One-shot Marton inner bound for classical-quantum broadcast channel

Jaikumar Radhakrishnan* Pranab Sen* Naqueeb Warsi*

Abstract

We consider the problem of communication over a classical-quantum broadcast channel with one sender and two receivers. Generalizing the classical inner bounds shown by Marton and the recent quantum asymptotic version shown by Savov and Wilde, we obtain one-shot inner bounds in the quantum setting. Our bounds are stated in terms of smooth min and max Rényi divergences. We obtain these results using a different analysis of the random codebook argument and employ a new one-shot classical mutual covering argument based on rejection sampling. These results give a full justification of the claims of Savov and Wilde in the classical-quantum asymptotic iid setting; the techniques also yield similar bounds in the information spectrum setting.

1 Introduction

We consider the problem of communication over a broadcast channel with one sender (Alice) and two receivers (Bob and Charlie). They have access to a channel that takes one input X (supplied by Alice) and produces two outputs Y and Z, received by Bob and Charlie respectively. The characteristics of the channel are given by $p(y, z \mid x)$. The goal is to obtain bounds on the rates at which Alice may transmit messages to Bob and Charlie.

Marton bound: The most general achievable rate region known in this situation was given by Marton [1], who showed the following.

Theorem 1. *Fix a discrete memoryless broadcast channel given by $p(y, z \mid x)$. Let a pair of random variables is (U, V) taking values in $\mathcal{U} \times \mathcal{V}$ and a function $f : \mathcal{U} \times \mathcal{V} \to \mathcal{X}$ be given; suppose the random variables (U, V, Y, Z) have joint probability mass function $p(u, v, y, z) = p(u, v)p(y, z \mid f(u, v))$. Let (R_1, R_2) be such that*

$$\begin{aligned} R_1 &< I[U; Y], \\ R_2 &< I[V; Z], \\ R_1 + R_2 &< I[U; Y] + I[V; Z] - I[U; V]. \end{aligned}$$

Then, the rate pair (R_1, R_2) is achievable.

A quantum version of the broadcast channel was considered by Savov and Wilde [2], where instead of $p(y, z \mid x)$, the channel is characterized by density matrices ρ_x^{BC} (note that the channel takes classical input, so x is classical). They formulated the following rate region in the classical-quantum setting.

Theorem 2. *Let $(\mathcal{X}, \mathcal{N} : x \mapsto \rho_x^{BC})$ be a classical-quantum broadcast channel. Let a pair of random variables (U, V) taking values in $\mathcal{U} \times \mathcal{V}$ and a function $f : \mathcal{U} \times \mathcal{V} \to \mathcal{X}$ be given; consider the state*

$$\rho^{UVBC} = \sum_{(u,v) \in \mathcal{U} \times \mathcal{V}} P_{UV}(u, v) |u\rangle\langle u|^U \otimes |v\rangle\langle v|^V \otimes \rho_{f(u,v)}^{BC}.$$

[*] School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai 400005, India. Email: {jaikumar,naqueeb}@tifr.res.in, pgdsen@tcs.tifr.res.in

Let (R_1, R_2) be such that

$$
\begin{aligned}
R_1 &< I[U;B],\\
R_2 &< I[V;C],\\
R_1 + R_2 &< I[U;B] + I[V;C] - I[U;V],
\end{aligned}
$$

where the information theoretic quantities above are computed with respect to the state ρ^{UVBC}. Then, the rate pair (R_1, R_2) is achievable. (See [3] for definitions.)

Prior to Savov and Wilde [2] Yard et al. proved superposition coding inner bounds for classical and quantum communication over a quantum broadcast channel in [4]. In [5] Dupuis et al. prove the Marton's inner bound for quantum transmission over a quantum broadcast channel.

Motivation: Our work is motivated by the work of Savov and Wilde [2] mentioned above, who base their proof on the presentation of Marton's bound in the book of El Gamal and Kim [6]. The argument proceeds as follows. The sender uses a randomly generated codebook. The codewords to be fed into the channel are arranged in a rectangular array. The rows are partitioned into 2^{nR_1} bands and the columns into 2^{nR_2} bands. There is one band of rows for each message m_1 that Alice might need to send to Bob, and one band of columns for each message m_2 that Alice might need to send to Charlie. On receiving (m_1, m_2), Alice picks a codeword from the intersection of the corresponding bands and feeds it into the channel. Bob and Charlie, on receiving their share of the channel output, try to determine the intended messages m_1 and m_2, that is, locate the corresponding row and column bands. El Gamal and Kim show that with high probability the correct bands can be identified by Bob and Charlie. Formally, this is done by applying the union bound to upper bound the probability of decoding a wrong band. This part of the proof is not straightforward to translate into the quantum setting. In fact, the argument presented by Savov and Wilde [2] leaves a gap. In a recent archive version of their work, Savov and Wilde [3] address this gap, and complete the justification of their claims in the asymptotic iid setting by employing an *over counting idea* from the previous version of this paper [7].

Our results: We consider the above problem of communication over a classical-quantum channel in the one-shot setting. We show the following version of Marton's bound.

Theorem 3. *Let $(\mathcal{X}, \mathcal{N} : x \mapsto \rho_x^{BC})$ be a classical-quantum broadcast channel. Let a pair of random variables is (U, V) taking values in $\mathcal{U} \times \mathcal{V}$ and a function $f : \mathcal{U} \times \mathcal{V} \to \mathcal{X}$ be given; consider the state*

$$
\rho^{UVBC} = \sum_{(u,v) \in \mathcal{U} \times \mathcal{V}} P_{UV}(u,v) |u\rangle\langle u|^U \otimes |v\rangle\langle v|^V \otimes \rho^{BC}_{f(u,v)}. \tag{1}
$$

Let (R_1, R_2), ε, ϵ_∞, ϵ_0 and $\tilde{\varepsilon}$ be such that

$$
R_1 \leq I_0^{\varepsilon_0}[U;B] - 3\log\frac{1}{\tilde{\varepsilon}} - 2, \tag{2}
$$

$$
R_2 \leq I_0^{\varepsilon_0}[V;C] - 3\log\frac{1}{\tilde{\varepsilon}} - 2, \tag{3}
$$

$$
R_1 + R_2 \leq I_0^{\varepsilon_0}[U;B] + I_0^{\varepsilon_0}[V;C] - I_\infty^{\varepsilon_\infty}[U;V] - 5\log\frac{1}{\tilde{\varepsilon}} - 5 \tag{4}
$$

where $\varepsilon_\infty \leq \frac{1}{4}$ and is such that $I_\infty^{\varepsilon_\infty}(U;V) \geq 0$ and $40\tilde{\varepsilon} + 16\varepsilon_0 \leq \varepsilon$. Then, there exists a (R_1, R_2, ε)-classical-quantum broadcast channel code. The information theoretic quantities mentioned in (2), (3) *and* (4) *are calculated with respect to the classical-quantum state given in (1). (Classical-quantum broadcast channel code and the quantities I_0^ε and I_∞^ε are defined in Section 2.)*

This result implies the results of Savov and Wilde [3]. Our method also yields the following one-shot version of Marton's inner bound in the classical setting.

Theorem 4. *Consider a classical broadcast channel given by $p(yz \mid x)$, where $x \in \mathcal{X}$, $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$. Suppose there is a pair of random variables $(U, V) \in \mathcal{U} \times \mathcal{V}$ and a function $f : \mathcal{U} \times \mathcal{V} \to \mathcal{X}$. Let (Y, Z) be random variables such that* $\Pr\{(Y, Z) = (y, z) \mid U = u, V = v\} = p(y, z \mid f(u, v))$*. Further, suppose (R_1, R_2) and ϵ, ϵ_0, ϵ_∞, and $\tilde{\epsilon}$ are such that*

$$R_1 \quad \leq \quad I_0^{\epsilon_0}[U;Y] - 3\log\frac{1}{\tilde{\epsilon}} - 2, \tag{5}$$

$$R_2 \quad \leq \quad I_0^{\epsilon_0}[V;Z] - 3\log\frac{1}{\tilde{\epsilon}} - 2, \tag{6}$$

$$R_1 + R_2 \quad \leq \quad I_0^{\epsilon_0}[U;Y] + I_0^{\epsilon_0}[V;Z] - I_\infty^{\epsilon_\infty}[U;V] - 5\log\frac{1}{\tilde{\epsilon}} - 5 \tag{7}$$

where $\varepsilon_\infty \leq \frac{1}{4}$ and is such that $I_\infty^{\varepsilon_\infty}(U;V) \geq 0$ and $37\tilde{\epsilon} + 8\epsilon_0 \leq \epsilon$. Then, there is a one-shot (R_1, R_2, ϵ)-classical broadcast code for the channel. (Classical broadcast channel code and the quantities I_0^ε and I_∞^ε are defined in Section 3.)

Techniques: Our proof follows along the lines of the proof in El Gamal and Kim [6] for the classical Marton bound. As before, we generate a rectangular array, whose rows and column indices are chosen independently according to the marginal distributions of U and V; furthermore, as in the original proof, we partition the rows and columns into bands of appropriate sizes. There are two major difficulties that one encounters.

(a) First, given a message pair (m_1, m_2), we do not have a natural analogue of joint typicality to help us choose a (u, v) pair from the subcodebook. Furthermore, in the iid settings it is well-known that if a jointly typical pair is used as input to the channel, the output is very likely to be jointly typical with the input; however, we cannot exploit such facts in the one-shot setting. Instead, we use rejection sampling to ensure that the resulting probability distribution is very close to the ideal joint distribution on (U, V) and the outputs for the channel.

(b) Second, the difficulty mentioned above with applying the union bound, particular to the asymptotic quantum setting, are present in the one-shot setting as well, and equally hard to overcome. The solution is somewhat technical, we observe that if the sizes of the bands are tuned carefully, we have the liberty to overestimate the probability of error, and obtain a good bound. *Our analysis technique in fact shows that if the receivers employ a standard pretty good measurement technique, the they not only decode the transmitted bands correctly, but also recover the row and column index pair that was used for transmission.* This analysis of error at the two receivers differs from the standard analysis in the normal point to point classical-quantum channel coding problem, as well as from the 'decoding up to the band' arguments of El Gamal and Kim.

The new methods seem necessary to deal with the additional complications that arise in the one-shot setting.

Related work: A one-shot version of Marton's bound was also proved by Verdú (see Theorem [8, Theorem 8]). Stated using the terminology above, the Verdu's inequalities can be stated as follows.

$$R_1 \leq I_0^{\epsilon_0}[U;Y] - \ln\frac{1}{\gamma}$$

$$R_2 \leq I_0^{\epsilon_0}[V;Z] - I_\infty^{\epsilon_\infty}[U;V] - 2\ln\frac{1}{\gamma}$$

$$\varepsilon \leq 2\epsilon_0 + \epsilon_\infty + 2\gamma + \exp(-\frac{1}{\gamma}).$$

Note that apart from the dependence on ϵ, our bounds generalize Verdú's. Furthermore, while the strategy employed by Verdú allows the decoding of the transmitted bands correctly with high probability, we achieve more by decoding the actual row and column.

The technical difficulty in ensuring unique decoding in several classical settings related to ours has been recongnised and addressed in several recent works [9, 10, 11, 12]. In particular, Minero, Lim and Kim [11, Lemma 1] achieve unique decoding for the rate regions associated with the Gelfand-Pinsker bound in the asymptotic setting by controlling the perturbations in distributions caused by conditioning on other events. Their analysis makes critical use of the asymptotic equipartition property (AEP) available in the asymptotic iid setting. We, working in the one-shot and non-iid setting do not have recourse to such tools. Instead, we observe that by carefully controlling the bad sizes while generating the code, one can simple overcount and bound the probability of error. We belive this method is applicable to other settings as well (see Remark 1 below). Furthermore, this mehtod of analysis works in the quantum setting with almost no change.

Asymptotic iid and non-iid bounds. Our bounds imply the bounds obtained earlier for the same problem in the iid settings. The asymptotic information spectrum setting pioneered by Han and Vérdu [13] and its quantum version due to Hayashi and Nagaoka [14] allows one to derive meaningful bounds on rates even in the absence of the iid assumption; however, the analysis is often more challenging in these settings. The bounds in our work are expressed using smooth Rényi quantities. The close relationship between these quantitites and the quantitites that typically arise in the information spectrum setting (see Datta and Renner [15]) allows us to conclude similar bounds in the asymptotic case, in both the non-iid (information spectrum) and iid settings.

Organisation of the paper

In Section 2 we give the definitions of the information theoretic quantities which appear in Theorem 3. We then prove this theorem in this section. In Section 3 we give the proof for Theorem 4. In Section 4 one-shot version of the mutual covering lemma is proven. In Section 5 we discuss the asymptotic behaviour of the one-shot bounds derived in this paper.

2 The classical-quantum one-shot bound

Definition 1 (Channel). *Let $\mathcal{X}$ be a finite alphabet. We model a classical-quantum broadcast channel between parties Alice, Bob and Charlie as a map*

$$\mathcal{N} : x \mapsto \rho_x^{BC}, \tag{8}$$

where $x \in \mathcal{X}$ is the input given to the channel by Alice, and ρ_x^{BC} is the joint state of Bob and Charlie in the Hilbert space $\mathcal{H}_A \times \mathcal{H}_B$. The resulting state of Bob is then modelled as $\rho_x^B = \mathrm{Tr}_C \rho_x^{BC}$, and the state of Charlie is modelled as $\rho_x^C = \mathrm{Tr}_B \rho_x^{BC}$.

Our goal is to use this channel to enable Alice to transmit a pair of messages $(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$ (for some large sets $\mathcal{M}_1$ and $\mathcal{M}_2$) such that Bob can recover m_1 and Charlie can recover m_2.

Definition 2 (Encoding, Decoding, Error). *An (R_1, R_2, ε)-classical-quantum broadcast channel code consists of*

- *an encoding function $F : [2^{R_1}] \times [2^{R_2}] \to \mathcal{X}$, and*
- *two decoding POVMs $\{\mathcal{T}_{m_1}^B : m_1 \in [2^{R_1}]\}$ and $\{\mathcal{T}_{m_2}^C : m_2 \in [2^{R_2}]\}$ such that the average probability of error*

$$\frac{1}{2^{R_1+R_2}} \sum_{(m_1,m_2)\in[2^{R_1}]\times[2^{R_2}]} p_e(m_1, m_2) \leq \varepsilon, \tag{9}$$

where

$$p_e(m_1, m_2) = \mathrm{Tr}\left[\left(\mathbb{I} - \mathcal{T}_{m_1}^B \otimes \mathcal{T}_{m_2}^C\right) \mathcal{N}(F(m_1, m_2))\right],$$

is the probability of error when Alice uses this scheme to transmit the message pair (m_1, m_2).

Our one-shot version of the Marton inner bound will be stated in terms of min and max Rényi divergences which are defined as follows.

Definition 3. *(Smooth quantum min Rényi divergence [16]) Let* $\rho^{UB} := \sum_{u \in \mathcal{U}} p_U(u) |u\rangle\langle u|^U \otimes \rho_u^B$ *be a classical quantum state. For* $\varepsilon \in [0,1)$ *the smooth min Rényi divergence between the systems* U *and* B *denoted is*

$$I_0^{\varepsilon}[U;B] := \sup_{\substack{0 \preceq \Gamma \preceq \mathbb{I} \\ \mathrm{Tr}\left[\Gamma \rho^{UB}\right] \geq 1-\varepsilon}} - \log \mathrm{Tr}\left[\Gamma\left(\rho^U \otimes \rho^B\right)\right].$$

Definition 4. *(Smooth max Rényi divergence [17]) For random variables* $(U,V) \sim p_{UV}$ *with range* $\mathcal{U} \times \mathcal{V}$ *and* $\epsilon \in [0,1)$ *we have*

$$I_{\infty}^{\varepsilon}[U;V] = \inf_{\substack{\mathcal{G} \subseteq \mathcal{U} \times \mathcal{V} \\ P_{UV}(\mathcal{G}) \geq 1-\varepsilon}} \sup_{(u,v) \in \mathcal{G}} \log \frac{p_{UV}(u,v)}{p_U(u) p_V(v)}.$$

We are now ready to prove Theorem 3.

2.1 Proof of Theorem 3

We need to describe the encoding function $F : [2^{R_1}] \times [2^{R_2}] \rightarrow \mathcal{X}$ and suitable POVMs that will be used for decoding. We will adapt the scheme suggested by Marton as presented in El Gamal and Kim [6], to the quantum one-shot setting.

The random codebook: Let ρ^{UVBC} and f be as in the statement of the theorem, and (R_1, R_2) satisfy the required inequalities. In the following we set

$$\begin{aligned} I_{\infty} &= I_{\infty}^{\varepsilon_{\infty}}[U;V]; \\ I_0^B &= I_0^{\varepsilon_0}[U;B]; \\ I_0^C &= I_0^{\varepsilon_0}[V;C]. \end{aligned}$$

Let Γ^{UB} be such that $\mathrm{Tr}\left[\Gamma^{UB}\rho^{UB}\right] \geq 1 - \varepsilon_0$ and $\mathrm{Tr}\left[\Gamma^{UB}\left(\rho^U \otimes \rho^B\right)\right] = 2^{-I_0^B}$. Similarly, let Γ^{VC} be such that $\mathrm{Tr}\left[\Gamma^{VC}\rho^{VC}\right] \geq 1 - \varepsilon_0$ and $\mathrm{Tr}\left[\Gamma^{VC}\left(\rho^V \otimes \rho^C\right)\right] = 2^{-I_0^C}$ where $\rho^{UB} = \mathrm{Tr}_{VC}\left[\rho^{UVBC}\right]$; $\rho^U = \mathrm{Tr}_{VBC}\left[\rho^{UVBC}\right]$; $\rho^B = \mathrm{Tr}_{UVC}\left[\rho^{UVBC}\right]$; $\rho^{VC} = \mathrm{Tr}_{UB}\left[\rho^{UVBC}\right]$; $\rho^V = \mathrm{Tr}_{UBC}\left[\rho^{UVBC}\right]$ and $\rho^C = \mathrm{Tr}_{UVB}\left[\rho^{UVBC}\right]$. Choose positive integers r_1 and r_2 such that

$$R_1 + r_1 \leq I_0^B - 2\log\frac{1}{\tilde{\varepsilon}} - 1; \tag{10}$$

$$R_2 + r_2 \leq I_0^C - 2\log\frac{1}{\tilde{\varepsilon}} - 1; \tag{11}$$

$$r_1, r_2 \geq \log\frac{1}{\tilde{\varepsilon}}; \tag{12}$$

$$r_1 + r_2 = \left\lceil I_{\infty} + \log\frac{1}{\tilde{\varepsilon}} \right\rceil. \tag{13}$$

[To see that such a choice exists we may, e.g., start with $r_1, r_2 = \lceil \log \frac{1}{\tilde{\varepsilon}} \rceil$. Clearly, the first three constraints are satisfied, and for the last, LHS $\leq$ RHS. Now, the assumptions in the theorem imply that we may increase r_1 and r_2 without violating the first two constraints until the last constraint is satisfied.]

Let $U[1]$, $U[2]$,..., $U[2^{R_1+r_1}]$ be drawn independently according to the distribution of U; similarly, let $V[1]$, $V[2]$,..., $V[2^{R_2+r_2}]$ be drawn according the distribution of V. These samples will be associated with rows and columns of a $2^{R_1+r_1} \times 2^{R_2+r_2}$ matrix $\mathcal{C}$, whose entries will be elements of $\mathcal{X} \cup \{\star\}$. The entry $\mathcal{C}[k,\ell]$ will be determined as follows.

For each pair (k,ℓ), let $\eta(k,\ell)$ be chosen independently and uniformly from $[0,1]$. Let $\mathbb{I}(k,\ell)$ be the 0-1 indicator random variable defined by

$$\mathbf{I}(k,\ell) = \mathbb{I}\left\{\eta(k,\ell) \leq \frac{p(U[k],V[\ell])}{2^{I_{\infty}} p(U[k]) p(V[\ell])}\right\}. \tag{14}$$

Then, $\mathcal{C}[k,\ell] = f(U[k], V[\ell])$ if $\mathbf{I}(k,\ell) = 1$, and $\mathcal{C}[k,\ell] = \star$ otherwise. Thus, $\mathcal{C}$ is a random matrix of entries, determined by the random choices of $(U[k], V[\ell], \eta(k,\ell))$ for $k = 1, 2, \ldots, 2^{R_1+r_1}$ and $\ell = 1, 2, \ldots, 2^{R_2+r_2}$; we will call this (the random matrix, together with all the associated random choices $U[k]$, $V[\ell]$ and $\eta(k,\ell)$) the random codebook $\mathcal{C}$. Later we will fix one realization of $\mathcal{C}$.

Our encoding function $F : [2^{R_1}] \times [2^{R_2}] \rightarrow \mathcal{X}$ will be based on $\mathcal{C}$. We partition the row indices of $\mathcal{C}$ into 2^{R_1} classes each with 2^{r_1} elements; let the i-th class $\mathcal{C}_1(i) = \{(i-1)2^{r_1}+1, (i-1)2^{r_1}+2, \ldots, i2^{r_1}\}$. Similarly, we partition the column indices into 2^{R_2} classes, where the j-th class $\mathcal{C}_2(j) = \{(j-1)2^{r_2}+1, (j-1)2^{r_2}+2, \ldots, j2^{r_2}\}$. $F(m_1, m_2)$ will be set to $\mathcal{C}[k,\ell] \neq \star$ for some $(k,\ell) \in \mathcal{C}_1(m_1) \times \mathcal{C}_2(m_2)$. However, we must ensure that the choice (k,ℓ) aids the decoding process.

Below, we will see that the POVMs used by Bob and Charlie will be based on operators defined as follows.

$$\Lambda_u^B \;:=\; \mathrm{Tr}_U \left[\Gamma^{UB} \left(|u\rangle\langle u| \otimes \mathbb{I}\right)\right] \tag{15}$$

$$\Lambda_v^C \;:=\; \mathrm{Tr}_V \left[\Gamma^{VC} \left(|v\rangle\langle v| \otimes \mathbb{I}\right)\right]. \tag{16}$$

Similar operator was used by Wang and Renner in [16] to design the decoding POVM elements for finding one-shot achievable rate for the point to point classical-quantum channels. Our choice of (k,ℓ) will be guided by these operators. If $\mathcal{C}[i,j] = x \neq \star$, then let

$$\begin{aligned}\alpha(i,j) &= \mathrm{Tr}\left[\Lambda^B_{U[i]} \rho^B_{f(U[i],V[j])}\right];\\ \beta(i,j) &= \mathrm{Tr}\left[\Lambda^C_{V[j]} \rho^C_{f(U[i],V[j])}\right].\end{aligned}$$

If $\mathcal{C}[i,j] = \star$, let $\alpha(i,j), \beta(i,j) = -\infty$. For a pair of messages (m_1, m_2), let $F(m_1, m_2) = \mathcal{C}[i,j]$, where $(i,j) \in \mathcal{C}_1(m_1) \times \mathcal{C}_2(m_2)$ is the lexicographically first pair such that $\alpha(i,j), \beta(i,j) \geq 1 - 4\varepsilon_0$; if no such (i,j) exists, then let $F(m_1, m_2)$ be the first element of $\mathcal{X}$.

Decoding: We first consider Bob's strategy for recovering m_1 on receiving the channel output σ^B. Fix a codebook. For each $k \in [2^{R_1+r_1}]$, we have the operator $\Lambda^B_{U[k]}$ defined above. Bob will *normalize* these operators, to obtain a POVM. The POVM element corresponding to k will be

$$\mathcal{T}_k^B = \left(\sum_{k' \in [2^{R_1+r_1}]} \Lambda^B_{U[k']}\right)^{-\frac{1}{2}} \Lambda^B_{U[k]} \left(\sum_{k' \in [2^{R_1+r_1}]} \Lambda^B_{U[k']}\right)^{-\frac{1}{2}}. \tag{17}$$

Bob measures his state using these operators to obtain an index $\tilde{k} \in [2^{r_1+R_1}]$ (we would like this to be k, the row index used by Alice). He outputs $\tilde{m}_1$ if $\tilde{k} \in \mathcal{C}_1(\tilde{m}_1)$. Similarly, for every $\ell \in [2^{R_2+r_2}]$ Charlie has the following POVM element

$$\mathcal{T}_\ell^C = \left(\sum_{\ell' \in [2^{R_2+r_2}]} \Lambda^C_{V[\ell']}\right)^{-\frac{1}{2}} \Lambda^C_{V[\ell]} \left(\sum_{\ell' \in [2^{R_2+r_2}]} \Lambda^C_{V[\ell']}\right)^{-\frac{1}{2}}. \tag{18}$$

Using this POVM, Charlie measures his state σ^C to obtain a column index $\tilde{\ell} \in [2^{r_2+R_2}]$, and outputs $\tilde{m}_2$ if $\tilde{\ell} \in \mathcal{C}_{\tilde{m}_2}$.

Joint typicality versus rejection sampling: In the standard argument [6], the indicator random variable $\mathbf{I}(k,\ell)$ stands for joint typicality of $U[k]$ and $V[\ell]$. The rejection sampling based on I_∞, has the same effect. We list below some of its properties.

(P1) $\mathbb{E}\{\mathbf{I}(k,\ell)\} \geq (1-\varepsilon_\infty)2^{-I_\infty}$.

(P2) For all u and v,

$$\begin{aligned}\mathbb{E}\{\mathbf{I}(k,\ell)\mid U[k]=u\} &\leq \sum_v p(v)\frac{p(u,v)}{2^{I_\infty}p(u)p(v)} \leq 2^{-I_\infty}; && (19)\\ \mathbb{E}\{\mathbf{I}(k,\ell)\mid V[\ell]=v\} &\leq \sum_u p(u)\frac{p(u,v)}{2^{I_\infty}p(u)p(v)} \leq 2^{-I_\infty}; && (20)\\ \mathbb{E}\{\mathbf{I}(k,\ell)\} &\leq 2^{-I_\infty}. && (21)\end{aligned}$$

(P3) If $\ell \neq \ell'$, then $\mathbf{I}(k,\ell)$ and $\mathbf{I}(k,\ell')$ are conditionally independent given $U[k]$; if $k \neq k'$, then $\mathbf{I}(k,\ell)$ and $\mathbf{I}(k',\ell)$ are conditionally independent given $V[\ell]$.

(P4) If $k \neq k'$ and $\ell \neq \ell'$, then $\mathbf{I}(k,\ell)$ and $\mathbf{I}(k',\ell)$ are independent.

Probability of error: Suppose a pair of messages $(m_1,m_2) \in [2^{R_1}] \times [2^{R_2}]$ is transmitted by Alice using the above scheme and is decoded as $(\tilde{m}_1,\tilde{m}_2)$ by Bob and Charlie. We wish to show that the probability (averaged over the choice of the codebook) that $(\tilde{m}_1,\tilde{m}_2) \neq (m_1,m_2)$ is at most ϵ. By the symmetry in the generation of the code book, it is enough to prove this claim for $(m_1,m_2)=(1,1)$. There are several sources of error: (i) Alice finds no suitable pair $(k,\ell) \in \mathcal{C}_1(1)\times\mathcal{C}_2(1)$; (ii) Alice finds a suitable pair, say (k^*,ℓ^*), but Bob's measurement is unable to determine the index k^* correctly, that is, $\tilde{k} \neq k^*$; (iii) Alice finds a suitable pair, say (k^*,ℓ^*), but but $\tilde{\ell} \neq \ell^*$. We will analyse these events separately. Consider the indicator random variable

$$\mathbf{J}(k,\ell) := \mathbb{I}\left\{\mathbf{I}(k,\ell)=1 \text{ and } \alpha(k,\ell),\beta(k,\ell) \geq 1-4\varepsilon_0\right\}, \tag{22}$$

and consider the three events corresponding to the three sources of error identified above

$$\begin{aligned}\mathcal{E}_1 &:= \text{for all } (k,\ell)\in\mathcal{C}_1(1)\times\mathcal{C}_2(1) : \mathbf{J}(k,\ell)=0;\\ \mathcal{E}_2 &:= \mathcal{E}_1^c \text{ and } \tilde{k}\neq k^*;\\ \mathcal{E}_3 &:= \mathcal{E}_1^c \text{ and } \tilde{\ell}\neq \ell^*.\end{aligned}$$

Consider $\mathcal{E}_1$: We claim

$$\Pr\{\mathcal{E}_1\} \leq 2^{-r_1-r_2+I_\infty+2} + 2^{-r_1+4} + 2^{-r_2+4}. \tag{23}$$

We first show a lower bound on $\mathbb{E}\{\mathbf{J}(k,\ell)\}$. We observed in (P1) above that $\Pr\{\mathbf{I}(k,\ell)=1\} \geq (1-\epsilon_\infty)2^{-I_\infty}$. We account for (and exclude) the probability of the events $\alpha(k,\ell) \leq 1-4\varepsilon_0$ and $\beta(k,\ell)\leq 1-4\varepsilon_0$. Let

$$\mathrm{Bad} = \{(u,v) : \mathrm{Tr}\left[\Lambda_u^B \rho^B_{f(u,v)}\right] \leq 1-4\epsilon_0\}.$$

Then, from the definition of Γ^{UB} and Λ_u^B, we have (using Markov's inequality) that

$$\Pr_{(U,V)}\{(U,V)\in\mathrm{Bad}\} = \sum_{(u,v)\in\mathrm{Bad}} p(u,v) \leq \frac{1}{4}.$$

Thus,

$$\Pr\{\mathbf{I}(k,\ell)=1 \text{ and } \alpha(k,\ell)\geq 1-4\epsilon_0\} = \sum_{(u,v)\in\mathrm{Bad}} p(u)p(v)\frac{p(u,v)}{2^{I_\infty}p(u)p(v)} \leq \left(\frac{1}{4}\right)2^{-I_\infty}.$$

Similarly, $\Pr\{\mathbf{I}(k,\ell)=1 \text{ and } \beta(k,\ell)\geq 1-4\epsilon_0\} \leq \left(\frac{1}{4}\right)2^{-I_\infty}$. Since $\varepsilon_\infty \leq \frac{1}{4}$, we have

$$\mathbb{E}\{\mathbf{J}(k,\ell)\} \geq \left(1-\varepsilon_\infty-\frac{1}{4}-\frac{1}{4}\right)2^{-I_\infty} \geq 2^{-I_\infty-2}.$$

Furthermore,

$$\mathbb{E}\{\mathbf{J}(k,\ell)\mathbf{J}(k',\ell')\} \leq \mathbb{E}\{\mathbf{I}(k,\ell)\mathbf{I}(k'\ell')\};$$

in particular, using properties (P2) and (P3) of $\mathbf{I}(k,\ell)$, we have for $k' \neq k$ and $\ell' \neq \ell$,

$$\mathbb{E}\{\mathbf{J}(k,\ell)\mathbf{J}(k',\ell)\}, \mathbb{E}\{\mathbf{J}(k,\ell)\mathbf{J}(k,\ell')\} \leq 2^{-2I_\infty}.$$

Also, $\mathbf{J}(k,\ell)$ and $\mathbf{J}(k',\ell')$ are independent whenever $k \neq k'$ and $\ell \neq \ell'$. By Lemma 1 (see Section 4, set $\alpha \leftarrow \frac{1}{4}$, $q \leftarrow 2^{I_\infty}$),

$$\Pr\{\mathcal{E}_1\} \leq 2^{-r_1-r_2+I_\infty+2} + \frac{2^{r_1}+2^{r_2}}{2^{r_1+r_2-4}} = 2^{-r_1-r_2+I_\infty+2} + 2^{-r_1+4} + 2^{-r_2+4}. \tag{24}$$

From (24) and our choice of the pair (r_1, r_2) it now follows that

$$\Pr\{\mathcal{E}_1\} \leq 36\tilde{\varepsilon}. \tag{25}$$

Consider $\mathcal{E}_2$ and $\mathcal{E}_3$: We claim that

$$\Pr\{\mathcal{E}_2\} \leq 8\varepsilon_0 + 2^{R_1+2r_1+r_2+2}2^{-I_\infty}2^{-I_0^B} \leq 8\varepsilon_0 + 2\tilde{\varepsilon}; \tag{26}$$

$$\Pr\{\mathcal{E}_3\} \leq 8\varepsilon_0 + 2^{R_2+2r_2+r_1+2}2^{-I_\infty}2^{-I_0^C} \leq 8\varepsilon_0 + 2\tilde{\varepsilon}. \tag{27}$$

To justify (26), we have the following inequalities (below k' takes ranges over $[2^{R_1+r_1}]$ and (k,ℓ) ranges over $\mathcal{C}_1(1) \times \mathcal{C}_2(1)$).

$$\begin{aligned}
\Pr\{\mathcal{E}_1^c \text{ and } \tilde{k} \neq k^*\} &= \mathbb{E}_{\mathcal{C}}\left[\mathbf{I}\{\mathcal{E}_1^c\}\mathrm{Tr}\left[\left(\mathbb{I} - \mathcal{T}_{k^*}^B\right)\rho^B_{f(U[k^*],V[\ell^*])}\right]\right] \\
&\overset{a}{\leq} 2\mathbb{E}_{\mathcal{C}}\left[\mathbf{I}\{\mathcal{E}_1^c\}\mathrm{Tr}\left[\left(\mathbb{I} - \Lambda^B_{U[k^*]}\right)\rho^B_{f(U[k^*],V[\ell^*])}\right]\right] \\
&\quad + \mathbb{E}_{\mathcal{C}}\left[4\sum_{k,\ell}\mathbf{I}\{k^* = k, \ell^* = \ell\}\sum_{k' \neq k^*}\mathrm{Tr}\left[\Lambda^B_{U[k']}\rho^B_{f(U[k^*],V[\ell^*])}\right]\right] \\
&\overset{b}{\leq} 8\varepsilon_0 + 4\sum_{k,\ell,k' \neq k}\mathbb{E}_{\mathcal{C}}\left[\mathbf{I}(k,l)\mathrm{Tr}\left[\Lambda^B_{U[k']}\rho^B_{f(U[k],V[\ell])}\right]\right] \\
&\overset{c}{=} 8\varepsilon_0 + 4\sum_{k,\ell,k' \neq k}\sum_{u,v,u'}2^{-I_\infty}P_U(u)P_V(v)\frac{P_{UV}(u,v)}{P_U(u)P_V(v)}P_U(u')\mathrm{Tr}\left[\Lambda^B_{u'}\rho^B_{f(u,v)}\right] \\
&\leq 8\varepsilon_0 + 2^{r_1+r_2+R_1+r_1+2}2^{-I_\infty}\mathrm{Tr}\left[\Gamma^{UB}\left(\rho^U \otimes \rho^B\right)\right] \\
&\overset{d}{\leq} 8\varepsilon_0 + 2^{R_1+2r_1+r_2+2}2^{-I_\infty}2^{-I_0^B}, \\
&\overset{e}{\leq} 8\varepsilon_0 + 2\tilde{\varepsilon},
\end{aligned}$$

where a follows from the Hayashi-Nagaoka operator inequality [14]; b follows from the definition of the event $\mathcal{E}_1$ and because our encoding ensures that $\alpha(k^*,\ell^*) \geq 1 - 4\epsilon_0$; c follows from the definition of $\mathbf{I}(k,l)$; d follows from the definition of ρ^U, ρ^B and Γ^B; e follows because R_1, r_1 and r_2 satisfy (10)–(13). Similarly, we justify (27). Thus, from (25), (26) and (27) it follows that

$$\Pr\left\{\tilde{M}_1 \neq 1 \cup \tilde{M}_2 \neq 1\right\} \leq 40\tilde{\varepsilon} + 16\varepsilon_0.$$

The above upper bound on the probability of error applies to every (m_1, m_2) as we average over the choices of the codebook; by linearity of expectation, this upper bound holds in expectation when (m_1, m_2) is chosen randomly. It follows that there is a fixed codebook for which the expected error is bounded by $40\tilde{\epsilon} + 16\varepsilon_0$. This completes the proof.

3 The classical one-shot bound

The proofs in this section are just translations of the proof for the quantum case presented above; we reproduce the common parts for the sake of completeness.

We use the following information-theoretic quantities in our theorem.

Definition 5. *A classical broadcast channel consists of an input alphabet $\mathcal{X}$, two output alphabets $\mathcal{Y}$ and $\mathcal{Z}$ and the probability transition function $p_{YZ|X}$.*

Definition 6. *An (R_1, R_2, ε)-code for a classical broadcast channel $C = \{p_{YZ|X}\}$ consists of*

- *an encoding function $F : [2^{R_1}] \times [2^{R_2}] \rightarrow \mathcal{X}$, and*
- *two decoding functions $D_1 : \mathcal{Y} \rightarrow [2^{R_1}]$ and $D_2 : \mathcal{Z} \rightarrow [2^{R_2}]$*

such that

$$\Pr\{(M_1, M_2) \neq (D_1(Y), D_2(Z))\} \leq \epsilon,$$

where (M_1, M_2) are uniformly distributed over $[2^{R_1}] \times [2^{R_2}]$, and Y and Z satisfy $\Pr\{(Y = y, Z = z) \mid M_1 = m_1, M_2 = m_2\} = p_{YZ|X}(yz \mid F(m_1, m_2))$.

3.1 One-shot Marton inner bound for the classical broadcast channel

Our one-shot version of the Marton inner bound will be stated in terms of min and max Rényi divergences which are defined as follows.

Definition 7. *(Smooth classical Min Rényi divergence [18])*
For random variables $(U, V) \sim p_{UV}$ with range $\mathcal{U} \times \mathcal{V}$ and $\epsilon \in [0, 1)$ we have the following.

$$I_0^{\varepsilon}[U; V] := \sup_{\substack{\mathcal{A} \subseteq \mathcal{U} \times \mathcal{V} \\ p_{UV}(\mathcal{A}) \geq 1 - \varepsilon}} - \log \sum_{(u,v) \in \mathcal{A}} p_U(u) p_V(v).$$

3.2 Code generation

We need to describe a function $F : [2^{R_1}] \times [2^{R_2}] \rightarrow \mathcal{X}$. We will adapt the scheme suggested by Marton as presented in El Gamal and Kim [6], to the one-shot setting.

The random codebook: Let (U, V, Y, Z) and f be as in the statement of the theorem, and (R_1, R_2) satisfy the required inequalities. In the following we set

$$\begin{aligned} I_\infty &= I_\infty^{\epsilon_\infty}[U; V]; \\ I_0^B &= I_0^{\epsilon_0}[U; Y]; \\ I_0^C &= I_0^{\epsilon_0}[V; Z]. \end{aligned}$$

Let $\mathcal{A}_1$ be the set in the definition of $I_0^B = I_0^{\epsilon_0}(U; Y)$ such that $p_{UY}(\mathcal{A}_1) \geq 1 - \epsilon_0$ and $\sum_{(u,y) \in \mathcal{A}_1} p_U(u) p_Y(y) \leq 2^{-I_0^B}$. Similarly, let $\mathcal{A}_2$ be the set such that $p_{VZ}(\mathcal{A}_2) \geq 1 - \epsilon_0$ and $\sum_{(v,z) \in \mathcal{A}_1} p_V(v) p_Z(z) \leq 2^{-I_0^C}$. Choose positive integers

r_1 and r_2 such that

$$\begin{aligned} R_1 + r_1 &\leq I_0^B - (2\log\frac{1}{\tilde{\epsilon}} + 1); & (28)\\ R_2 + r_2 &\leq I_0^C - (2\log\frac{1}{\tilde{\epsilon}} + 1); & (29)\\ r_1, r_2 &\geq \log\frac{1}{\tilde{\epsilon}}; & (30)\\ r_1 + r_2 &= \left\lceil I_\infty + \log\frac{1}{\tilde{\epsilon}}\right\rceil. & (31)\end{aligned}$$

[The existence of such r_1 and r_2 was justified in the quantum case.]

Remark 1. *The main difference from the usual calculation is in the (31). One usually imposes a lower bound on $r_1 + r_2$ in order to ensure the* covering property *that with high probability there is a codeword available in the intersection of the two message bands (see $\mathcal{E}_1$ below). The value for $r_1 + r_2$ set in (31) suffices to ensure that such a code word is avaialable with high probability. However, by insisting that $r_1 + r_2$ not exceed the value by too much (we require equality in (31)), we ensure that there are not too many such codewords to choose from, which intuitively makes the decoding unambiguous.*

Let $U[1]$, $U[2]$,$\ldots$, $U[2^{R_1+r_1}]$ be drawn independently according to the distribution of U; similarly, let $V[1]$, $V[2]$,$\ldots$, $V[2^{R_2+r_2}]$ be drawn according the distribution of V. These samples will be associated with rows and columns of a $2^{R_1+r_1} \times 2^{R_2+r_2}$ matrix $\mathcal{C}$, whose entries will be elements of $\mathcal{X} \cup \{\star\}$. The entry $\mathcal{C}[k,\ell]$ will be determined as follows.

For each pair (k,ℓ), let $\eta(k,\ell)$ be chosen independently and uniformly from $[0,1]$. Let $\mathbf{I}(k,\ell)$ be the 0-1 indicator random variable defined by

$$\mathbf{I}(k,\ell) = \mathbb{I}\left\{\eta(k,\ell) \leq \frac{p(U[k],V[\ell])}{2^{I_\infty} p(U[k])p(V[\ell])}\right\}.$$

Then, $\mathcal{C}[k,\ell] = f(U[k],V[\ell])$ if $\mathbf{I}(k,\ell) = 1$, and $\mathcal{C}[k,\ell] = \star$ otherwise. Thus, $\mathcal{C}$ is a random matrix of entries, determined by the random choices of $(U[k],V[\ell],\eta(k,\ell))$ for $k = 1,2,\ldots,2^{R_1+r_1}$ and $\ell = 1,2,\ldots,2^{R_2+r_2}$; we will call this (the random matrix, together with all the associated random choices $U[k]$, $V[\ell]$ and $\eta(k,\ell)$) the random codebook $\mathcal{C}$. Later we will fix one realization of $\mathcal{C}$.

Our encoding function $F : [2^{R_1}] \times [2^{R_2}] \to \mathcal{X}$ will be based on $\mathcal{C}$. We partition the row indices of $\mathcal{C}$ into 2^{R_1} classes each with 2^{r_1} elements; let the i-th class $\mathcal{C}_1(i) = \{(i-1)2^{r_1}+1, (i-1)2^{r_1}+2, \ldots, i2^{r_1}\}$. Similarly, we partition the column indices into 2^{R_2} classes, where the j-th class $\mathcal{C}_2(j) = \{(j-1)2^{r_2}+1, (j-1)2^{r_2}+2, \ldots, j2^{r_2}\}$. If $\mathcal{C}[i,j] = x \neq \star$, then let

$$\begin{aligned} \alpha(i,j) &= \sum_{y:(U[i],y)\in\mathcal{A}_1} p(y \mid x);\\ \beta(i,j) &= \sum_{z:(V[j],z)\in\mathcal{A}_2} p(z \mid x);\end{aligned}$$

if $\mathcal{C}[i,j] = \star$, let $\alpha(i,j), \beta(i,j) = -\infty$. For a pair of messages (m_1,m_2), let $F(m_1,m_2) = \mathcal{C}[i,j]$, where $(i,j) \in \mathcal{C}_1(m_1) \times \mathcal{C}_2(m_2)$ is the lexicographically first pair such that $\alpha(i,j), \beta)(i,j) \geq 1 - 4\epsilon_0$; if no such (i,j) exists, then let $F(m_1,m_2)$ be the first element of $\mathcal{X}$.

Joint typicality versus rejection sampling: In the standard argument [6], the indicator random variable $\mathbf{I}(k,\ell)$ stands for joint typicality of $U[k]$ and $V[\ell]$. The rejection sampling based on I_∞, has the same effect. We list below its properties.

1. $\mathbb{E}\{\mathbf{I}(k,\ell)\} \geq (1-\epsilon_\infty)2^{-I_\infty}$.

2. For all u and v,

$$\begin{aligned} \mathbb{E}\{\mathbf{I}(k,\ell) \mid U[k]=u\} &\le \sum_v p(v)\frac{p(u,v)}{2^{I_\infty}p(u)p(v)} \le 2^{-I_\infty}; && (32)\\ \mathbb{E}\{\mathbf{I}(k,\ell) \mid V[\ell]=v\} &\le \sum_u p(u)\cdot\frac{p(u,v)}{2^{I_\infty}p(u)p(v)} \le 2^{-I_\infty}; && (33)\\ \mathbb{E}\{\mathbf{I}(k,\ell)\} &\le 2^{-I_\infty}. && (34) \end{aligned}$$

3. If $\ell \neq \ell'$, then $\mathbf{I}(k,\ell)$ and $\mathbf{I}(k,\ell')$ are conditionally independent given $U[k]$; if $k \neq k'$, then $\mathbf{I}(k,\ell)$ and $\mathbf{I}(k',\ell)$ are conditionally independent given $V[\ell]$.

4. If $k \neq k'$ and $\ell \neq \ell'$, then $\mathbf{I}(k,\ell)$ and $\mathbf{I}(k',\ell)$ are independent.

Decoding: We first consider Bob's strategy for recovering m_1 on receiving the channel output $\tilde{y}$: let $D_1(\tilde{y})$ be the smallest $\tilde{m}_1$ such that there is a $k \in \mathcal{C}_{\tilde{m}_1}$ such that $(U[k],\tilde{y}) \in \mathcal{A}_1$. Similarly, Charlie's strategy is determined using the set $\mathcal{A}_2$: $D_2(\tilde{z})$ is the smallest $\tilde{m}_2$ such that there is an $\ell \in \mathcal{C}_{\tilde{m}_2}$ such that $(V[\ell],\tilde{z}) \in \mathcal{A}_2$. In both cases, if an appropriate $U[k]$ or $V[\ell]$ is not found, the answer 1 is returned. (In fact, we will show that whp there is a unique such (k,ℓ) in $[2^{R_1+r_1}]$; if there is a pair (k,ℓ) satisfying the above requirements, but different from the one used by Alice, then we will consider it an error.)

3.3 Proof of Theorem 4

Suppose a pair of messages $(m_1,m_2) \in [2^{R_1}] \times [2^{R_2}]$ is transmitted by Alice using the above scheme and is decoded as $(\tilde{m}_1,\tilde{m}_2)$ by Bob and Charlie. We wish to show that the probability (averaged over the choice of the codebook) that $(\tilde{m}_1,\tilde{m}_2) \neq (m_1,m_2)$ is at most ϵ. By the symmetry in the generation of the code book, it is enough to prove this claim for $(m_1,m_2)=(1,1)$.

We identify three sources of error. First, we regard as error those cases where there is no pair $(k,\ell) \in \mathcal{C}_1(1)\times\mathcal{C}_2(1)$ for which $\mathbf{I}(k,\ell)=1$ and $\alpha(i,j),\beta(i,j) \ge 1-4\epsilon_0$; second, it may happen that even though such a pair (k^*,ℓ^*) is found, we have $(U[k^*],\tilde{y}) \notin \mathcal{A}_1$ (here, as before, $\tilde{y}$ refers to the channel output received by Bob) or $(V[\ell^*],\tilde{z}) \notin \mathcal{A}_2$; third, Alice or Bob may not recover (k,ℓ) uniquely, for it may happen that $(U[k'],\tilde{y}) \in \mathcal{A}_1$, for some $k' \neq k$ or $(V[\ell'],\tilde{z}) \in \mathcal{A}_2$, for some $\ell' \neq \ell$. Consider the indicator random variable

$$\mathbf{J}(k,\ell) = \mathbb{I}\{\mathbf{I}(k,\ell)=1 \text{ and } \alpha(k,\ell),\beta(k,\ell) \ge 1-4\epsilon_0\},$$

and the events

$$\begin{aligned} \mathcal{E}_1 &= \text{for all } (k,\ell) \in \mathcal{C}_1(1)\times\mathcal{C}_2(1) : \mathbf{J}(k,\ell)=0;\\ \mathcal{E}_{2,B} &= \mathcal{E}_1^c \text{ and } (U[k^*],\tilde{y}) \notin \mathcal{A}_1;\\ \mathcal{E}_{2,C} &= \mathcal{E}_1^c \text{ and } (V[\ell^*],\tilde{z}) \notin \mathcal{A}_2;\\ \mathcal{E}_{3,B} &= \mathcal{E}_1^c \text{ and for some } k' \neq k^* : (U[k'],\tilde{y}) \in \mathcal{A}_1;\\ \mathcal{E}_{3,C} &= \mathcal{E}_1^c \text{ and for some } \ell' \neq \ell^* : (V[\ell'],\tilde{z}) \in \mathcal{A}_2. \end{aligned}$$

Clearly, if we exclude all the above events, then Bob and Charlie indeed recover the pair (k^*,ℓ^*) used by Alice.

Claim 1. *(i)* $\Pr\{\mathcal{E}_1\} \le 2^{-r_1-r_1+I_\infty+2} + 2^{-r_1+4} + 2^{-r_2+4} \le 36\tilde{\varepsilon}$.

(ii) $Pr\{\mathcal{E}_{2,B}\}, \Pr\{\mathcal{E}_{2,C}\} \le 4\epsilon_0$*;*

(iii) $Pr\{\mathcal{E}_{3,B}\} \le 2^{2r_1+r_2+R_1}2^{-I_\infty}2^{-I_0^B} \le \frac{\tilde{\varepsilon}}{2}$ *and* $\Pr\{\mathcal{E}_{3,C}\} \le 2^{r_1+2r_2+R_2}2^{-I_\infty}2^{-I_0^C} \le \frac{\tilde{\varepsilon}}{2}$.

Then, from our claim and the union bound, we conclude

$$\Pr\{\text{error}\} \leq 37\tilde{\varepsilon} + 8\varepsilon_0.$$

The above upper bound on the probability of error applies to every (m_1, m_2) as we average over the choices of the codebook; by linearity of expectation, this upper bound holds in expectation when (m_1, m_2) is chosen randomly. It follows that there is a fixed codebook for which the expected error is bounded by $37\tilde{\epsilon} + 8\epsilon_0$.

It remains to establish the claim above.

Consider $\mathcal{E}_1$: Since $\Pr[\alpha(k,\ell) \leq (1-4\epsilon_0)], \Pr[\beta(k,\ell) \leq (1-4\epsilon_0)] \leq \frac{1}{4}$ and $\epsilon_\infty \leq \frac{1}{4}$, we have

$$\mathbb{E}\{\mathbf{J}(k,\ell) = 1\} \geq (\frac{1}{2} - \epsilon_\infty)2^{-I_\infty} \geq 2^{-I_\infty - 2}.$$

Furthermore,

$$\mathbb{E}\{\mathbf{J}(k,\ell)\mathbf{J}(k',\ell')\} \leq \mathbb{E}\{\mathbf{I}(k,\ell)\mathbf{I}(k',\ell')\};$$

in particular, for $k' \neq k$ and $\ell' \neq \ell$,

$$\mathbb{E}\{\mathbf{J}(k,\ell)\mathbf{J}(k',\ell)\}, \mathbb{E}\{\mathbf{J}(k,\ell)\mathbf{J}(k,\ell')\} \leq 2^{-2I_\infty}.$$

Also, $\mathbf{J}(k,\ell)$ and $\mathbf{J}(k',\ell')$ are independent whenever $k \neq k'$ and $\ell \neq \ell'$. By Lemma 1,

$$\Pr\{\mathcal{E}_1\} \leq 2^{-r_1 - r_2 + I_\infty + 2} + \frac{2^{r_1} + 2^{r_2}}{2^{r_1 + r_2 - 4}}.$$

Thus, from our choice of r_1 and r_2 it now easily follows that $\Pr[\mathcal{E}_1] \leq 36\tilde{\varepsilon}$.

Consider $\mathcal{E}_{2,B}$, $\mathcal{E}_{2,C}$: It follows immediately from the definition of $\mathcal{E}_1$, that

$$\Pr\{\mathcal{E}_{2,B}\}, \Pr\{\mathcal{E}_{2,C}\} \leq 4\epsilon_0.$$

Consider $\mathcal{E}_{3,B}$, $\mathcal{E}_{3,C}$: We will focus on $\mathcal{E}_{3,B}$; similar arguments are applicable to $\mathcal{E}_{3,C}$ Fix a codebook. For $(k,\ell) \in \mathcal{C}_1(1) \times \mathcal{C}_2(1)$ such that $\mathbf{I}(k,\ell) = 1$, let

$$\begin{aligned}
\Gamma(k,\ell) &= \{y : (U(k'), y) \in \mathcal{A}_1 \text{ for some } k' \neq k\}; \\
\gamma(k,\ell) &= \sum_{y \in \Gamma(k,\ell)} p(y \mid \mathcal{C}[k,\ell]) \\
&\leq \sum_{k' \neq k} \sum_{y:(U(k'),y) \in \mathcal{A}_1} p(y \mid \mathcal{C}[k,\ell]).
\end{aligned}$$

Then, (here (k,ℓ) ranges over $\mathcal{C}_1(1) \times \mathcal{C}_2(1)$)

$$\begin{aligned}
\Pr\{\mathcal{E}_{3,B}\} &\leq \sum_{k,\ell} \mathbb{I}\{k^* = k, \ell^* = \ell\}\gamma(k,\ell) \\
&\leq \sum_{k,\ell} \mathbf{I}[k,\ell]\gamma(k,\ell).
\end{aligned}$$

[Note the last inequality involves over counting; we can afford it because of the upper bound on $r_1 + r_2$ imposed through (31).] Now, averaging over all code books, we have (below k' takes ranges over $[2^{R_1 + r_1}]$ and (k,ℓ) ranges over

$\mathcal{C}_1(1) \times \mathcal{C}_2(1))$

$$
\begin{aligned}
\Pr\{\mathcal{E}_{3,B}\} &\leq \sum_{k,\ell} \mathbb{E}\{\mathbf{I}[k,\ell]\gamma(k,\ell)\} \\
&\leq \sum_{k,\ell,k'\neq k} \sum_{u,v,u'} \sum_{y:(u',y)\in\mathcal{A}_1} p(u)p(v)\left(\frac{p(u,v)}{2^{I_\infty}p(u)p(v)}\right) p(u')p(y|f(u,v)) \\
&\leq \sum_{k,\ell,k'\neq k} \sum_{u,v,u'} \sum_{y:(u',y)\in\mathcal{A}_1} 2^{-I_\infty} p(u,v)p(u')p(y \mid f(u,v)) \\
&\leq \sum_{k,\ell,k'\neq k} \sum_{(u',y)\in\mathcal{A}_1} \left(\sum_{u,v} 2^{-I_\infty} p(u,v)p(y \mid f(u,v))\right) p(u') \\
&\leq \sum_{k,\ell,k'\neq k} 2^{-I_\infty} \left(\sum_{(u',y)\in\mathcal{A}_1} p(y)p(u')\right) \\
&\leq \sum_{k,\ell,k'\neq k} 2^{-I_\infty} 2^{-I_0} \\
&\leq 2^{r_1+r_2+R_1+r_1} 2^{-I_\infty} 2^{-I_0}. \qquad (35)
\end{aligned}
$$

Thus, from (35) and by our choice of R_1, r_1 and r_2 as mentioned in (28)–(31) the desired result follows, i.e.,

$$\Pr\{\mathcal{E}_{3,B}\} \leq \frac{\tilde{\varepsilon}}{2}.$$

4 Mutual covering

Lemma 1. *Suppose* $0 < q \leq 1$. *Let* $Z = \sum_{k=1}^{r} \sum_{\ell=1}^{s} \mathbf{J}(k,\ell)$, *where the 0-1 random variables* $\mathbf{J}(k,\ell) \in \{0,1\}$ *are such that*

$$
\begin{aligned}
\mathbb{E}\{\mathbf{J}(k,\ell)\} &\geq \alpha q\,; \\
\mathbb{E}\{\mathbf{J}(k,\ell)\mathbf{J}(k,\ell')\} &\leq q^2 \text{ whenever } \ell \neq \ell'\,; \\
\mathbb{E}\{\mathbf{J}(k,\ell)\mathbf{J}(k',\ell)\} &\leq q^2 \text{ whenever } k \neq k'\,;
\end{aligned}
$$

furthermore, $\mathbf{J}(k,\ell)$ *and* $\mathbf{J}(k',\ell')$ *are independent whenever* $k \neq k'$ *and* $\ell \neq \ell'$. *Then,*

$$\Pr\{Z = 0\} \leq \frac{1}{\alpha rsq} + \frac{r+s}{\alpha^2 rs}.$$

Proof. We will use Chebyshev's inequality. We have

$$
\begin{aligned}
\mathbb{E}\{Z\} &\geq \alpha rsq\,; && (36) \\
\mathrm{Var}\{Z\} &= \mathbb{E}\{Z^2\} - \mathbb{E}\{Z^2\} && (37) \\
&\leq \sum_{(k,\ell),(k'\ell')} (\mathbb{E}\{\mathbf{J}(k,\ell)\mathbf{J}(k',\ell')\} - \mathbb{E}\{\mathbf{J}(k,\ell)\}\mathbb{E}\{(k'\ell')\}) && (38) \\
&\leq \mathbb{E}\{Z\} + rs(r+s)q^2, && (39)
\end{aligned}
$$

where we used the fact that $\mathbb{E}\{\mathbf{J}(k,\ell)\mathbf{J}(k',\ell')\} - \mathbb{E}\{\mathbf{J}(k,\ell)\}\mathbb{E}\{(k'\ell')\}$ whenever $k \neq k'$ and $\ell \neq \ell'$. Then, by Chebyshev's inequality, we have

$$\begin{aligned}
\Pr\{Z=0\} &\leq \frac{\mathrm{Var}\{Z\}}{\mathbb{E}\{Z\}^2} \\
&\leq \frac{\mathbb{E}\{Z\}+rs(r+s)q^2}{\mathbb{E}\{Z\}^2} \\
&\leq \frac{1}{\alpha rsq}+\frac{r+s}{\alpha^2 rs}.
\end{aligned}$$

This completes the proof.

□

5 Asymptotics

As stated in the introduction our analysis immediately implies similar bounds invthe asymptotic iid and information spectrum settings. In this section, we formally verify these claims for appropriate classical-quantum channels; similar bounds also follow in the classical setting, but we do not discuss them separately.

Asymptotic iid setting: The bound derived by Savov and Wilde [3] in the iid setting, which was restated as Theorem 2 in the introduction, follows from Theorem 3 because of the following well-known asymptotic convergence results.

Theorem 5. *(Ogawa and Nagaoka [19]) Let ρ^{UB} be a classical-quantum state, and let $\rho^{U_nB_n}$ be its n-fold tensor. Then, for all $\varepsilon > 0$, we have*

$$I[U;B] = \frac{1}{n}\lim_{n\to\infty} I_0^{\epsilon}[U_n;B_n], \tag{40}$$

where $I[U;B]$ is computed with respect to the state ρ^{UB} and $I_0^{\varepsilon}[U_n;B_n]$ is computed with respect to the state $\rho^{U_nB_n}$.

Theorem 6. *(Datta [20]) For a pair of classical random variables (U,V), let represent (U_n,V_n) be n indepenedent copies of (U,V). Then,*

$$I[U;V] = \frac{1}{n}\lim_{n\to\infty} I_{\infty}^{\epsilon}[U_n;V_n]. \tag{41}$$

Remark 2. *Note that though Theorem 2 was formulated in [2]; its complete justification appeared later in [3]; their analysis which works directly in the asymptotic setting makes crucial use of the over counting argument and the analysis techniques that first appeared in our work.*

Asymptotic non-iid setting: We first review the basic definitions in the asymptotic non-iid setting and formulate the rate region. In this setting, we again have an infinite sequence of state (with respect to which the asymptotic analysis will be performed), but successive states will not be obtained by independent repetitions of a basic state. The analogs of the quantities $I_0^{\epsilon}[U;B]$ and $I_{\infty}^{\epsilon}[U;V]$ in this setting are as follows.

Let $\{U_n\}_{n=1}^{n}$ be a sequence of random variables, where U_n takes values in $\mathcal{U}^n$. Furthermore, for each n and each $u_n \in \mathcal{U}^n$, let $\rho_{u_n}^{B_n}$ be a quantum state in the Hilbert space $\mathcal{H}_n$. Let $\rho^{UB} := \left\{\rho^{U_nB_n}\right\}_{n=1}^{\infty}$ be a sequence of classical-quantum states, where

$$\rho^{U_nB_n} := \sum_{u_n\in\mathcal{U}_n} P_{U_n}(u_n)|u_n\rangle\langle u_n|^{U_n}\otimes\rho_{u_n}^{B_n}. \tag{42}$$

With respect to this sequence the analog of $\underline{\mathbf{I}}(\mathbf{U};\mathbf{B})$ is defined as follows.

Definition 8. *(Spectral inf quantum mutual information rate) The spectral inf mutual information rate for ρ^{UB} is*

$$\underline{\mathbf{I}}(\mathbf{U};\mathbf{B}) := \sup\left\{\gamma : \lim_{n\to\infty}\sum_{U_n\in\mathcal{U}_n} P_{U_n}(U_n)\mathrm{Tr}\left[\left\{\rho_{u_n}^{B_n}\succeq 2^{n\gamma}\rho^{B_n}\right\}\rho^{B_n}\right]=1\right\},$$

where $\rho^{B_n} = \mathrm{Tr}_{U_n}\left[\rho^{U_nB_n}\right]$ *and* $\left\{\rho^{B_n}_{u_n} \succeq 2^{n\gamma}\rho^{B_n}\right\}$ *is the projector onto the positive Eigen space of the operator* $\rho^{B_n}_{u_n} - 2^{n\gamma}\rho^{B_n}$.

Let $(\mathbf{U},\mathbf{V}) := \{(U_n,V_n)\}_{n=1}^{\infty}$ be a sequence of pairs of random variables where (U_n,V_n) takes values in $\mathcal{U}^n \times \mathcal{V}^n$.

Definition 9. *(Spectral sup classical mutual information rate) The spectral sup classical mutual information rate between* $\mathbf{U}$ *and* $\mathbf{V}$ *is*

$$\overline{I}(\mathbf{U};\mathbf{V}) := \inf\left\{\lambda : \lim_{n\to\infty} \Pr\left\{\frac{1}{n}\log\frac{P_{U_nV_n}}{P_{U_n}P_{V_n}} > \lambda\right\} = 0\right\},$$

where the probability is calculated with respect to $P_{U_nV_n}$.

With this, we may formulate the Marton inner bound in the information spectrum setting as follows.

Theorem 7. *Let* $\left\{\mathcal{N}(x^n) := \rho^{B_nC_n}_{x^n}\right\}_{n=1}^{\infty}$ *be a sequence of general classical-quantum broadcast channel. Let* $\{f_n\}_{n=1}^{\infty}$ *be a sequence of functions where for every* n, $f_n : (\mathcal{U}\times\mathcal{V})^n \to \mathcal{X}^n$*; consider the sate*

$$\rho^{U_nV_nB_nC_n} = \sum_{(U_nV_n)\in(\mathcal{U}\times\mathcal{V})^n} P_{U_nV_n}|U_n\rangle\langle U_n|^{U_n} \otimes |V_n\rangle\langle V_n|^{V_n} \otimes \rho^{B_nC_n}_{f_n(U_n,V_n)}.$$

Let (R_1,R_2) *be such that*

$$R_1 < \underline{\mathbf{I}}(\mathbf{U};\mathbf{B}) \tag{43}$$

$$R_2 < \underline{\mathbf{I}}(\mathbf{V};\mathbf{C}) \tag{44}$$

$$R_1 + R_2 < \underline{\mathbf{I}}(\mathbf{U};\mathbf{B}) + \underline{\mathbf{I}}(\mathbf{V};\mathbf{C}) - \overline{I}(\mathbf{U};\mathbf{V}). \tag{45}$$

Then (R_1,R_2) *is achievable (see [3] for the definition of achievable rate pair.).*

Proof. The proof immediately follows from Theorem 3 and from the observation that for every $\varepsilon \in (0,1), \gamma_n < \underline{\mathbf{I}}(\mathbf{U};\mathbf{B}), \lambda_n > \overline{I}(\mathbf{U};\mathbf{V})$ and for n large enough we have

$$\frac{1}{n}I_0^{\varepsilon}(U_n;B_n) \geq \gamma_n$$

$$\frac{1}{n}I_{\infty}^{\varepsilon}(U_n;V_n) \leq \lambda_n.$$

□

Acknowledgments

We are grateful to Vinod Prabhakaran and Mark Wilde for their comments on this work.

References

[1] K. Marton, “A coding theorem for the discrete memoryless broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, 1979.

[2] I. Savov and M. M. Wilde, “Classical codes for quantum broadcast channels,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, (Cambridge, MA, USA), pp. 721–725, July 2012.

[3] I. Savov and M. M. Wilde, “Classical codes for quantum broadcast channels.” http://arxiv.org/abs/1303.0808v3, Nov. 2014.

[4] J. Yard, P. Hayden, and I. Devetak, “Quantum broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 57, pp. 7147–7162, Oct. 2011.

[5] F. Dupuis, P. Hayden, and K. Li, "A father protocol for quantum broadcast channels," *IEEE Trans. Inf. Theory*, vol. 56, pp. 2946–2956, June 2010.

[6] A. El Gamal and Y. H. Kim, *Network Information Theory*. Cambridge, U.K: Cambridge University Press, 2012.

[7] J. Radhakrishnan, P. Sen, and N. Warsi, "One-shot marton inner bound for classical-quantum broadcast channel." http://arxiv.org/abs/1410.3248v1, Oct. 2014.

[8] S. Verdú, "Non-asymptotic achievability bounds in the multiuser information theory," in *Proc. 50th Allerton Conf. Comm. Cont. Comp.*, (Monticello, USA), Oct. 2012.

[9] S. S. Bidokhti and V. M. Prabhakaran, "Is non-unique decoding necessary?," *IEEE Trans. Inf. Theory*, vol. 60, pp. 2594–2610, May 2014.

[10] A. Lapidoth and S. Tinguely, "Sending a bivariate Gaussian over a Gaussian MAC," *IEEE Trans. Inf. Theory*, vol. 56, pp. 2714–2752, June 2010.

[11] P. Minero, S. H. Lim, and Y. H. Kim, "Hybrid coding: An interface for joint source-channel coding and network communication," *arXiv:1306.0530*, 2013.

[12] P. Grover, A. B. Wagner, and A. Sahai, "Information embedding and the triple role of control," *arXiv:1306.5018*, 2013.

[13] T. S. Han and S. Verdú, "Approximation theory of output statistics,," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 752–772, 1993.

[14] M. Hayashi and H. Nagaoka, "General formulas for capacity of claasical-quantum channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1753–1768, 2003.

[15] N. Datta and R. Renner, "Smooth Rényi entropies and the quantum information spectrum," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2807–2815, 2009.

[16] L. Wang and R. Renner, "One-shot classical-quantum capacity and hypothesis testing," *Phys. Rev. Lett.*, vol. 108, pp. 200501–200505, May 2012.

[17] N. A. Warsi, "One-shot source coding with coded side information available at the decoder," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, (Istanbul), July 2013.

[18] L. Wang, R. Colbeck, and R. Renner, "Simple channel coding bounds," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, (Seoul, Korea), June 2009.

[19] 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.

[20] N. Datta, "Min- and max-relative entropies and a new entangelement monotone," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2816–2826, June 2009.