq 0$),

d follows because $R_s = \bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] + 2\gamma$, e follows because $(e^{-xy}) \leq 1 - x + e^{-y}$ ($0 \leq x \leq 1, y \geq 0$) and union bound and f follows because n is large enough such that $\Pr \{ \mathcal{T}_n^c(p_{S^n \tilde{S}^n}) \} \leq \varepsilon$. Let us now bound each of the last two terms on the R.H.S. of (17) as follows:

Consider $\Pr \left\{ g_1(\tilde{S}^n) \geq \sqrt{\varepsilon} \right\}$:

$$\Pr \left\{ g_1(\tilde{S}^n) \geq \sqrt{\varepsilon} \right\} \\ \stackrel{a}{\leq} \frac{\mathbb{E}[g_1(\tilde{S}^n)]}{\sqrt{\varepsilon}} \\ \stackrel{b}{=} \frac{\sum_{\tilde{s}^n} \sum_{u^n: (u^n, \tilde{s}^n) \notin \mathcal{T}_n(p_{U^n \tilde{S}^n})} p_{U^n \tilde{S}^n}(u^n, \tilde{s}^n)}{\sqrt{\varepsilon}} \\ = \frac{\Pr \left\{ (U^n, \tilde{S}^n) \notin \mathcal{T}_n(p_{U^n \tilde{S}^n}) \right\}}{\sqrt{\varepsilon}} \\ \stackrel{c}{\leq} \frac{\varepsilon}{\sqrt{\varepsilon}} \\ = \sqrt{\varepsilon}, \quad (18)$$

where a follows from Markov inequality; b follows from the definition of $g_1(\tilde{s}^n)$ and taking expectation over the random variable \tilde{S}^n and c follows under the assumption that n is large enough such that $\Pr \left\{ (U^n, \tilde{S}^n) \notin \mathcal{T}_n(p_{U^n \tilde{S}^n}) \right\} \leq \varepsilon$.

Consider $\Pr \left\{ g_2(\tilde{S}^n) \geq \varepsilon^{\frac{1}{4}} \right\}$:

$$\Pr \left\{ g_2(\tilde{S}^n) \geq \varepsilon^{\frac{1}{4}} \right\} \\ \stackrel{a}{\leq} \frac{\mathbb{E}[g_2(\tilde{S}^n)]}{\varepsilon^{\frac{1}{4}}} \\ \stackrel{b}{=} \frac{\sum_{\tilde{s}^n} \sum_{u^n: \text{Tr}[\Lambda_{u^n} \rho_{u^n, \tilde{s}^n}^{B^n}] \leq 1 - \sqrt{\varepsilon}} p_{U^n \tilde{S}^n}(u^n, \tilde{s}^n)}{\varepsilon^{\frac{1}{4}}} \\ = \frac{\Pr \left\{ \text{Tr} \left[\Lambda_{U^n} \rho_{U^n, \tilde{S}^n}^{B^n} \right] \leq 1 - \sqrt{\varepsilon} \right\}}{\varepsilon^{\frac{1}{4}}} \\ \stackrel{c}{\leq} \frac{\sqrt{\varepsilon}}{\varepsilon^{\frac{1}{4}}} \\ = \varepsilon^{\frac{1}{4}}, \quad (19)$$

where a follows from Markov inequality; b follows from the definition of $g_2(\tilde{s}^n)$ and by taking the expectation over the random variable \tilde{S}^n and c follows because of the following set of inequalities:

$$\Pr \left\{ \text{Tr} \left[\Lambda_{U^n} \rho_{U^n, \tilde{S}^n}^{B^n} \right] \leq 1 - \sqrt{\varepsilon} \right\} \\ \stackrel{a}{\leq} \frac{1 - \mathbb{E} \text{Tr} \left[\Lambda_{U^n} \rho_{U^n, \tilde{S}^n}^{B^n} \right]}{\sqrt{\varepsilon}} \\ \stackrel{b}{=} \frac{1}{\sqrt{\varepsilon}} \\ - \frac{\mathbb{E} \text{Tr} \left[\text{Tr}_{U^n} \left[\Pi^{U^n B^n} (|U^n\rangle\langle U^n| \otimes \mathbb{I}) \right] \rho_{U^n, \tilde{S}^n}^{B^n} \right]}{\sqrt{\varepsilon}}$$

$$\begin{aligned} &\stackrel{c}{=} \frac{1 - \text{Tr} [\Pi^{U^n B^n} \Theta^{U^n B^n}]}{\sqrt{\varepsilon}} \\ &\stackrel{d}{\leq} \sqrt{\varepsilon}, \end{aligned}$$

where a follows from the Markov inequality, b follows from the definition of Λ_{U^n} mentioned in (8), c follows from the definition of $\Theta^{U^n B^n}$ mentioned in (4) and d follows under the assumption that n is large enough such that $\text{Tr} [\Pi^{U^n B^n} \Theta^{U^n B^n}] \geq 1 - \varepsilon$. Thus, it now follows from (17), (18) and (19) that

$$\Pr \{\mathcal{E}_1\} \leq \varepsilon + \sqrt{\varepsilon} + \varepsilon^{\frac{1}{4}} + \exp(-2^{n\gamma}). \quad (20)$$

Consider $\Pr \{\mathcal{E}_1^c \cap \mathcal{E}_2\}$:

$$\begin{aligned} &\Pr \{\mathcal{E}_1^c \cap \mathcal{E}_2\} \\ &= \mathbb{E}_{(k^*, \tilde{S}^n[k^*], S^n)} \mathbf{I}(\mathcal{E}_1^c) \Pr \{\forall \ell \in \mathcal{C}(1) : \mathbf{J}(k^*, \ell) = 0\} \\ &\stackrel{a}{=} \mathbb{E}_{(k^*, \tilde{S}^n[k^*], S^n)} \mathbf{I}(\mathcal{E}_1^c) (1 - \Pr \{\mathbf{J}(k^*, \ell) = 1\})^{2^{nr}} \\ &\stackrel{b}{=} \mathbb{E}_{(k^*, \tilde{S}^n[k^*], S^n)} \mathbf{I}(\mathcal{E}_1^c) \\ &\quad \left(1 - \Pr \{\mathbf{I}(k^*, \ell) = 1, g(k^*, \ell) > 1 - \sqrt{\varepsilon}\} \right)^{2^{nr}} \\ &\leq \mathbb{E}_{(k^*, \tilde{S}^n[k^*], S^n)} \mathbf{I}(\mathcal{E}_1^c) \\ &\quad \left(1 - 2^{-n(\bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] + \gamma)} \sum_{\substack{u^n: (\tilde{S}^n[k^*], u^n) \in \mathcal{T}_n(p_{U^n \tilde{S}^n}) \\ \text{Tr}[\Lambda_{u^n} \rho_{u^n, \tilde{S}^n[k^*]}^{B^n}] > 1 - \sqrt{\varepsilon}}} p_{U^n}(u^n) \frac{p_{\tilde{S}^n|U^n}(\tilde{S}^n[k^*], u^n)}{p_{\tilde{S}^n}(\tilde{S}^n[k^*]) p_{U^n}(u^n)} \right)^{2^{nr}} \\ &\stackrel{c}{\leq} \mathbb{E}_{(k^*, \tilde{S}^n[k^*], S^n)} \mathbf{I}(\mathcal{E}_1^c) \\ &\quad \exp \left(-2^{n(r - (\bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] + \gamma))} \sum_{\substack{u^n: (\tilde{S}^n[k^*], u^n) \in \mathcal{T}_n(p_{U^n \tilde{S}^n}) \\ \text{Tr}[\Lambda_{u^n} \rho_{u^n, \tilde{S}^n[k^*]}^{B^n}] > 1 - \sqrt{\varepsilon}}} p_{U^n|\tilde{S}^n}(u^n | \tilde{S}^n[k^*]) \right) \\ &\stackrel{d}{\leq} \mathbb{E}_{(k^*, \tilde{S}^n[k^*], S^n)} \mathbf{I}(\mathcal{E}_1^c) \\ &\quad \exp \left(-2^{n\gamma} \sum_{\substack{u^n: (\tilde{S}^n[k^*], u^n) \in \mathcal{T}_n(p_{U^n \tilde{S}^n}) \\ \text{Tr}[\Lambda_{u^n} \rho_{u^n, \tilde{S}^n[k^*]}^{B^n}] > 1 - \sqrt{\varepsilon}}} p_{U^n|\tilde{S}^n}(u^n | \tilde{S}^n[k^*]) \right) \\ &\stackrel{e}{\leq} \mathbb{E}_{(k^*, \tilde{S}^n[k^*], S^n)} \mathbf{I}(\mathcal{E}_1^c) \\ &\quad \left(1 - \sum_{\substack{u^n: (\tilde{S}^n[k^*], u^n) \in \mathcal{T}_n(p_{U^n \tilde{S}^n}) \\ \text{Tr}[\Lambda_{u^n} \rho_{u^n, \tilde{S}^n[k^*]}^{B^n}] > 1 - \sqrt{\varepsilon}}} p_{U^n|\tilde{S}^n}(u^n | \tilde{S}^n[k^*]) \right) \\ &\quad + \mathbb{E}_{(k^*, \tilde{S}^n[k^*], S^n)} \mathbf{I}(\mathcal{E}_1^c) \exp(-2^{n\gamma}) \\ &\stackrel{f}{\leq} \mathbb{E}_{(k^*, \tilde{S}^n[k^*], S^n)} \mathbf{I}(\mathcal{E}_1^c) \sum_{\substack{u^n: (\tilde{S}^n[k^*], u^n) \in \mathcal{T}_n(p_{U^n \tilde{S}^n}) \\ u^n: (\tilde{S}^n[k^*], u^n) \notin \mathcal{T}_n(p_{U^n \tilde{S}^n})}} p_{U^n|\tilde{S}^n}(u^n | \tilde{S}^n[k^*]) \\ &\quad + \mathbb{E}_{(k^*, \tilde{S}^n[k^*], S^n)} \mathbf{I}(\mathcal{E}_1^c) \exp(-2^{n\gamma}) \\ &\quad + \mathbb{E}_{(k^*, \tilde{S}^n[k^*], S^n)} \mathbf{I}(\mathcal{E}_1^c) \sum_{\substack{u^n: \text{Tr}[\Lambda_{u^n} \rho_{u^n, \tilde{S}^n[k^*]}^{B^n}] \leq 1 - \sqrt{\varepsilon}}} p_{U^n|\tilde{S}^n}(u^n | \tilde{S}^n[k^*]) \\ &\stackrel{g}{\leq} \sqrt{\varepsilon} + \varepsilon^{\frac{1}{4}} + \exp(-2^{n\gamma}), \end{aligned} \quad (21)$$

where a follows because $\eta(k^*, 1), \dots, \eta(k^*, 2^{nr})$ are independent and identically distributed and the fact that $U^n[1], \dots, U^n[2^{nr}]$ are independent and subject to identical distribution p_{U^n} , b follows from the definition of $\mathbf{J}(k^*, \ell)$; c follows from the fact that $(1-x)^y \leq e^{-xy}$ ($0 \leq x \leq 1, y \geq 0$), d follows because $r = \bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] + 2\gamma$, e follows because $e^{-xy} \leq 1 - x + e^{-y}$ ($0 \leq x \leq 1, y \geq 0$); f follows because of the union bound and g follows because if the event \mathcal{E}_1^c happens then $\sum_{u^n: (\tilde{S}^n[k^*], u^n) \notin \mathcal{T}_n(p_{U^n \tilde{S}^n})} p_{U^n|\tilde{S}^n}(u^n | \tilde{S}^n[k^*]) < \sqrt{\varepsilon}$ and $\sum_{u^n: \text{Tr}[\Lambda_{u^n} \rho_{u^n, \tilde{S}^n[k^*]}^{B^n}] \leq 1 - \sqrt{\varepsilon}} p_{U^n|\tilde{S}^n}(u^n | \tilde{S}^n[k^*]) < \varepsilon^{\frac{1}{4}}$.

Consider $\Pr \{\mathcal{E}_1^c \cap \mathcal{E}_2^c, \tilde{\ell} \neq \ell^*\}$:

$$\begin{aligned} &\Pr \{\mathcal{E}_1^c \cap \mathcal{E}_2^c, \tilde{\ell} \neq \ell^*\} \\ &= \mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} \left[\mathbf{I} \{\mathcal{E}_1^c\} \mathbf{I} \{\mathcal{E}_2^c\} \right. \\ &\quad \left. \text{Tr} \left[(\mathbb{I} - \beta_n(\ell^*)) \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*]), S^n}^{B^n} \right] \right] \\ &\leq 2 \mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} \left[\mathbf{I} \{\mathcal{E}_1^c\} \mathbf{I} \{\mathcal{E}_2^c\} \right. \\ &\quad \left. \text{Tr} \left[(\mathbb{I} - \Lambda_{U^n[\ell^*]}) \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*]), S^n}^{B^n} \right] \right] \\ &\quad + 4 \mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} \sum_{\ell' \neq \ell^*} \left[\mathbf{I} \{\mathcal{E}_1^c\} \right. \\ &\quad \left. \mathbf{I} \{\mathcal{E}_2^c\} \text{Tr} \left[\Lambda_{U^n(\ell')} \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*]), S^n}^{B^n} \right] \right], \end{aligned} \quad (22)$$

where the inequality above follows from the Hayashi-Nagaoka operator inequality [11]. In what follows we show that for n large enough,

$$\begin{aligned} &2 \mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} \left[\mathbf{I} \{\mathcal{E}_1^c\} \mathbf{I} \{\mathcal{E}_2^c\} \right. \\ &\quad \left. \text{Tr} \left[(\mathbb{I} - \Lambda_{U^n[\ell^*]}) \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*]), S^n}^{B^n} \right] \right] \\ &\leq \frac{2\sqrt{\varepsilon}}{\left(1 - \varepsilon - \sqrt{\varepsilon} - \varepsilon^{\frac{1}{4}} \right)}, \end{aligned} \quad (23)$$

and

$$\begin{aligned} &4 \mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} \sum_{\ell' \neq \ell^*} \left[\mathbf{I} \{\mathcal{E}_1^c\} \mathbf{I} \{\mathcal{E}_2^c\} \right. \\ &\quad \left. \text{Tr} \left[\Lambda_{U^n(\ell')} \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*]), S^n}^{B^n} \right] \right] \leq 4\varepsilon. \end{aligned} \quad (24)$$

We would like to highlight here that the proof pertaining to the derivation of (23) is nontrivial and requires careful analysis of the probabilistic terms involved. In fact, the proof for the derivation of (24) is nontrivial as well. However, we would borrow the idea of *over-counting* from [15] to bound (24). The reason for all this nontriviality is because k^* and ℓ^* are random.

$2\mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} \left[\mathbf{I} \{ \mathcal{E}_1^c \} \mathbf{I} \{ \mathcal{E}_2^c \} \text{Tr} \left[\left(\mathbb{I} - \Lambda_{U^n[\ell^*]} \right) \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*], S^n)}^{B^n} \right] \right]$ is bounded as follows:

$$\begin{aligned}
& 2\mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} \left[\mathbf{I} \{ \mathcal{E}_1^c \} \mathbf{I} \{ \mathcal{E}_2^c \} \right. \\
& \quad \left. \text{Tr} \left[\left(\mathbb{I} - \Lambda_{U^n[\ell^*]} \right) \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*], S^n)}^{B^n} \right] \right] \\
&= 2\mathbb{E}_{(k^*, \ell^*)} \sum_{u^n, \tilde{s}^n, x^n, s^n} \Pr \left\{ \mathbf{I}(k^*) = 1, \mathbf{J}(k^*, \ell^*) = 1, \right. \\
& \quad \left. U^n[\ell^*] = u^n, \tilde{S}^n[k^*] = \tilde{s}^n, X^n = x^n, \right. \\
& \quad \left. S^n = s^n \mid (k^*, \ell^*) \right\} \\
& \quad \text{Tr} \left[\left(\mathbb{I} - \Lambda_{(U^n[\ell^*]=u^n)} \right) \rho_{X^n=x^n, S^n=s^n}^{B^n} \right] \\
&\leq 2\mathbb{E}_{(k^*, \ell^*)} \sum_{\tilde{s}^n, s^n, u^n, x^n} \Pr \left\{ \tilde{S}^n[k^*] = \tilde{s}^n, S^n = s^n \right. \\
& \quad \left. \mid k^*, \mathbf{I}(k^*) = 1 \right\} \\
& \quad \Pr \left\{ U^n[\ell^*] = u^n \mid S^n = s^n, (k^*, \ell^*), \right. \\
& \quad \left. \tilde{S}^n[k^*] = \tilde{s}^n, \mathbf{I}(k^*) = 1, \mathbf{J}(k^*, \ell^*) = 1 \right\} \\
& \quad \Pr \left\{ X^n = x^n \mid U^n[\ell^*] = u^n, \tilde{S}^n[k^*] = \tilde{s}^n, \right. \\
& \quad \left. S^n = s^n \right\} \\
& \quad \text{Tr} \left[\left(\mathbb{I} - \Lambda_{(U^n[\ell^*]=u^n)} \right) \rho_{X^n=x^n, S^n=s^n}^{B^n} \right], \tag{25}
\end{aligned}$$

where the above inequality follows because X^n given $(U^n[\ell^*] = u^n, \tilde{S}^n[k^*] = \tilde{s}^n, S^n = s^n)$ is conditionally independent of $(\mathbf{I}(k^*), \mathbf{J}(k^*, \ell^*))$. To get the required bound on $2\mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} [\mathbf{I} \{ \mathcal{E}_1^c \} \mathbf{I} \{ \mathcal{E}_2^c \} \text{Tr}[(\mathbb{I} - \Lambda_{U^n[\ell^*]}) \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*], S^n)}^{B^n}]]$ we will now first show that $\Pr \left\{ \tilde{S}^n[k^*] = \tilde{s}^n, S^n = s^n \mid k^* = k, \mathbf{I}(k^*) = 1 \right\} \leq \frac{p_{\tilde{S}^n S^n}(\tilde{s}^n, s^n)}{(1 - \varepsilon - \sqrt{\varepsilon} - \varepsilon^{\frac{1}{4}})}$. Towards this notice the following set of inequalities:

$$\begin{aligned}
& \Pr \left\{ \tilde{S}^n[k^*] = \tilde{s}^n, S^n = s^n \mid \mathbf{I}(k^*) = 1, k^* = k \right\} \\
&= \Pr \left\{ \tilde{S}^n[k] = \tilde{s}^n, S^n = s^n \mid \mathbf{I}(k) = 1, k^* = k \right\} \\
&\stackrel{a}{=} \Pr \left\{ \tilde{S}^n[k] = \tilde{s}^n, S^n = s^n \mid \mathbf{I}(k) = 1 \right\} \\
&= \frac{\Pr \left\{ \mathbf{I}(k) = 1 \mid \tilde{S}^n[k] = \tilde{s}^n, S^n = s^n \right\}}{\Pr \{ \mathbf{I}(k) = 1 \}} \\
& \quad \times \frac{\Pr \left\{ \tilde{S}^n[k] = \tilde{s}^n, S^n = s^n \right\}}{\Pr \{ \mathbf{I}(k) = 1 \}}
\end{aligned}$$

$$\begin{aligned}
& \leq \frac{2^{-n(\bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma)} p_{\tilde{S}^n S^n}(\tilde{s}^n, s^n)}{p_{\tilde{S}^n}(\tilde{s}^n) p_{S^n}(s^n)} p_{\tilde{S}^n}(\tilde{s}^n) p_{S^n}^n(s^n) \\
& \leq \frac{2^{-n(\bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma)}}{\left(1 - \varepsilon - \sqrt{\varepsilon} - \varepsilon^{\frac{1}{4}} \right)} \\
& = \frac{p_{\tilde{S}^n S^n}(\tilde{s}^n, s^n)}{\left(1 - \varepsilon - \sqrt{\varepsilon} - \varepsilon^{\frac{1}{4}} \right)}, \tag{26}
\end{aligned}$$

where a follows because $(\tilde{S}^n[k], S^n)$ is conditionally independent of k^* given the indicator random variable $\mathbf{I}(k)$ and b follows from the definition of $\mathbf{I}(k)$ and from the fact that $\tilde{S}^n[k]$ is independent of S^n and because of the following set of inequalities:

$$\begin{aligned}
& \Pr \{ \mathbf{I}(k) = 1 \} \\
& \geq \Pr \{ \zeta(k) = 1 \} - \Pr \left\{ \zeta(k) = 1, g_1(\tilde{S}^n[k]) \geq \sqrt{\varepsilon} \right\} \\
& \quad - \Pr \left\{ \zeta(k) = 1, g_2(\tilde{S}^n[k]) \geq \varepsilon^{\frac{1}{4}} \right\} \\
& \stackrel{a}{\geq} \sum_{(\tilde{s}^n, s^n) \in \mathcal{T}_n(p_{\tilde{S}^n S^n})} \Pr \left\{ \tilde{S}^n[k] = \tilde{s}^n \right\} \Pr \{ S^n = s^n \} \\
& \quad - \frac{p_{S^n \tilde{S}^n}(s^n, \tilde{s}^n)}{2^{n(\bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma)} p_{\tilde{S}^n}(\tilde{s}^n) p_{S^n}(s^n)} \\
& \quad - \sum_{s^n, \tilde{s}^n: g_1(\tilde{s}^n) \geq \sqrt{\varepsilon}} \Pr \left\{ \tilde{S}^n[k] = \tilde{s}^n \right\} \Pr \{ S^n = s^n \} \\
& \quad - \frac{p_{S^n \tilde{S}^n}(s^n, \tilde{s}^n)}{2^{n(\bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma)} p_{\tilde{S}^n}(\tilde{s}^n) p_{S^n}(s^n)} \\
& \quad - \sum_{s^n, \tilde{s}^n: g_2(\tilde{s}^n) \geq \varepsilon^{\frac{1}{4}}} \Pr \left\{ \tilde{S}^n[k] = \tilde{s}^n \right\} \Pr \{ S^n = s^n \} \\
& \quad - \frac{p_{S^n \tilde{S}^n}(s^n, \tilde{s}^n)}{2^{n(\bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma)} p_{\tilde{S}^n}(\tilde{s}^n) p_{S^n}(s^n)} \\
& = 2^{-n(\bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma)} \left(\Pr \{ \mathcal{T}_n(p_{\tilde{S}^n S^n}) \} - \Pr \left\{ g_1(\tilde{S}^n) \geq \sqrt{\varepsilon} \right\} \right. \\
& \quad \left. - \Pr \left\{ g_2(\tilde{S}^n) \geq \varepsilon^{\frac{1}{4}} \right\} \right) \\
& \geq 2^{-n(\bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma)} \left(1 - \varepsilon - \sqrt{\varepsilon} - \varepsilon^{\frac{1}{4}} \right),
\end{aligned}$$

where a follows from the definition of $\mathbf{I}(k)$ and b follows from the definition of the set $\mathcal{T}_n(p_{\tilde{S}^n S^n})$ and under the assumption that n is large enough such that $\Pr \{ \mathcal{T}_n(p_{\tilde{S}^n S^n}) \} \geq 1 - \varepsilon$ and from (18) and (19). Combining (25) and (26) we now bound $2\mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} [\mathbf{I} \{ \mathcal{E}_1^c \} \mathbf{I} \{ \mathcal{E}_2^c \} \text{Tr}[(\mathbb{I} - \Lambda_{U^n[\ell^*]}) \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*], S^n)}^{B^n}]]$ as follows:

$$\begin{aligned}
& 2\mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} \left[\mathbf{I} \{ \mathcal{E}_1^c \} \mathbf{I} \{ \mathcal{E}_2^c \} \right. \\
& \quad \left. \text{Tr} \left[\left(\mathbb{I} - \Lambda_{U^n[\ell^*]} \right) \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*], S^n)}^{B^n} \right] \right] \\
& \stackrel{a}{\leq} 2 \sum_{(k, \ell)} \Pr \{ k^* = k, \ell^* = \ell \} \sum_{\tilde{s}^n, s^n, u^n, x^n} \Pr \left\{ \tilde{S}^n[k^*] = \tilde{s}^n, \right. \\
& \quad \left. S^n = s^n \mid k^* = k, \mathbf{I}(k^*) = 1 \right\} \Pr \left\{ U^n[\ell^*] = u^n \right.
\end{aligned}$$

$$\begin{aligned}
& | S^n = s^n, (k^* = k, \ell^* = \ell), \tilde{S}^n[k^*] = \tilde{s}^n, \\
& \mathbf{I}(k^*) = 1, \mathbf{J}(k^*, \ell^*) = 1 \Big\} \Pr \left\{ X^n = x^n \right. \\
& \left. | U^n[\ell^*] = u^n, \tilde{S}^n[k^*] = \tilde{s}^n, S^n = s^n \right\} \\
& \text{Tr} \left[\left(\mathbb{I} - \Lambda_{(U^n[\ell^*]=u^n)} \right) \rho_{X^n=x^n, S^n=s^n}^{B^n} \right] \\
& \leq \frac{2}{\left(1 - \varepsilon - \sqrt{\varepsilon} - \varepsilon^{\frac{1}{4}}\right)} \sum_{(k, \ell)} \Pr \{k^* = k, \ell^* = \ell\} \\
& \sum_{\tilde{s}^n, u^n} p_{\tilde{S}^n}(\tilde{s}^n) \Pr \left\{ U^n[\ell^*] = u^n \mid (k^* = k, \ell^* = \ell), \right. \\
& \left. \tilde{S}^n[k^*] = \tilde{s}^n, \mathbf{J}(k^*, \ell^*) = 1 \right\} \\
& \text{Tr} \left[\left(\mathbb{I} - \Lambda_{(U^n[\ell^*]=u^n)} \right) \sum_{x^n, s^n} p_{S^n|\tilde{S}^n}(s^n \mid \tilde{s}^n) \right. \\
& \left. p_{X^n|U^n \tilde{S}^n}(x^n \mid u^n, \tilde{s}^n) \rho_{X^n=x^n, S^n=s^n}^{B^n} \right] \\
& \stackrel{c}{=} \frac{2}{\left(1 - \varepsilon - \sqrt{\varepsilon} - \varepsilon^{\frac{1}{4}}\right)} \sum_{(k, \ell)} \Pr \{k^* = k, \ell^* = \ell\} \sum_{\tilde{s}^n} p_{\tilde{S}^n}(\tilde{s}^n) \\
& \sum_{u^n} \Pr \left\{ U^n[\ell^*] = u^n \mid k^* = k, \ell^* = \ell, \tilde{S}^n[k^*] = \tilde{s}^n, \right. \\
& \left. \mathbf{J}(k^*, \ell^*) = 1 \right\} \\
& \text{Tr} \left[\left(\mathbb{I} - \Lambda_{(U^n[\ell^*]=u^n)} \right) \rho_{(U^n[\ell^*]=u^n, \tilde{S}^n[k^*]=\tilde{s}^n)}^{B^n} \right] \\
& \stackrel{d}{\leq} \frac{2\sqrt{\varepsilon}}{\left(1 - \varepsilon - \sqrt{\varepsilon} - \varepsilon^{\frac{1}{4}}\right)} \sum_{(k, \ell)} \Pr \{k^* = k, \ell^* = \ell\} \\
& \sum_{\tilde{s}^n, u^n} p_{\tilde{S}^n}(\tilde{s}^n) \Pr \left\{ U^n[\ell^*] = u^n \mid (k^* = k, \ell^* = \ell), \right. \\
& \left. \tilde{S}^n[k^*] = \tilde{s}^n, \mathbf{J}(k^*, \ell^*) = 1 \right\} \\
& = \frac{2\sqrt{\varepsilon}}{\left(1 - \varepsilon - \sqrt{\varepsilon} - \varepsilon^{\frac{1}{4}}\right)}, \tag{27}
\end{aligned}$$

where a follows from (25); b follows from (26) and from the fact that given $(U^n[\ell^*] = u^n, \tilde{S}^n[k^*] = \tilde{s}^n)$, (k^*, ℓ^*) is deterministic, thus, $X^n \mid \{U^n[\ell^*] = u^n, \tilde{S}^n[k^*] = \tilde{s}^n, S^n = s^n\} \sim p_{X^n|U^n \tilde{S}^n S^n}(x^n \mid u^n, \tilde{s}^n, s^n) = p_{X^n|U^n \tilde{S}^n}(x^n \mid u^n, \tilde{s}^n)$ and $U^n[\ell^*]$ is conditionally independent of $(S^n, \mathbf{I}(k^*))$ given $(\tilde{S}^n[k^*] = \tilde{s}^n, \mathbf{J}(k^*, \ell^*) = 1)$; c follows because from the definition of $\rho_{u^n \tilde{s}^n}^{B^n}$ mentioned in (3) and d follows because for $\mathbf{J}(k^*, \ell^*) = 1$, we have $\text{Tr} \left[\left(\mathbb{I} - \Lambda_{(U^n[\ell^*]=u^n)} \right) \rho_{(U^n[\ell^*]=u^n, \tilde{S}^n[k^*]=\tilde{s}^n)}^{B^n} \right] < \sqrt{\varepsilon}$.

$4\mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} \sum_{\ell' \neq \ell^*} [\mathbf{I}\{\mathcal{E}_1^c\} \mathbf{I}\{\mathcal{E}_2^c\}] \text{Tr} \left[\Lambda_{U^n(\ell')} \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*]), S^n}^{B^n} \right]$ is bounded as follows:

$$4\mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} \sum_{\ell' \neq \ell^*} \left[\mathbf{I}\{\mathcal{E}_1^c\} \mathbf{I}\{\mathcal{E}_2^c\} \right]$$

$$\begin{aligned}
& \text{Tr} \left[\Lambda_{U^n(\ell')} \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*]), S^n}^{B^n} \right] \\
& \stackrel{a}{=} 4\mathbb{E}_{(k^*, \ell^*, U^n[\ell^*], \tilde{S}^n[k^*], X^n, S^n)} \sum_{\ell' \neq \ell^*} \sum_{k, \ell} \mathbf{I}\{k^* = k\} \\
& \mathbf{I}\{k^* = k, \ell^* = \ell\} \text{Tr} \left[\Lambda_{U^n(\ell')} \rho_{X^n(U^n[\ell^*], \tilde{S}^n[k^*]), S^n}^{B^n} \right] \\
& \stackrel{b}{\leq} 4 \sum_{k, \ell, \ell' \neq \ell} \mathbb{E}_{(U^n[\ell], U^n[\ell']|X^n, S^n)} \mathbf{I}(k) \mathbf{J}(k, \ell) \\
& \text{Tr} \left[\Lambda_{U^n(\ell')} \rho_{X^n(U^n[\ell], \tilde{S}^n[k]), S^n}^{B^n} \right] \\
& = 4 \sum_{k, \ell' \neq \ell} \sum_{(u^n, u^n, \tilde{s}^n, x^n, s^n)} p_{U^n}(u^n) p_{U^n}(u^n) p_{\tilde{S}^n}(\tilde{s}^n) p_{S^n}(s^n) \\
& p_{X^n|U^n \tilde{S}^n}(x^n \mid u^n, \tilde{s}^n) \Pr \left\{ \mathbf{I}(k) = 1 \mid S^n = s^n, \right. \\
& \left. \tilde{S}^n[k] = \tilde{s}^n \right\} \Pr \left\{ \mathbf{J}(k, \ell) = 1 \mid \mathbf{I}(k) = 1, S^n = s^n, \right. \\
& \left. U^n[\ell] = u^n, \tilde{S}^n[k] = \tilde{s}^n \right\} \text{Tr} \left[\Lambda_{u^n} \rho_{X^n=x^n, S^n=s^n}^{B^n} \right] \\
& \stackrel{c}{\leq} 4 \sum_{k, \ell, \ell' \neq \ell} \sum_{(u^n, u^n, \tilde{s}^n, x^n, s^n)} p_{U^n}(u^n) p_{U^n}(u^n) p_{\tilde{S}^n}(\tilde{s}^n) p_{S^n}(s^n) \\
& p_{X^n|U^n \tilde{S}^n}(x^n \mid u^n, \tilde{s}^n) 2^{-n(\bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \gamma)} 2^{-n(\bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] + \gamma)} \\
& \frac{p_{S^n \tilde{S}^n}(s^n, \tilde{s}^n)}{p_{S^n}(s^n) p_{\tilde{S}^n}(\tilde{s}^n)} \frac{p_{U^n \tilde{S}^n}(u^n, \tilde{s}^n)}{p_{U^n}(u^n) p_{\tilde{S}^n}(\tilde{s}^n)} \text{Tr} \left[\Lambda_{u^n} \rho_{X^n=x^n, S^n=s^n}^{B^n} \right] \\
& \stackrel{d}{=} 4 \times 2^{-n(\bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] + \bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + 2\gamma)} \sum_{k, \ell, \ell' \neq \ell} \sum_{u^n, \tilde{s}^n} p_{U^n \tilde{S}^n}(u^n \tilde{s}^n) \\
& \sum_{u^n} p_{U^n}(u^n) \text{Tr} \left[\Lambda_{u^n} \rho_{u^n, \tilde{s}^n}^{B^n} \right] \\
& \stackrel{e}{=} 4 \times 2^{-n(\bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] + \bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + 2\gamma)} \sum_{k, \ell, \ell' \neq \ell} \sum_{u^n, \tilde{s}^n} p_{U^n \tilde{S}^n}(u^n \tilde{s}^n) \\
& \sum_{u^n} p_{U^n}(u^n) \text{Tr} \left[\text{Tr}_{U^n} \left[\Pi^{U^n B^n} (|u^n\rangle\langle u^n| \otimes \mathbb{I}) \right] \rho_{u^n, \tilde{s}^n}^{B^n} \right] \\
& \stackrel{f}{\leq} 4 \times 2^{-n(\bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] + \bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] + \bar{\mathbf{I}}[\mathbf{U}; \mathbf{B}] + \gamma)} 2^{n(R+2r+R_S)} \\
& \stackrel{g}{=} 4 \times 2^{-n\gamma}, \\
& \stackrel{h}{\leq} 4\varepsilon, \tag{28}
\end{aligned}$$

where a follows from the union bound; b follows because $\mathbf{I}\{k^* = k\} = \mathbf{I}\{k^* = k\} \mathbf{I}(k) \leq \mathbf{I}(k)$ and $\mathbf{I}\{k^* = k, \ell^* = \ell\} = \mathbf{I}\{k^* = k, \ell^* = \ell\} \mathbf{J}(k, \ell) \leq \mathbf{J}(k, \ell)$, c follows from the definition of $\mathbf{I}(k)$ and $\mathbf{J}(k, \ell)$, d follows from the definition of $\rho_{u^n, \tilde{s}^n}^{B^n}$ mentioned in (3), e follows from the definition of Λ_{u^n} mentioned in (8), f follows from the definition of $\Pi^{U^n B^n}$ mentioned in (7) and because k ranges over $[1 : 2^{nR_S}]$, ℓ ranges over $[1 : 2^{nr}]$ and ℓ' ranges over $[1 : 2^{n(R+r)}]$ and g follows because of our choice of R, r and R_S and h follows under the assumption that n is large enough such that $2^{-n\gamma} \leq \varepsilon$. Thus, it now follows from (16), (20), (21), (22), (27) and (28) that

$$\begin{aligned}
& \Pr \{\tilde{m} \neq 1\} \\
& \leq 6\varepsilon + 3\sqrt{\varepsilon} + 3\varepsilon^{\frac{1}{4}} + 3\exp(-2^{n\gamma})
\end{aligned}$$

$$+ \frac{2\sqrt{\varepsilon}}{\left(1 - \varepsilon - \sqrt{\varepsilon} - \varepsilon^{\frac{1}{4}}\right)}.$$

This completes the proof for achievability.

B. Converse

The proof in this section is motivated by [16]. Suppose (R, R_S) be an achievable rate pair. It then follows from Definition 4 that there exists an $(n, M_n, M_{e,n}, \varepsilon_n)$ code such that $R \leq \liminf_{n \rightarrow \infty} \frac{\log M_n}{n}$ and $R_S \geq \limsup_{n \rightarrow \infty} \frac{\log M_{e,n}}{n}$. Let

$$\tilde{S}^n = f_{e,n}(S^n), \quad (29)$$

where $f_{e,n}$ is defined in Definition 3. Also, let U^n represent an arbitrary random variable denoting the uniform choice of a message in $[1 : M_n]$. Notice that the message random variable is independent of S^n . Hence, from the definition of U^n and \tilde{S}^n it now follows that U^n and \tilde{S}^n are independent of each other. Thus,

$$\bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] = 0. \quad (30)$$

Further, notice that in the setting of the problem and for the choice of U^n and \tilde{S}^n fixed above the following classical-quantum state is induced

$$\begin{aligned} & \sigma^{S^n U^n X^n B^n} \\ &= \sum_{(s^n, \tilde{s}^n, u^n, x^n)} p_{S^n}(s^n) p_{\tilde{S}^n | S^n}(\tilde{s}^n | s^n) p_{U^n}(u^n) \\ & \quad p_{X^n | \tilde{S}^n U^n}(x^n | \tilde{s}^n, u^n) |s^n\rangle \langle s^n|^{S^n} \otimes |\tilde{s}^n\rangle \langle \tilde{s}^n|^{S^n} \\ & \quad \otimes |u^n\rangle \langle u^n|^{U^n} \otimes |x^n\rangle \langle x^n|^{X^n} \otimes \rho_{x^n, s^n}^{B^n}. \end{aligned} \quad (31)$$

We will now first prove the lower bound on R_S . Towards this notice that from (29) and from the definition of $f_{e,n}$ it follows that the cardinality of the set over which the random variable \tilde{S}^n takes values cannot be larger than $M_{e,n}$. Thus, it now follows that from [21, Lemma 2.6.2] that

$$\Pr \left\{ \frac{1}{n} \log \frac{1}{p_{\tilde{S}^n}} \geq \frac{1}{n} \log M_{e,n} + \gamma \right\} \leq 2^{-n\gamma}, \quad (32)$$

where $\gamma > 0$ is an arbitrary constant. Furthermore, since $\frac{1}{n} \log \frac{p_{\tilde{S}^n | S^n}}{p_{\tilde{S}^n}} \leq \frac{1}{n} \log \frac{1}{p_{\tilde{S}^n}}$ it now follows from (32) that

$$\Pr \left\{ \frac{1}{n} \log \frac{p_{\tilde{S}^n | S^n}}{p_{\tilde{S}^n}} \geq \frac{1}{n} \log M_{e,n} + \gamma \right\} \leq 2^{-n\gamma}. \quad (33)$$

Thus, from the Definition 1 and (33) it now follows that there exists an n_0 such that for every $n > n_0$, we have $\bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] \leq \frac{1}{n} \log M_{e,n} + \gamma$. This further implies that

$$\begin{aligned} \bar{\mathbf{I}}[\mathbf{S}; \tilde{\mathbf{S}}] &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log M_{e,n} + 2\gamma \\ &\leq R_S + 2\gamma, \end{aligned}$$

where the last inequality follows from the definition of R_S .

We now prove the upper bound on R . Let $\rho_{u^n}^{B^n} = \sum_{(s^n, \tilde{s}^n, x^n)} p_{S^n}(s^n) p_{\tilde{S}^n | S^n}(\tilde{s}^n | s^n) p_{X^n | U^n \tilde{S}^n}(x^n | u^n, \tilde{s}^n) \rho_{x^n, s^n}^{B^n}$ and $\rho^{B^n} = \mathbb{E} \left[\rho_{U^n}^{B^n} \right]$, where U^n as mentioned above is uniformly distributed over the set $[1 : M_n]$ and \tilde{S}^n is

as defined in (29). Fix $\gamma > 0$. It now follows from Definition 3 and [9, Lemma 4] that ε_n satisfies the following bound,

$$\varepsilon_n \geq \sum_{u^n \in \mathcal{U}^n} p_{U^n}(u^n) \text{Tr} \left[\rho_{u^n}^{B^n} \left\{ \rho_{u^n}^{B^n} \preceq 2^{n(\frac{1}{n} \log M_n - \gamma)} \rho^{B^n} \right\} \right] - 2^{-n\gamma}. \quad (34)$$

Letting $\frac{1}{n} \log M_n = \mathbf{I}[\mathbf{U}; \mathbf{B}] + 2\gamma$ in (34) it holds that

$$\varepsilon_n \geq \sum_{u^n \in \mathcal{U}^n} p_{U^n}(u^n) \text{Tr} \left[\rho_{u^n}^{B^n} \left\{ \rho_{u^n}^{B^n} \preceq 2^{n(\mathbf{I}[\mathbf{U}; \mathbf{B}] + \gamma)} \rho^{B^n} \right\} \right] - 2^{-n\gamma}. \quad (35)$$

However, while $2^{-n\gamma} \rightarrow 0$ as $n \rightarrow \infty$, the definition of $\mathbf{I}[\mathbf{U}; \mathbf{B}]$ implies the existence of $\varepsilon_0 > 0$ and infinitely many n 's satisfying $\sum_{u^n \in \mathcal{U}^n} p_{U^n}(u^n) \text{Tr} \left[\rho_{u^n}^{B^n} \left\{ \rho_{u^n}^{B^n} \prec 2^{n(\frac{1}{n} \log M_n - \gamma)} \rho^{B^n} \right\} \right] > \varepsilon_0$. This further implies that $R \leq \liminf_{n \rightarrow \infty} \frac{\log M_n}{n}$ must satisfy the following for arbitrarily small probability of error:

$$\begin{aligned} R &\leq \mathbf{I}[\mathbf{U}; \mathbf{B}] + 2\gamma \\ &\leq \mathbf{I}[\mathbf{U}; \mathbf{B}] - \bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] + 2\gamma \\ &\leq \sup_{\{\omega_n\}_{n=1}^{\infty}} \left(\mathbf{I}[\mathbf{U}; \mathbf{B}] - \bar{\mathbf{I}}[\mathbf{U}; \tilde{\mathbf{S}}] \right) + 2\gamma, \end{aligned}$$

where the second inequality follows from (30) and in the last inequality the supremum is taken over all sequence of classical-quantum states having the following form for every n ,

$$\begin{aligned} \omega_n := & \sum_{(s^n, \tilde{s}^n, u^n, x^n)} p_{S^n}(s^n) p_{\tilde{S}^n | S^n}(\tilde{s}^n | s^n) p_{U^n | \tilde{S}^n}(u^n | \tilde{s}^n) \\ & p_{X^n | \tilde{S}^n U^n}(x^n | \tilde{s}^n, u^n) |s^n\rangle \langle s^n|^{S^n} \\ & \otimes |\tilde{s}^n\rangle \langle \tilde{s}^n|^{S^n} \otimes |u^n\rangle \langle u^n|^{U^n} \otimes |x^n\rangle \langle x^n|^{X^n} \\ & \otimes \rho_{x^n, s^n}^{B^n} \end{aligned}$$

This completes the proof for the converse.

V. CONCLUSION AND ACKNOWLEDGEMENT

We extended the result of Heggard and El Gamal [17] to the quantum case in the information-spectrum setting. The proof in [17] is based on the covering lemma [19, Lemma 3.3], conditional typicality lemma [19] and Markov lemma [19, Lemma 12.1]. We have shown in this paper that quantum information-spectrum generalization of the result of Heggard and El Gamal [17] can be derived even in the absence of these powerful lemmas. A natural open question which arises from the problem studied in this manuscript is to study the problem of communication when side information is available at the decoder end. This study along with the techniques presented in this paper will then lead to the establishing of the capacity region when the side information is available both at the encoder and decoder end.

This work was supported by the Engineering and Physical Sciences Research Council (Grant No. EP/M013243/1).

REFERENCES

- [1] M. M. Wilde, "From classical to quantum Shannon theory." <http://arxiv.org/abs/1106.1445>, 2011.

- [2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2nd ed., 2006.
- [3] R. M. Gray, *Entropy and Information Theory*. New York, USA: Springer-Verlag, 1990.
- [4] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, pp. 1147–1157, 1994.
- [5] S. Miyake and F. Kanaya, "Coding theorems on correlated general sources," *IEICE Trans. Fundamentals*, vol. E78-A(9), pp. 1063–1070, Sept. 1995.
- [6] K.-I. Iwata and J. Muramatsu, "An information-spectrum approach to rate-distortion function with side information," *IEICE Trans. on Fundamentals of Electronics, Communication and Computers*, vol. E85-A, pp. 1387–1395, June 2002.
- [7] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1562–1575, April 2006.
- [8] M. Bloch and J. N. Laneman, "On the secrecy capacity of arbitrary wiretap channel," in *Proc. Allerton Conf. Commun. Control, Computing*, (Monticello, IL, USA), Sept. 2008.
- [9] H. Nagaoka and M. Hayashi, "An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses," *IEEE Trans. Inf. Theory*, vol. 53, pp. 534–549, Feb. 2007.
- [10] T. Ogawa and H. Nagaoka, "Strong converse and Stein's lemma in quantum hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2428–2433, Nov. 2000.
- [11] M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1753–1768, 2003.
- [12] M. Hayashi, "General formulas for fixed-length entanglement concentration," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1753–1768, May 2003.
- [13] G. Bowen and N. Datta, "Beyond i.i.d. in quantum information theory." <http://arxiv.org/abs/quant-ph/0604013>, Oct. 2006.
- [14] N. Datta and R. Renner, "Smooth Rényi entropies and the quantum information spectrum," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2807–2815, 2009.
- [15] J. Radhakrishnan, P. Sen, and N. Warsi, "One-shot marton inner bound for classical-quantum broadcast channel." <http://arxiv.org/abs/1410.3248>, Oct. 2014.
- [16] V. Y. F. Tan, "A formula for the capacity of the general Gel'fand-Pinsker channel," *IEEE Trans. Commun.*, vol. 62, pp. 1857–1870, June 2014.
- [17] C. Heggard and A. El Gamal, "On the capacity of computer memory with defects," *IEEE Trans. Inf. Theory*, vol. 29, pp. 731–739, Sept. 1983.
- [18] S. Gel'fand and M. Pinsker, "Coding for channel with random parameters," *Prob. of Control and Inf. Th.*, vol. 9, no. 1, pp. 19–31, 1980.
- [19] A. El Gamal and Y. H. Kim, *Network Information Theory*. Cambridge, U.K: Cambridge University Press, 2012.
- [20] L. Wang and R. Renner, "One-shot classical-quantum capacity and hypothesis testing," *Phys. Rev. Lett.*, vol. 108, pp. 200501–200505, May 2012.
- [21] T. S. Han, *Information-Spectrum Methods in Information Theory*. Berlin, Germany: Springer-Verlag, 2003.



Justin P. Coon received a B.S. degree (with distinction) in electrical engineering from the Calhoun Honours College, Clemson University, USA and a Ph.D. in communications from the University of Bristol, UK in 2000 and 2005, respectively. He has worked in research positions in industry and academia, and is currently an Associate Professor with the Department of Engineering Science, Oxford University, and a Tutorial Fellow of Oriel College. Dr. Coon is the recipient of Toshiba's Distinguished Research Award and two "best paper" awards. He has served as an

Editor for the IEEE Transactions on Wireless Communications (2007 - 2013) and the IEEE Transactions on Vehicular Technology (2013 - 2016). Presently, he serves as an Editor for IEEE Wireless Communications Letters (2016 - present). Coon's research interests include communication theory and network theory.



Naqeeb Ahmad Warsi graduated from the Faculty of Engineering and Technology, Jamia Millia Islamia, New Delhi with a B.Tech in Electronics and Communication Engineering in 2006. After that, he worked as a scientist for the Space Application Center (Indian Space Research Organization) until 2009. He then went on to obtain a PhD in Computer and System science from the Tata Institute of Fundamental Research at Mumbai in 2015. After receiving his PhD he worked as a Postdoctoral Researcher in the Department of Engineering Science at the University

of Oxford. Currently, he is an Assistant Professor at IITD, Delhi and is a research fellow jointly with Nanyang Technological University and National University of Singapore.