10-22-2012

Partial decode-forward for quantum relay channels

Ivan Savov
Université McGill

Mark M. Wilde
Université McGill

Mai Vu
Université McGill

Recommended Citation

Savov, I., Wilde, M., & Vu, M. (2012). Partial decode-forward for quantum relay channels. *IEEE International Symposium on Information Theory - Proceedings*, 731-735. https://doi.org/10.1109/ISIT.2012.6284655

Partial decode-forward for quantum relay channels

Ivan Savov*, Mark M. Wilde* and Mai Vu†
* School of Computer Science and † Electrical and Computer Engineering Department, McGill University, *Montréal, Canada*

***Abstract*—A relay channel is one in which a Source and Destination use an intermediate Relay station in order to improve communication rates. We propose the study of relay channels with classical inputs and quantum outputs and prove that a "partial decode and forward" strategy is achievable. We divide the channel uses into many blocks and build codes in a randomized, block-Markov manner within each block. The Relay performs a standard Holevo-Schumacher-Westmoreland quantum measurement on each block in order to decode part of the Source's message and then forwards this partial message in the next block. The Destination performs a novel "sliding-window" quantum measurement on two adjacent blocks in order to decode the Source's message. This strategy achieves non-trivial rates for classical communication over a quantum relay channel.**

I. Introduction

Suppose that a Source wishes to communicate with a remote Destination. Suppose further that a Relay is available that can decode the messages transmitted by the Source during one time slot and *forward* them to the Destination during the next time slot. With the Relay's help, the Source and Destination can improve communication rates because the Destination can decode the intended messages in parallel from the channel outputs at two consecutive time slots. In this way, useful information is received both from the Source and the Relay.

The relay channel has been studied extensively in the context of classical information theory [1], [2], [3]. There, the discrete memoryless relay channel is modelled as a conditional probability distribution $p(y_1, y|x, x_1)$, where y_1 and y are the respective outputs at the Relay and Destination whenever the Source and Relay input symbols x and x_1. Two important families of coding strategies exist for relay channels: compress-and-forward and decode-and-forward [1], [3]. The partial decode-and-forward strategy differs from the decode-and-forward strategy in that it has the Relay decode only *part* of the message from the Source [1].

The study of quantum channels with information-theoretic techniques has been an active area for some time now [4]. Theoretical interest has focused on classical-quantum channels of the form $(\mathcal{X},\ \mathcal{N}^{X\to B}(x)\equiv\rho_x^B,\ \mathcal{H}^B)$, where, for each of the inputs $x \in \mathcal{X}$, there corresponds an output quantum state, described by a density operator ρ_x^B in a finite-dimensional Hilbert space $\mathcal{H}^B$. Classical-quantum channels are a useful abstraction for studying general quantum channels and correspond to the transmitters being restricted to classical encodings. In this setting, single-letter formulas characterize the capacity of point-to-point [5], [6] and multiple-access channels [7] and give achievable rates for other network channels [8], [9], [10].

The study of quantum channels finds practical applications in optical communications. Bosonic channels model the quantum aspects of optical communication channels, where information is encoded into continuous degrees of freedom. It is known that collective quantum measurements on bosonic-channel outputs outperform classical strategies, particularly in the low-photon-number regime [11]. In other words, quantum measurements are *necessary* to achieve their ultimate capacity. Ref. [11] also demonstrates that classical encoding is *sufficient* to achieve the Holevo capacity of the lossy bosonic channel, giving further motivation for the theoretical study of classical-quantum models.

In this paper, we develop a "partial decode and forward" strategy for classical-quantum relay channels. Our results here are the first extension of the quantum simultaneous decoding techniques used in [8], [9] to multi-hop networks. In the partial-decode-and-forward strategy given here, the Relay decodes part of the Source's message in one block and forwards it in the next. The Destination performs a novel "sliding-window" quantum measurement to decode both parts of the Source's message in two consecutive blocks [12], [2] and in doing so allows for the Source and Destination to achieve non-trivial communication rates. We state our main result in the Section II, introduce the necessary background on quantum systems and quantum decoding in Section III, and give the proof in Section IV. We conclude and discuss open problems in Section V.

II. Results

A classical-quantum relay channel $\mathcal{N}$ is a map with two classical inputs x and x_1 and two output quantum systems B_1 and B. For each pair of possible input symbols $(x, x_1) \in \mathcal{X} \times \mathcal{X}_1$, the channel prepares a density operator $\rho_{x,x_1}^{B_1B}$ defined on the tensor-product Hilbert space $\mathcal{H}^{B_1} \otimes \mathcal{H}^B$:

$$\rho_{x,x_1}^{B_1B} \equiv \mathcal{N}^{XX_1\to B_1B}(x, x_1), \tag{1}$$

where B_1 is the Relay output and B is the Destination output.

The theorem below captures the main result of our paper:

Theorem 1 (Partial decode-forward inner bound)**.** *Let* $\{\rho_{x,x_1}\}$ *be a cc-qq relay channel as in* (1)*. Then a rate* R *is achievable, provided that the following inequality holds:*

$$R \leq \max_{p(u,x,x_1)} \min \left\{ \begin{array}{c} I(XX_1;B)_\theta\,, \\ I(U;B_1|X_1)_\theta + I(X;B|X_1U)_\theta \end{array} \right\}, \tag{2}$$

where the information quantities are with respect to the classical-quantum state $\theta^{UXX_1B_1B} \equiv$

$$\sum_{x,u,x_1} p\,(u,x,x_1)\,|u\rangle\langle u|^U \otimes |x\rangle\langle x|^X \otimes |x_1\rangle\langle x_1|^{X_1} \otimes \rho_{x,x_1}^{B_1B}. \tag{3}$$

Our code construction employs codebooks $\{x_1^n\}$, $\{u^n\}$, and $\{x^n\}$ generated according to the distribution $p(x_1)p(u|x_1)p(x|u,x_1)$. We split the message for each block into two parts $(m,\ell) \in \mathcal{M} \times \mathcal{L}$ such that the rate $R = R_m + R_\ell$. The Relay fully decodes the message ℓ and re-encodes it directly in the next block (without using binning). The Destination exploits a "sliding-window" decoding strategy [12], [2] by performing a collective measurement on two consecutive blocks. In this approach, the message pair (m_j, ℓ_j) sent during block j is decoded from the outputs of blocks j and $j+1$, using an "AND-measurement."

III. Preliminaries

In this section, we introduce the notation used in our paper and some background information on quantum decoding.

1) Quantum systems: We denote quantum systems as B_1 and B and the corresponding Hilbert spaces as $\mathcal{H}^{B_1}$ and $\mathcal{H}^B$. We represent quantum states of a system B with a density operator ρ^B, which is a positive semi-definite operator with unit trace. Let $H(B)_\rho \equiv -\mathrm{Tr}\left[\rho^B \log_2 \rho^B\right]$ denote the von Neumann entropy of the state ρ^B. In order to describe the "distance" between two quantum states, we use the notion of *trace distance*. The trace distance between states σ and ρ is $\|\sigma - \rho\|_1 = \mathrm{Tr}|\sigma - \rho|$, where $|X| = \sqrt{X^\dagger X}$ [4]. Two states can *substitute* for one another up to a penalty proportional to the trace distance between them:

Lemma 1. *Let $0 \leq \rho, \sigma, \Lambda \leq I$. Then*

$$\mathrm{Tr}\left[\Lambda\rho\right] \leq \mathrm{Tr}\left[\Lambda\sigma\right] + \left\|\rho - \sigma\right\|_1. \tag{4}$$

Proof: This follows from a variational characterization of trace distance as the distinguishability of the states under an optimal measurement M [4]: $\|\rho - \sigma\|_1 = 2\max_{0\leq M\leq I} \mathrm{Tr}\left[M(\rho - \sigma)\right]$. ∎

2) Quantum decoding: In a communication scenario, the decoding operations performed by the receivers correspond to quantum measurements on the outputs of the channel. A quantum measurement is a positive operator-valued measure (POVM) $\{\Lambda_m\}_{m\in\mathcal{M}}$ on the system B^n. To be a valid POVM, the set $\{\Lambda_m\}$ of $|\mathcal{M}|$ operators should all be positive semi-definite and sum to the identity: $\Lambda_m \geq 0,\ \sum_m \Lambda_m = I$.

Suppose we are given positive operators $\{P_m\}_{m\in\mathcal{M}}$ that are apt at detecting ($\mathrm{Tr}[P_m\ \rho_m] \geq 1-\epsilon$) and distinguishing ($\mathrm{Tr}[P_m\ \rho_{m'\neq m}] \leq \epsilon$) the output states produced by each message. We can construct a valid POVM (known as the square-root measurement [5], [6]) by *normalizing* these operators:

$$\Lambda_m \equiv \left(\sum_k P_k\right)^{-1/2} P_m \left(\sum_k P_k\right)^{-1/2}. \tag{5}$$

The error analysis of a square-root measurement is greatly simplified by using the Hayashi-Nagaoka operator inequality.

Lemma 2 (Hayashi-Nagaoka [13])**.** *If S and T are operators such that $0 \leq T$ and $0 \leq S \leq I$, then*

$$I - (S+T)^{-\frac{1}{2}} S (S+T)^{-\frac{1}{2}} \ \leq\ 2(I-S) \ +\ 4T. \tag{6}$$

3) Error analysis: In the context of our coding strategy, we analyze the average probability of error at the Relay:

$$\bar{p}_e^R \equiv \frac{1}{|\mathcal{L}|} \sum_{\ell_j} \mathrm{Tr}\Big\{\Big(I - \Gamma_{\ell_j}^{B^n_{1(j)}}\Big)\, \rho_{\ell_j}^{B^n_{1(j)}}\Big\},$$

and the average probability of error at the Destination:

$$\bar{p}_e^D \equiv \frac{1}{|\mathcal{M}||\mathcal{L}|} \sum_{m_j,\ell_j} \mathrm{Tr}\Big[\Big(I - \Lambda_{m_j,\ell_j}^{B^n_{(j)}B^n_{(j+1)}}\Big)\, \rho_{m_j,\ell_j}^{B^n_{(j)}B^n_{(j+1)}}\Big]. \tag{7}$$

The operators $\left(I - \Gamma_{\ell_j}\right)$ and $\left(I - \Lambda_{m_j,\ell_j}\right)$ correspond to the complements of the correct decoding outcomes.

Definition 1. *An (n, R, ϵ) partial-decode-and-forward code for the quantum relay channel consists of two codebooks $\{x^n(m_j,\ell_j)\}_{m_j\in\mathcal{M},\ell_j\in\mathcal{L}}$ and $\{x_1^n(\ell_j)\}_{\ell_j\in\mathcal{L}}$ and decoding POVMs $\left\{\Gamma_{\ell_j}\right\}_{\ell_j\in\mathcal{L}}$ and $\left\{\Lambda_{m_j,\ell_j}\right\}_{m_j\in\mathcal{M},\ell_j\in\mathcal{L}}$ such that the average probability of error is bounded from above as $\bar{p}_e = \bar{p}_e^R + \bar{p}_e^D \leq \epsilon$.*

A rate R is *achievable* if there exists an $(n, R-\delta, \epsilon)$ quantum relay channel code for all $\epsilon, \delta > 0$ and sufficiently large n.

IV. Achievability proof

The channel is used for b blocks, each indexed by $j \in \{1,\ldots,b\}$. Our error analysis shows that:

- The Relay can decode the message ℓ_j during block j.
- The Destination can simultaneously decode (m_j, ℓ_j) from a collective measurement on the output systems of blocks j and $j+1$.

The error analysis at the Relay is similar to that of the Holevo-Schumacher-Westmoreland theorem [5], [6]. The message ℓ_j can be decoded reliably, if the rate R_ℓ obeys the following inequality:

$$R_\ell \leq I\left(U; B_1|X_1\right)_\theta. \tag{8}$$

We give a proof in the Appendix.

The decoding at the Destination is a variant of the quantum simultaneous decoder from [8], [9]. To decode the message (m_j, ℓ_j), the Destination performs a "sliding-window" decoder, implemented as an "AND-measurement" on the outputs of blocks j and $j+1$. This coding technique does not require binning at the Relay or backwards decoding at the Destination [12], [2].

In this section, we give the details of the coding strategy and analyze the probability of error at the Destination.

Codebook construction. Fix a distribution $p(u,x,x_1)$ and independently generate a different codebook for each block j:

- Randomly and independently generate 2^{nR_ℓ} sequences $x_1^n(\ell_{j-1})$, $\ell_{j-1} \in \left[1 : 2^{nR_\ell}\right]$, according to $\prod_{i=1}^n p(x_{1i})$.
- For each $x_1^n(\ell_{j-1})$, randomly and conditionally independently generate 2^{nR_ℓ} sequences $u^n(\ell_j|\ell_{j-1})$, $\ell_j \in \left[1 : 2^{nR_\ell}\right]$ according to $\prod_{i=1}^n p\left(u_i|x_{1i}(\ell_{j-1})\right)$.

- For each $x_1^n(\ell_{j-1})$ and each corresponding $u^n(\ell_j|\ell_{j-1})$, randomly and conditionally independently generate 2^{nR_m} sequences $x^n(m_j|\ell_j,\ell_{j-1})$, $m_j \in \left[1:2^{nR_m}\right]$, according to the distribution: $\prod_{i=1}^{n} p\big(x_i|x_{1i}(\ell_{j-1}), u_i(\ell_j|\ell_{j-1})\big)$.

Transmission. The transmission of (ℓ_j, m_j) to the Destination happens during blocks j and $j+1$. At the beginning of block j, we assume that the Relay has correctly decoded the message ℓ_{j-1}. During block j, the Source inputs the new messages m_j and ℓ_j, and the Relay forwards the old message ℓ_{j-1}. That is, their inputs to the channel for block j are the codewords $x^n(m_j,\ell_j,\ell_{j-1})$ and $x_1^n(\ell_{j-1})$, leading to the following state at the channel outputs:

$$\rho^{(j)}_{m_j,\ell_j,\ell_{j-1}} \equiv \rho^{B^n_{1(j)}B^n_{(j)}}_{x^n(m_j,\ell_j,\ell_{j-1}),x_1^n(\ell_{j-1})}.$$

During block $j+1$, the Source transmits (m_{j+1},ℓ_{j+1}) given ℓ_j, whereas the Relay sends ℓ_j, leading to the state:

$$\rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j} \equiv \rho^{B^n_{1(j+1)}B^n_{(j+1)}}_{x^n(m_{j+1},\ell_{j+1},\ell_j),x_1^n(\ell_j)}.$$

Our shorthand notation is such that the states are identified by the messages that they encode, and the codewords are implicit.

Decoding at the Destination. We now determine a decoding POVM that the Destination can perform on the output systems spanning blocks j and $j+1$. The Destination is trying to recover messages ℓ_j and m_j given knowledge of ℓ_{j-1}.

First let us consider forming decoding operators for block $j+1$. Consider the state obtained by tracing over the systems X, U, and B_1 in (3):

$$\theta^{X_1B} = \sum_{x_1} p\left(x_1\right) |x_1\rangle\langle x_1|^{X_1} \otimes \tau^B_{x_1},$$

where $\tau^B_{x_1} \equiv \sum_{u,x} p\left(u|x_1\right) p\left(x|x_1,u\right) \rho^B_{x,x_1}$. Also, let $\bar{\tau}^B$ denote the following state: $\bar{\tau}^B \equiv \sum_{x_1} p\left(x_1\right) \tau^B_{x_1}$. Corresponding to the above states are conditionally typical projectors [4] of the following form:

$$\Pi^{(j+1)}_{\tau_{\ell_j}} \equiv \Pi^{B^n_{(j+1)}}_{\tau_{x_1^n(\ell_j)}}, \qquad \Pi^{(j+1)}_{\bar{\tau}} \equiv \Pi^{B^n_{(j+1)}}_{\bar{\tau}},$$

which we combine to form the positive operator:

$$P^{B^n_{(j+1)}}_{\ell_j|\ell_{j-1}} \equiv \Pi^{(j+1)}_{\bar{\tau}} \ \Pi^{(j+1)}_{\tau_{\ell_j}} \ \Pi^{(j+1)}_{\bar{\tau}}, \tag{9}$$

that acts on the output systems $B^n_{(j+1)}$ of block $j+1$.

Let us now form decoding operators for block j. Define the conditional typical projector for the state $\rho^{(j)}_{m_j,\ell_j,\ell_{j-1}}$ as

$$\Pi^{(j)}_{\rho_{m_j,\ell_j|\ell_{j-1}}} \equiv \ \Pi^{B^n_{(j)}}_{\rho_{x^n(m_j,l_j,l_{j-1}),x_1^n(l_{j-1})}}. \tag{10}$$

The state obtained from (3) by tracing over X and B_1 is

$$\theta^{UX_1B} = \sum_{u,x_1} p\left(u|x_1\right) p\left(x_1\right) |u\rangle\langle u|^U \otimes |x_1\rangle\langle x_1|^{X_1} \otimes \bar{\rho}^B_{u,x_1},$$

where $\bar{\rho}^B_{u,x_1} \equiv \sum_x p\left(x|x_1,u\right) \rho^B_{x,x_1}$. Define also the doubly averaged state $\bar{\bar{\rho}}^B_{x_1} \equiv \sum_{u,x} p\left(x|x_1,u\right) p\left(u|x_1\right) \rho^B_{x,x_1}$.

The following conditionally typical projectors will be useful in our decoding scheme:

$$\Pi^{(j)}_{\bar{\rho}_{\ell_j|\ell_{j-1}}} \equiv \Pi^{B^n_{(j)}}_{\bar{\rho}_{u^n(l_j,l_{j-1}),x_1^n(l_{j-1})}}, \quad \Pi^{(j)}_{\bar{\bar{\rho}}_{|\ell_{j-1}}} \equiv \Pi^{B^n_{(j)}}_{\bar{\bar{\rho}}_{x_1^n(l_{j-1})}}.$$

We can then form a positive operator "sandwich":

$$P^{B^n_{(j)}}_{m_j,\ell_j|\ell_{j-1}} \equiv \Pi^{(j)}_{\bar{\bar{\rho}}_{|\ell_{j-1}}} \Pi^{(j)}_{\bar{\rho}_{\ell_j|\ell_{j-1}}} \Pi^{(j)}_{\rho_{m_j,\ell_j|\ell_{j-1}}} \Pi^{(j)}_{\bar{\rho}_{\ell_j|\ell_{j-1}}} \Pi^{(j)}_{\bar{\bar{\rho}}_{|\ell_{j-1}}}. \tag{11}$$

Finally, we combine the positive operators from (9) and (11) to form the "sliding-window" positive operator:

$$P^{B^n_{(j)}B^n_{(j+1)}}_{m_j,\ell_j|\ell_{j-1}} = P^{B^n_{(j)}}_{m_j,\ell_j|\ell_{j-1}} \otimes P^{B^n_{(j+1)}}_{\ell_j|\ell_{j-1}}, \tag{12}$$

from which we can build the Destination's square-root measurement $\Lambda^{B^n_{(j)}B^n_{(j+1)}}_{m_j,\ell_j|\ell_{j-1}}$ using the formula in (5). This measurement is what we call the "AND-measurement."

Error analysis at the Destination. In this section, we prove that the Destination can correctly decode the message pair (m_j,ℓ_j) by employing the measurement $\{\Lambda^{B^n_{(j)}B^n_{(j+1)}}_{m_j,\ell_j|\ell_{j-1}}\}$ on the output state $\rho^{(j)}_{m_j\ell_j\ell_{j-1}} \otimes \rho^{(j+1)}_{m_{j+1}\ell_{j+1}\ell_j}$ spanning blocks j and $j+1$. The average probability of error for the Destination is given in (7). For now, we consider the error analysis for a single message pair (m_j,ℓ_j):

$$\begin{aligned}
\bar{p}^D_e \equiv & \operatorname{Tr}\left[\left(I - \Lambda^{B^n_{(j)}B^n_{(j+1)}}_{m_j,\ell_j|\ell_{j-1}}\right) \rho^{(j)}_{m_j\ell_j\ell_{j-1}} \otimes \rho^{(j+1)}_{m_{j+1}\ell_{j+1}\ell_j}\right]. \\
& \leq 2 \operatorname{Tr}\left\{\left(I - P^{B^n_{(j)}B^n_{(j+1)}}_{m_j,\ell_j|\ell_{j-1}}\right) \ \rho^{(j)}_{m_j\ell_j\ell_{j-1}} \otimes \rho^{(j+1)}_{m_{j+1}\ell_{j+1}\ell_j}\right\} \\
& \quad + 4 \sum_{(\ell'_j,m'_j)\neq(\ell_j,m_j)} \operatorname{Tr}\left\{P^{B^n_{(j)}B^n_{(j+1)}}_{m'_j,\ell'_j|\ell_{j-1}} \ \rho^{(j)}_{m_j\ell_j\ell_{j-1}} \otimes \rho^{(j+1)}_{m_{j+1}\ell_{j+1}\ell_j}\right\},
\end{aligned}$$

where use the Hayashi-Nagaoka inequality (Lemma 2) to decompose the error operator $(I - \Lambda^{B^n_{(j)}B^n_{(j+1)}}_{m_j,\ell_j|\ell_{j-1}})$ into two components: (I) a term corresponding to the probability that the correct detector does not "click": $(I - P^{B^n_{(j)}B^n_{(j+1)}}_{m_j,\ell_j|\ell_{j-1}})$, and (II) another term corresponding to the probability that a wrong detector "clicks": $\sum_{(\ell'_j,m'_j)} P^{B^n_{(j)}B^n_{(j+1)}}_{m'_j,\ell'_j|\ell_{j-1}}$. These two errors are analogous to the classical error events in which an output sequence y^n is either not jointly typical with the correct codeword or is jointly typical with another codeword.

We will bound the expectation of the average probability of error $\mathbb{E}_{U^nX^nX_1^n}\{\bar{p}^D_e\}$, using the properties of typical projectors [4], and the following lemmas:

Lemma 3. *For any operators $0 \leq P^A, Q^B \leq I$, we have:*

$$(I^{AB} - P^A \otimes Q^B) \leq (I^A - P^A) \otimes I^B + I^A \otimes (I^B - Q^B).$$

Proof: Expand and rearrange $(I-P)\otimes(I-Q) \geq 0$. ■

Lemma 4 (Gentle Operator Lemma for Ensembles [14])**.** *Let $\{p(x), \rho_x\}$ be an ensemble and let $\bar{\rho} \equiv \sum_x p(x)\,\rho_x$. If an operator Λ, where $0 \leq \Lambda \leq I$, has high overlap with the average state, $\operatorname{Tr}[\,\Lambda\,\bar{\rho}\,] \geq 1-\epsilon$, then the subnormalized state $\sqrt{\Lambda}\rho_x\sqrt{\Lambda}$ is close in trace distance to the original state ρ_x on average: $\mathbb{E}_X\left\{\left\|\sqrt{\Lambda}\rho_X\sqrt{\Lambda} - \rho_X\right\|_1\right\} \leq 2\sqrt{\epsilon}$.*

The first term (I) is bounded as follows:

$$\begin{aligned}
&\mathrm{Tr}\Big[\Big(I - P^{B^n_{(j)}B^n_{(j+1)}}_{m_j,\ell_j|\ell_{j-1}}\Big)\ \rho^{(j)}_{m_j\ell_j\ell_{j-1}}\otimes\rho^{(j+1)}_{m_{j+1}\ell_{j+1}\ell_j}\Big]\\
&= \mathrm{Tr}\Big[\Big(I - P^{B^n_{(j)}}_{m_j,\ell_j|\ell_{j-1}}\otimes P^{B^n_{(j+1)}}_{\ell_j|\ell_{j-1}}\Big)\rho^{(j)}_{m_j\ell_j\ell_{j-1}}\otimes\rho^{(j+1)}_{m_{j+1}\ell_{j+1}\ell_j}\Big]\\
&\le \underbrace{\mathrm{Tr}\Big[\Big(I - P^{B^n_{(j)}}_{m_j,\ell_j|\ell_{j-1}}\Big)\rho^{(j)}_{m_j,\ell_j,\ell_{j-1}}\Big]}_{\alpha}\underbrace{\mathrm{Tr}\Big[\rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big]}_{=1}\\
&\quad+\underbrace{\mathrm{Tr}\Big[\rho^{(j)}_{m_j,\ell_j,\ell_{j-1}}\Big]}_{=1}\underbrace{\mathrm{Tr}\Big[\Big(I - P^{B^n_{(j+1)}}_{\ell_j|\ell_{j-1}}\Big)\rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big]}_{\beta},
\end{aligned}$$

where the inequality follows from Lemma 3.

We proceed to bound the term β as follows:

$$\begin{aligned}
\beta &= \mathrm{Tr}\Big[\Big(I - P^{B^n_{(j+1)}}_{\ell_j|\ell_{j-1}}\Big)\rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big]\\
&= \mathrm{Tr}\Big[\Big(I - \Pi^{(j+1)}_{\bar\tau}\Pi^{(j+1)}_{\tau_{\ell_j}}\Pi^{(j+1)}_{\bar\tau}\Big)\rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big]\\
&= 1 - \mathrm{Tr}\Big[\Pi^{(j+1)}_{\bar\tau}\Pi^{(j+1)}_{\tau_{\ell_j}}\Pi^{(j+1)}_{\bar\tau}\ \rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big]\\
&\le 1 - \mathrm{Tr}\Big[\Pi^{(j+1)}_{\tau_{\ell_j}}\ \rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big]\\
&\qquad+\Big\|\Pi^{(j+1)}_{\bar\tau}\rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Pi^{(j+1)}_{\bar\tau} - \rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big\|_1,
\end{aligned}$$

where the inequality follows from Lemma 1.

By taking the expectation over the code randomness, we obtain the upper bound:

$$\begin{aligned}
\mathop{\mathbb{E}}_{U^nX^nX_1^n}\{\beta\} &= 1 - \mathop{\mathbb{E}}_{X_1^n}\mathrm{Tr}\left[\Pi^{(j+1)}_{\tau_{\ell_j}}\mathop{\mathbb{E}}_{U^nX^n|X_1^n}\Big\{\rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big\}\right]\\
&\quad+\mathop{\mathbb{E}}_{U^nX^nX_1^n}\Big\|\Pi^{(j+1)}_{\bar\tau}\rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Pi^{(j+1)}_{\bar\tau} - \rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big\|_1\\
&\le 1 - (1-\epsilon) + 2\sqrt{\epsilon}.
\end{aligned}$$

The inequality follows from $\mathbb{E}_{U^nX^n|X_1^n}\Big\{\rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big\} = \tau_{\ell_j}$, the properties of typical projectors [4]: $\mathbb{E}_{X_1^n}\,\mathrm{Tr}[\Pi^{(j+1)}_{\tau_{\ell_j}}\ \tau_{\ell_j}] \ge 1-\epsilon$, $\mathrm{Tr}[\Pi^{(j+1)}_{\bar\tau}\ \bar\tau] \ge 1-\epsilon$ and Lemma 4.

The error term α is bounded in a similar fashion.

We can split the sum in the second type of error, (II), as $\sum_{(\ell'_j,m'_j)\neq(\ell_j,m_j)}(\cdot) = \sum_{m'_j\neq m_j}(\cdot) + \sum_{\ell'_j\neq\ell_j,\ m'_j}(\cdot)$:

$$\begin{aligned}
&\sum_{(\ell'_j,m'_j)\neq(\ell_j,m_j)}\mathrm{Tr}\Big[P^{B^n_{(j)}B^n_{(j+1)}}_{m'_j,\ell'_j|\ell_{j-1}}\ \rho^{(j)}_{m_j\ell_j\ell_{j-1}}\otimes\rho^{(j+1)}_{m_{j+1}\ell_{j+1}\ell_j}\Big]\\
&= \underbrace{\sum_{m'_j\neq m_j}\mathrm{Tr}\Big[P^{B^n_{(j)}B^n_{(j+1)}}_{m'_j,\ell_j|\ell_{j-1}}\ \rho^{(j)}_{m_j\ell_j\ell_{j-1}}\otimes\rho^{(j+1)}_{m_{j+1}\ell_{j+1}\ell_j}\Big]}_{(A)}\\
&\quad+\underbrace{\sum_{l'_j\neq l_j,\ m'_j}\mathrm{Tr}\Big[P^{B^n_{(j)}B^n_{(j+1)}}_{m'_j,\ell'_j|\ell_{j-1}}\ \rho^{(j)}_{m_j\ell_j\ell_{j-1}}\otimes\rho^{(j+1)}_{m_{j+1}\ell_{j+1}\ell_j}\Big]}_{(B)}.
\end{aligned}$$

We now analyze the two terms (A) and (B) separately.

a) Matching ℓ_j, wrong m_j: By performing the error analysis for the case where ℓ_j is decoded correctly, but m_j is decoded incorrectly, we obtain the bound $R_m < I(X;B|UX_1) = H(B|UX_1) - H(B|UXX_1) - \delta$, using the following properties of typical projectors [4]:

$$\Pi^{(j)}_{\rho^{(j)}_{m'_j,\ell_j|\ell_{j-1}}} \le 2^{n[H(B|UXX_1)+\delta]}\rho^{(j)}_{m'_j,\ell_j,\ell_{j-1}}, \quad (13)$$

$$\Pi^{(j)}_{\bar\rho_{\ell_j|\ell_{j-1}}}\bar\rho^{(j)}_{\ell_j,\ell_{j-1}}\Pi^{(j)}_{\bar\rho_{\ell_j|\ell_{j-1}}} \le 2^{-n[H(B|UX_1)-\delta]}\Pi^{(j)}_{\bar\rho_{\ell_j|\ell_{j-1}}}. \quad (14)$$

Consider the first term:

$$\begin{aligned}
(A) &= \sum_{m'_j\neq m_j}\mathrm{Tr}\Big[P^{B^n_{(j)}B^n_{(j+1)}}_{m'_j,\ell_j|\ell_{j-1}}\ \rho^{(j)}_{m_j\ell_j\ell_{j-1}}\otimes\rho^{(j+1)}_{m_{j+1}\ell_{j+1}\ell_j}\Big]\\
&= \sum_{m'_j\neq m_j}\mathrm{Tr}\Big[\Big(P^{B^n_{(j)}}_{m'_j,\ell_j|\ell_{j-1}}\otimes P^{B^n_{(j+1)}}_{\ell_j|\ell_{j-1}}\Big)\rho^{(j)}_{m_j,\ell_j,\ell_{j-1}}\otimes\rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big]\\
&\le \sum_{m'_j\neq m_j}\mathrm{Tr}\Big[P^{B^n_{(j)}}_{m'_j,\ell_j|\ell_{j-1}}\otimes I^{B^n_{(j+1)}}\ \ \rho^{(j)}_{m_j,\ell_j,\ell_{j-1}}\otimes\rho^{(j+1)}_{m_{j+1},\ell_{j+1},\ell_j}\Big]\\
&= \sum_{m'_j\neq m_j}\mathrm{Tr}\Big[P^{B^n_{(j)}}_{m'_j,\ell_j|\ell_{j-1}}\ \ \rho^{(j)}_{m_j,\ell_j,\ell_{j-1}}\Big]\\
&= \sum_{m'_j\neq m_j}\mathrm{Tr}\Big[\Pi^{(j)}_{\bar\rho_{|\ell_{j-1}}}\underbrace{\Pi^{(j)}_{\bar\rho_{\ell_j|\ell_{j-1}}}\overbrace{\Pi^{(j)}_{\rho^{(j)}_{m'_j,\ell_j|\ell_{j-1}}}}^{①}\Pi^{(j)}_{\bar\rho_{\ell_j|\ell_{j-1}}}}_{②}\Pi^{(j)}_{\bar\rho_{|\ell_{j-1}}}\rho^{(j)}_{m_j,\ell_j,\ell_{j-1}}\Big]
\end{aligned}$$

We now upper bound expression ① using (13) and take the conditional expectation with respect to X^n:

$$\mathop{\mathbb{E}}_{X^n|U^nX_1^n}\Big\{\rho^{(j)}_{m'_j,\ell_j,\ell_{j-1}}\Big\} = \bar\rho^{(j)}_{\ell_j,\ell_{j-1}},$$

which is independent of the state $\rho^{(j)}_{m_j,\ell_j,\ell_{j-1}}$ since $m'_j \neq m_j$. The resulting expression in ② has the state $\bar\rho^{(j)}_{\ell_j,\ell_{j-1}}$ sandwiched between its typical projector on both sides, and so we can use (14). After these steps, we obtain the upper bound:

$$\begin{aligned}
&\mathop{\mathbb{E}}_{X^n|U^nX_1^n}\{(A)\} \le 2^{n[H(B|XUX_1)+\delta]}\ 2^{-n[H(B|UX_1)-\delta]}\times\\
&\qquad\mathop{\mathbb{E}}_{X^n|U^nX_1^n}\sum_{m'_j\neq m_j}\mathrm{Tr}\Big[\Pi^{(j)}_{\bar\rho_{|\ell_{j-1}}}\Pi^{(j)}_{\bar\rho_{\ell_j|\ell_{j-1}}}\Pi^{(j)}_{\bar\rho_{|\ell_{j-1}}}\ \rho^{(j)}_{m_j,\ell_j,\ell_{j-1}}\Big]\\
&\le 2^{n[H(B|XUX_1)+\delta]}2^{-n[H(B|UX_1)-\delta]}\sum_{m'_j\neq m_j}\mathrm{Tr}\Big[\ \rho^{(j)}_{m_j,\ell_j,\ell_{j-1}}\ \Big]\\
&\le |\mathcal{M}|\ 2^{-n[I(X;B|UX_1)-2\delta]}. \qquad (15)
\end{aligned}$$

The first inequality follows because each operator inside the trace is positive and less than the identity.

b) Wrong ℓ_j (and thus wrong m_j): We obtain the bound $R \equiv R_\ell + R_m \le I(XX_1;B) = I(X_1;B) + I(UX;B|X_1)$ from the "AND-measurement" and the following inequalities:

$$\mathrm{Tr}[\Pi^{(j+1)}_{\tau_{\ell_j}}] \le 2^{n[H(B|X_1)+\delta]}, \quad (16)$$

$$\Pi^{(j+1)}_{\bar\tau}\ \bar\tau\ \ \Pi^{(j+1)}_{\bar\tau} \le 2^{-n[H(B)-\delta]}\Pi^{(j+1)}_{\bar\tau}, \quad (17)$$

$$\mathrm{Tr}[\Pi^{(j)}_{\rho_{m_j,\ell_j|\ell_{j-1}}}] \le 2^{n[H(B|UXX_1)+\delta]}, \quad (18)$$

$$\Pi^{(j)}_{\bar\rho_{|\ell_{j-1}}}\ \bar\rho^{(j)}_{|\ell_{j-1}}\ \Pi^{(j)}_{\bar\rho_{|\ell_{j-1}}} \le 2^{-n[H(B|X_1)-\delta]}\Pi^{(j)}_{\bar\rho_{|\ell_{j-1}}}. \quad (19)$$

Consider the following term:

$$
\begin{aligned}
(B) &= \sum_{\ell'_j \neq \ell_j, m'_j} \mathrm{Tr}\Big[P^{B^n_{(j)} B^n_{(j+1)}}_{m'_j, \ell'_j | \ell_{j-1}} \quad \rho^{(j)}_{m_j \ell_j \ell_{j-1}} \otimes \rho^{(j+1)}_{m_{j+1} \ell_{j+1} \ell_j} \Big] \\
&= \sum_{\ell'_j \neq \ell_j, m'_j} \mathrm{Tr}\Big[\Big(P^{B^n_{(j)}}_{m'_j, \ell'_j | \ell_{j-1}} \otimes P^{B^n_{(j+1)}}_{\ell'_j | \ell_{j-1}} \Big) \rho^{(j)}_{m_j \ell_j \ell_{j-1}} \otimes \rho^{(j+1)}_{m_{j+1} \ell_{j+1} \ell_j} \Big] \\
&= \sum_{\ell'_j \neq \ell_j, m'_j} \underbrace{\mathrm{Tr}\Big[P^{B^n_{(j)}}_{m'_j, \ell'_j | \ell_{j-1}} \rho^{(j)}_{m_j, \ell_j, \ell_{j-1}} \Big]}_{(B1)} \underbrace{\mathrm{Tr}\Big[P^{B^n_{(j+1)}}_{\ell'_j | \ell_{j-1}} \rho^{(j+1)}_{m_{j+1}, \ell_{j+1}, \ell_j} \Big]}_{(B2)}
\end{aligned}
$$

We want to calculate the expectation of (B) under the code randomness $\mathbb{E}_{U^n X^n X_1^n}$. The random variables in different blocks are independent, and so we can analyze the expectations of the terms $(B1)$ and $(B2)$ separately.

Consider first the calculation in block j, which leads to the following bound on the expectation of $(B1)$:

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{U^n X^n X_1^n} \{(B1)\} = \mathop{\mathbb{E}}_{U^n X^n X_1^n} \Big\{ \mathrm{Tr}\Big[P^{B^n_{(j)}}_{m'_j, \ell'_j | \ell_{j-1}} \rho^{(j)}_{m_j, \ell_j, \ell_{j-1}} \Big] \Big\} \\
&= \mathop{\mathbb{E}}_{U^n X^n X_1^n} \mathrm{Tr} \left[\begin{array}{c} \Pi^{(j)}_{\bar{\rho}_{\ell'_j | \ell_{j-1}}} \Pi^{(j)}_{\rho_{m'_j, \ell'_j | \ell_{j-1}}} \Pi^{(j)}_{\bar{\rho}_{\ell'_j | \ell_{j-1}}} \times \\ \Pi^{(j)}_{\bar{\bar{\rho}}_{|\ell_{j-1}}} \rho^{(j)}_{m_j, \ell_j, \ell_{j-1}} \Pi^{(j)}_{\bar{\bar{\rho}}_{|\ell_{j-1}}} \end{array} \right] \\
&= \mathop{\mathbb{E}}_{X_1^n} \mathrm{Tr} \left[\begin{array}{c} \mathop{\mathbb{E}}_{U^n X^n | X_1^n} \{ \Pi^{(j)}_{\bar{\rho}_{\ell'_j | \ell_{j-1}}} \Pi^{(j)}_{\rho_{m'_j, \ell'_j | \ell_{j-1}}} \Pi^{(j)}_{\bar{\rho}_{\ell'_j | \ell_{j-1}}} \} \times \\ \Pi^{(j)}_{\bar{\bar{\rho}}_{|\ell_{j-1}}} \underbrace{\mathop{\mathbb{E}}_{U^n X^n | X_1^n} \Big\{ \rho^{(j)}_{m_j, \ell_j, \ell_{j-1}} \Big\}}_{③} \Pi^{(j)}_{\bar{\bar{\rho}}_{|\ell_{j-1}}} \end{array} \right] \\
&= \mathop{\mathbb{E}}_{X_1^n} \mathrm{Tr} \left[\begin{array}{c} \mathop{\mathbb{E}}_{U^n X^n | X_1^n} \{ \Pi^{(j)}_{\bar{\rho}_{\ell'_j | \ell_{j-1}}} \Pi^{(j)}_{\rho_{m'_j, \ell'_j | \ell_{j-1}}} \Pi^{(j)}_{\bar{\rho}_{\ell'_j | \ell_{j-1}}} \} \times \\ \underbrace{\Pi^{(j)}_{\bar{\bar{\rho}}_{|\ell_{j-1}}} \bar{\bar{\rho}}^{(j)}_{|\ell_{j-1}} \Pi^{(j)}_{\bar{\bar{\rho}}_{|\ell_{j-1}}}}_{④} \end{array} \right] \\
&\leq 2^{-n[H(B|X_1)-\delta]} \mathop{\mathbb{E}}_{U^n X^n X_1^n} \mathrm{Tr}\Big[\Pi^{(j)}_{\bar{\rho}_{\ell'_j | \ell_{j-1}}} \Pi^{(j)}_{\rho_{m'_j, \ell'_j | \ell_{j-1}}} \Pi^{(j)}_{\bar{\rho}_{\ell'_j | \ell_{j-1}}} \Pi^{(j)}_{\bar{\bar{\rho}}_{|\ell_{j-1}}} \Big] \\
&\leq 2^{-n[H(B|X_1)-\delta]} \mathop{\mathbb{E}}_{U^n X^n X_1^n} \mathrm{Tr}\Big[\Pi^{(j)}_{\rho_{m'_j, \ell'_j | \ell_{j-1}}} \Big] \\
&\leq 2^{-n[H(B|X_1)-\delta]} \mathop{\mathbb{E}}_{U^n X^n X_1^n} 2^{n[H(B|X_1 U X)+\delta]} \\
&= 2^{-n[I(UX;B|X_1)-2\delta]}
\end{aligned}
$$

The result of the expectation in ③ is $\bar{\bar{\rho}}^{(j)}_{|\ell_{j-1}}$, and we can bound the expression in ④ using (19). The first inequality follows because all the other terms in the trace are positive operators less than the identity. The final inequality follows from (18).

Now we consider the expectation of the second term:

$$
\begin{aligned}
\mathop{\mathbb{E}}_{U^n X^n X_1^n} \{(B2)\} &= \mathop{\mathbb{E}}_{U^n X^n X_1^n} \Big\{ \mathrm{Tr}\Big\{ P^{B^n_{(j+1)}}_{\ell'_j | \ell_{j-1}} \ \rho^{(j+1)}_{m_{j+1}, \ell_{j+1}, \ell_j} \Big\} \Big\} \\
&= \mathrm{Tr}\Big\{ P^{B^n_{(j+1)}}_{\ell'_j | \ell_{j-1}} \mathop{\mathbb{E}}_{U^n X^n X_1^n} \Big\{ \rho^{(j+1)}_{m_{j+1}, \ell_{j+1}, \ell_j} \Big\} \Big\} \\
&= \mathrm{Tr}\Big\{ P^{B^n_{(j+1)}}_{\ell'_j | \ell_{j-1}} \ \bar{\tau}^{\otimes n} \Big\} \\
&= \mathrm{Tr}\Big\{ \Pi^{(j+1)}_{\bar{\tau}} \Pi^{(j+1)}_{\tau_{\ell'_j}} \ \Pi^{(j+1)}_{\bar{\tau}} \ \bar{\tau}^{\otimes n} \Big\} \\
&= \mathrm{Tr}\Big\{ \Pi^{(j+1)}_{\tau_{\ell'_j}} \ \Pi^{(j+1)}_{\bar{\tau}} \bar{\tau}^{\otimes n} \Pi^{(j+1)}_{\bar{\tau}} \Big\} \\
&\leq 2^{-n[H(B)-\delta]} \mathrm{Tr}\Big\{ \Pi^{(j+1)}_{\tau_{\ell'_j}} \ \Pi^{(j+1)}_{\bar{\tau}} \Big\} \\
&\leq 2^{-n[H(B)-\delta]} 2^{n[H(B|X_1)+\delta]} = 2^{-n[I(X_1;B)-2\delta]}.
\end{aligned}
$$

Combining the upper bounds on $(B1)$ and $(B2)$ gives our final upper bound:

$$
\begin{aligned}
\mathop{\mathbb{E}}_{U^n X^n X_1^n} \{(B)\} &= \mathop{\mathbb{E}}_{U^n X^n X_1^n} \sum_{\ell'_j \neq \ell_j, m'_j} (B1) \times (B2) \\
&\leq \sum_{\ell'_j \neq \ell_j, \ m'_j} 2^{-n[I(UX;B|X_1)-2\delta]} \times 2^{-n[I(X_1;B)-2\delta]} \\
&\leq |\mathcal{L}||\mathcal{M}| \ 2^{-n[I(X_1;B)+I(UX;B|X_1)-4\delta]}. \qquad (20)
\end{aligned}
$$

By choosing the size of message sets to satisfy equations (8), (15) and (20), the expectation of the average probability of error becomes arbitrarily small for n sufficiently large. ■

V. Discussion

We proved the achievability of the rates given by the partial decode and forward inner bound, thus extending the study of classical-quantum channels to multi-hop scenarios. An interesting open question is to determine a compress-and-forward strategy for the quantum setting. Another avenue for research would be to consider *quantum* communication scenarios, and results here might have applications for the design of quantum repeaters [15].

I. Savov acknowledges support from FQRNT and NSERC. M. M. Wilde acknowledges support from the Centre de Recherches Mathématiques.

References

[1] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, 1979.

[2] L. Xie and P. Kumar, "An achievable rate for the multiple-level relay channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1348–1358, 2005.

[3] A. El Gamal and Y.-H. Kim, "Lecture notes on network information theory," January 2010, arXiv:1001.3404.

[4] M. M. Wilde, *From Classical to Quantum Shannon Theory*, 2011, arXiv:1106.1445.

[5] A. S. Holevo, "The capacity of the quantum channel with general signal states," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 269–273, 1998.

[6] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A*, vol. 56, pp. 131–138, 1997.

[7] A. Winter, "The capacity of the quantum multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 3059–3065, 2001.

[8] O. Fawzi, P. Hayden, I. Savov, P. Sen, and M. M. Wilde, "Classical communication over a quantum interference channel," February 2011, arXiv:1102.2624.

[9] P. Sen, "Achieving the Han-Kobayashi inner bound for the quantum interference channel by sequential decoding," arXiv:1109.0802.

[10] I. Savov and M. M. Wilde, "Classical codes for quantum broadcast channels," 2011, arXiv:1111.3645.

[11] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, "Classical capacity of the lossy bosonic channel: The exact solution," *Phys. Rev. Lett.*, vol. 92, no. 2, p. 027902, January 2004.

[12] A. Carleial, "Multiple-access channels with different generalized feedback signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 841–850, 1982.

[13] M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1753–1768, 2003.

[14] A. Winter, "Coding theorem and strong converse for quantum channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2481–2485, 1999.

[15] D. Collins, N. Gisin, and H. De Riedmatten, "Quantum relays for long distance quantum cryptography," *Journal of Modern Optics*, vol. 52, no. 5, pp. 735–753, 2005.

APPENDIX

A. Decoding at the Relay

In this section we give the details of the POVM construction and the error analysis for the Relay decoder.

POVM Construction. During block j, the Relay wants to decode the message ℓ_j encoded in $u^n(\ell_j, \ell_{j-1})$, given the knowledge of the message ℓ_{j-1} from the previous block. Consider the state obtained by tracing over the systems X and B in (3):

$$\theta^{UX_1B_1} = \sum_{u,x_1} p\left(u|x_1\right) p\left(x_1\right) |u\rangle \langle u|^U \otimes |x_1\rangle \langle x_1|^{X_1} \otimes \sigma_{u,x_1}^{B_1},$$

where $\sigma_{u,x_1}^{B_1} \equiv \sum_x p\left(x|x_1,u\right) \mathrm{Tr}_B \left\{\rho_{x,x_1}^{B_1 B}\right\}$. Further tracing over the system U leads to the state

$$\theta^{X_1B_1} = \sum_{x_1} p\left(x_1\right) |x_1\rangle \langle x_1|^{X_1} \otimes \bar{\sigma}_{x_1}^{B_1},$$

where $\bar{\sigma}_{x_1} \equiv \sum_u p\left(u|x_1\right) \sigma_{u,x_1}^{B_1}$. Corresponding to the above conditional states are conditionally typical projectors of the following form

$$\Pi_{\sigma_{\ell_j|\ell_{j-1}}} \equiv \Pi^{B^n_{1(j)}}_{\sigma_{u^n(\ell_j,\ell_{j-1}),x_1^n(\ell_{j-1})}}, \quad \Pi_{\bar{\sigma}_{|\ell_{j-1}}} \equiv \Pi^{B^n_{1(j)}}_{\bar{\sigma}_{x_1^n(\ell_{j-1})}}.$$

The Relay constructs a square-root measurement $\{\Gamma_{\ell_j}\}$ using formula (5) and the following positive operators:

$$P^{B^n_{1(j)}}_{\ell_j|\ell_{j-1}} \equiv \Pi_{\bar{\sigma}_{|\ell_{j-1}}} \Pi_{\sigma_{\ell_j|\ell_{j-1}}} \Pi_{\bar{\sigma}_{|\ell_{j-1}}} \tag{21}$$

Error analysis. In this section we show that during block j the Relay will be able to decode ℓ_j from the state $\rho^{B^n_{1(j)}}_{x^n(m_j,l_j,l_{j-1}),x_1^n(l_{j-1})}$, provided the rate $R_\ell < I(U;B_1|X_1) = H(B_1|X_1) - H(B_1|UX_1) - \delta$. The bound follows from the following properties of typical projectors:

$$\mathrm{Tr}[\Pi_{\sigma_{\ell_j|\ell_{j-1}}}] \leq 2^{n[H(B_1|UX_1)+\delta]} \tag{22}$$

$$\Pi_{\bar{\sigma}_{|\ell_{j-1}}} \bar{\sigma}\ \Pi_{\bar{\sigma}_{|\ell_{j-1}}} \leq 2^{-n[H(B_1|X_1)-\delta]} \Pi_{\bar{\sigma}_{|\ell_{j-1}}}, . \tag{23}$$

The average probability of error at the Relay is given by:

$$\bar{p}_e^R \equiv \frac{1}{|\mathcal{L}|} \sum_{\ell_j} \mathrm{Tr}\Big\{ \Big(I - \Gamma^{B^n_{1(j)}}_{\ell_j|\ell_{j-1}} \Big) \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \Big\},$$

We consider the probability of error for a single message ℓ_j and begin by applying the Hayashi-Nagaoka operator inequality (Lemma 2) to split the error into two terms:

$$\begin{aligned}
\bar{p}_e^R &\equiv \mathrm{Tr}\Big[\Big(I - \Gamma^{B^n_{1(j)}}_{\ell_j|\ell_{j-1}} \Big) \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \Big] \\
&\leq \underbrace{2\,\mathrm{Tr}\left[\Big(I - P^{B^n_{1(j)}}_{\ell_j|\ell_{j-1}} \Big)\ \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \right]}_{(I)} \\
&\quad + \underbrace{4 \sum_{\ell_j' \neq \ell_j} \mathrm{Tr}\left[P^{B^n_{1(j)}}_{\ell_j'|\ell_{j-1}}\ \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \right]}_{(II)}.
\end{aligned}$$

We will bound the expectation of the average probability of error by bounding the individual terms. We bound the first term as follows:

$$\begin{aligned}
(\mathrm{I}) &= \mathrm{Tr}\left[\Big(I - P^{B^n_{1(j)}}_{\ell_j|\ell_{j-1}} \Big)\ \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \right] \\
&= \mathrm{Tr}\Big[\Big(I - \Pi_{\bar{\sigma}_{|\ell_{j-1}}} \Pi_{\sigma_{\ell_j|\ell_{j-1}}} \Pi_{\bar{\sigma}_{|\ell_{j-1}}} \Big) \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \Big] \\
&= 1 - \mathrm{Tr}\Big[\Pi_{\bar{\sigma}_{|\ell_{j-1}}} \Pi_{\sigma_{\ell_j|\ell_{j-1}}} \Pi_{\bar{\sigma}_{|\ell_{j-1}}}\ \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \Big] \\
&\leq 1 - \mathrm{Tr}\Big[\Pi_{\sigma_{\ell_j|\ell_{j-1}}}\ \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \Big] \\
&\qquad + \left\| \Pi_{\bar{\sigma}_{|\ell_{j-1}}} \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \Pi_{\bar{\sigma}_{|\ell_{j-1}}} - \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \right\|_1,
\end{aligned}$$

where the inequality follows from Lemma 1.

By taking the expectation over the code randomness we obtain the bound

$$\begin{aligned}
\mathop{\mathbb{E}}_{U^nX^nX_1^n}(\mathrm{I}) &= 1 - \mathop{\mathbb{E}}_{U^nX_1^n} \mathrm{Tr}\left[\Pi_{\sigma_{\ell_j|\ell_{j-1}}} \mathop{\mathbb{E}}_{X^n|U^nX_1^n} \Big\{ \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \Big\} \right] \\
&\quad + \mathop{\mathbb{E}}_{U^nX^nX_1^n} \left\| \Pi_{\bar{\sigma}_{|\ell_{j-1}}} \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \Pi_{\bar{\sigma}_{|\ell_{j-1}}} - \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \right\|_1 \\
&= 1 - \mathop{\mathbb{E}}_{U^nX_1^n} \mathrm{Tr}\Big[\Pi_{\sigma_{\ell_j|\ell_{j-1}}} \sigma_{\ell_j,\ell_{j-1}} \Big] \\
&\quad + \mathop{\mathbb{E}}_{U^nX^nX_1^n} \left\| \Pi_{\bar{\sigma}_{|\ell_{j-1}}} \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \Pi_{\bar{\sigma}_{|\ell_{j-1}}} - \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \right\|_1 \\
&\leq 1 - \mathop{\mathbb{E}}_{U^nX_1^n} \mathrm{Tr}\Big[\Pi_{\sigma_{\ell_j|\ell_{j-1}}} \sigma_{\ell_j,\ell_{j-1}} \Big] + 2\sqrt{\epsilon} \\
&\leq 1 - (1-\epsilon) + 2\sqrt{\epsilon} = \epsilon + 2\sqrt{\epsilon}.
\end{aligned}$$

The first inequality follows from Lemma 4 and the property

$$\mathrm{Tr}\Big[\Pi_{\bar{\sigma}_{|\ell_{j-1}}}\ \bar{\sigma} \Big] \geq 1 - \epsilon. \tag{24}$$

The second inequality follows from:

$$\mathrm{Tr}\Big[\Pi_{\sigma_{\ell_j|\ell_{j-1}}} \sigma_{\ell_j,\ell_{j-1}} \Big] \geq 1 - \epsilon. \tag{25}$$

To bound the second term we proceed as follows:

$$\begin{aligned}
\mathop{\mathbb{E}}_{U^nX^nX_1^n} \{(II)\} &= \mathop{\mathbb{E}}_{U^nX^nX_1^n} \sum_{\ell_j' \neq \ell_j} \mathrm{Tr}\left[P^{B^n_{1(j)}}_{\ell_j'|\ell_{j-1}}\ \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \right] \\
&= \mathop{\mathbb{E}}_{X_1^n} \sum_{\ell_j' \neq \ell_j} \mathrm{Tr}\left[\mathop{\mathbb{E}}_{U^nX^n|X_1^n} \Big\{ P^{B^n_{1(j)}}_{\ell_j'|\ell_{j-1}} \Big\} \mathop{\mathbb{E}}_{U^nX^n|X_1^n} \{ \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \} \right] \\
&= \mathop{\mathbb{E}}_{X_1^n} \sum_{\ell_j' \neq \ell_j} \mathrm{Tr}\left[\mathop{\mathbb{E}}_{U^nX^n|X_1^n} \Big\{ P^{B^n_{1(j)}}_{\ell_j'|\ell_{j-1}} \Big\}\ \bar{\sigma}_{|\ell_{j-1}} \right]
\end{aligned}$$

The expectation can be broken up because $\ell_j' \neq \ell_j$ and thus the U^n codewords are independent. We have also used

$$\mathop{\mathbb{E}}_{U^nX^n|X_1^n} \{ \rho^{B^n_{1(j)}}_{m_j,\ell_j,\ell_{j-1}} \} = \bar{\sigma}_{|\ell_{j-1}}. \tag{26}$$

We continue by expanding the operator $P^{B_{1(j)}^n}_{\ell_j'|\ell_{j-1}}$ as follows:

$$
\begin{aligned}
&= \mathop{\mathbb{E}}_{U^n X^n X_1^n} \sum_{\ell_j' \neq \ell_j} \mathrm{Tr}\left[\Pi_{\bar{\sigma}_{|\ell_{j-1}}} \Pi_{\sigma_{\ell_j'|\ell_{j-1}}} \Pi_{\bar{\sigma}_{|\ell_{j-1}}} \bar{\sigma}_{|\ell_{j-1}}\right] \\
&= \mathop{\mathbb{E}}_{U^n X^n X_1^n} \sum_{\ell_j' \neq \ell_j} \mathrm{Tr}\left[\Pi_{\sigma_{\ell_j'|\ell_{j-1}}} \underbrace{\Pi_{\bar{\sigma}_{|\ell_{j-1}}} \bar{\sigma}_{|\ell_{j-1}} \Pi_{\bar{\sigma}_{|\ell_{j-1}}}}_{⑤}\right] \\
&\leq \mathop{\mathbb{E}}_{U^n X^n X_1^n} \sum_{\ell_j' \neq \ell_j} \mathrm{Tr}\left[\Pi_{\sigma_{\ell_j'|\ell_{j-1}}} \; 2^{-n[H(B_1|X_1)-\delta]} \Pi_{\bar{\sigma}_{|\ell_{j-1}}}\right] \\
&\leq 2^{-n[H(B_1|X_1)-\delta]} \mathop{\mathbb{E}}_{U^n X^n X_1^n} \sum_{\ell_j' \neq \ell_j} \mathrm{Tr}\left[\Pi_{\sigma_{\ell_j'|\ell_{j-1}}}\right] \\
&\leq 2^{-n[H(B_1|X_1)-\delta]} \mathop{\mathbb{E}}_{U^n X^n X_1^n} \sum_{\ell_j' \neq \ell_j} 2^{n[H(B_1|UX_1)+\delta]} \\
&\leq |\mathcal{L}| \; 2^{-n[I(U;B_1|X_1)-2\delta]}.
\end{aligned}
$$

The first inequality follows from using (23) on the expression ⑤. The second inequality follows from the fact that $\Pi_{\bar{\sigma}_{|\ell_{j-1}}}$ is a positive operator less than the identity. More precisely we have

$$
\begin{aligned}
\Pi_{\sigma_{\ell_j'|\ell_{j-1}}} \Pi_{\bar{\sigma}_{|\ell_{j-1}}} &= \Pi_{\sigma_{\ell_j'|\ell_{j-1}}} \Pi_{\bar{\sigma}_{|\ell_{j-1}}} \Pi_{\sigma_{\ell_j'|\ell_{j-1}}} \\
&\leq \Pi_{\sigma_{\ell_j'|\ell_{j-1}}} \; I \; \Pi_{\sigma_{\ell_j'|\ell_{j-1}}} \\
&= \Pi_{\sigma_{\ell_j'|\ell_{j-1}}}.
\end{aligned}
$$

The penultimate inequality follows from (22).

Thus if we choose $R_\ell \leq I(U;B_1|X_1) - 3\delta$, we can make the expectation of the average probability of error vanish in the limit of many uses of the channel.

Proof conclusion. Note that the gentle operator lemma for ensembles is used several times in the proof to guarantee that the effect of acting with one of the projectors from the "measurement sandwich" does not disturb the state too much. Furthermore, because each of the outputs blocks is operated on twice, we again depend on the gentle operator lemma to guarantee that the state disturbance is asymptotically negligible.