Communicating over a Classical-Quantum MAC with State Information Distributed at the Senders

Arun Padakandla

Abstract

We consider the problem of communicating over a classical-quantum (CQ) multiple access channel with classical state information non-causally available at the transmitters, henceforth referred to as a QMSTx. We undertake a Shannon-theoretic study and focus on the problem of characterizing inner bounds to the capacity region of a QMSTx. We propose a new coding scheme based on *union coset codes* - codes possessing algebraic closure properties and derive a new inner bound that subsumes the largest known inner bound based on IID random coding. We identify examples for which the derived inner bound is strictly larger.

I. Introduction

Consider the scenario depicted in Fig. 1, wherein a pair of distributed transmitters (Txs) are required to communicate independent classical messages over a classical-quantum (CQ) multiple access channel (MAC). In addition to the symbols X_1 and X_2 input by Tx 1 and 2 respectively, the quantum state provided to the receiver (Rx) is governed by a pair of classical jointly distributed random states S_1, S_2 whose evolution over time is independent and identically distributed (IID). Specifically, if the channel is in state s_1, s_2 and Txs $1, 2$ choose input symbols x_1, x_2 respectively, then the Rx receives the quantum state $\rho_{x_1 x_2 s_1 s_2}$. Tx j is provided the entire sequence of realizations of the component S_j non-causally, while the Rx remains uninformed of the states. We undertake a Shannon theoretic study of this quantum channel, henceforth referred to as a QMSTx. Our focus is on the problems of designing efficient coding schemes and characterizing inner bounds to the capacity region of the QMSTx.

A QMSTx is a CQ MAC with Tx state information. The study of channels with Tx state information has evinced considerable interest [1]–[5] over the years and has had a large influence on the design of efficient coding schemes in information theory. The investigation of a point-to-point (PTP) classical channel with Tx state information (CSTx) (Fig. 2) led Gel'fand and Pinsker [2] to their ingenious technique of partitioning channel codes. This Gel'fand-Pinsker technique forms a core component of the current known best coding scheme for the classical [6] and quantum broadcast channels [7], [8] and is also employed in other network scenarios [9]–[11]. In fact, the Gel'fand-Pinsker technique has found utility even in storage applications with defective memory cells [12]. These works and their impact motivate our study of the QMSTx. As we shall discuss, QMSTx facilitates the development of a new unconventional coding scheme. Specifically, we design and analyze a new decoding POVM that yields a strictly larger inner bound to the capacity region of the QMSTx in comparison to the conventional one.

The Gel'fand-Pinsker technique remains to be the best known technique to exploit Tx state information and is optimal for communication over both the CSTx and its CQ analogue - the QSTx. See Fig. 2. In regards to the Rx, since a pair of independent messages need to be communicated over a QMSTx, it is natural to build a simultaneous (joint) decoding POVM. We are thus led to a natural coding scheme for a QMSTx wherein the two Txs incorporate Gel'fand-Pinsker's channel code partitioning technique and the Rx adopts a joint decoding POVM to recover the messages. The conventional long established approach in information theory is to incorporate these techniques by building IID random codes, also referred to herein as unstructured (IID) codes. One is thus led to partitioning the two unstructured IID random codes, building the corresponding joint decoding POVM, analyzing performance to characterize inner bounds to the capacity region of a QMSTx. The reader is referred to Thm. 1 for a characterization of this *unstructured coding* inner bound.

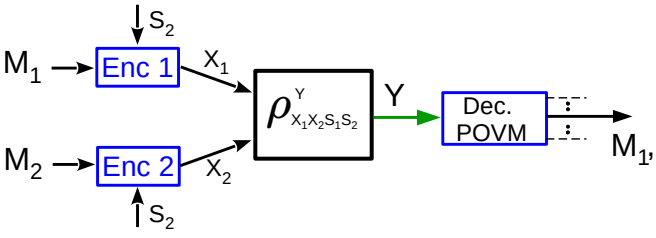

Fig. 1. A QMSTx wherein two Txs observe jointly distributed random classical states that evolve IID over time and are required to communicate independent messages to a Rx.

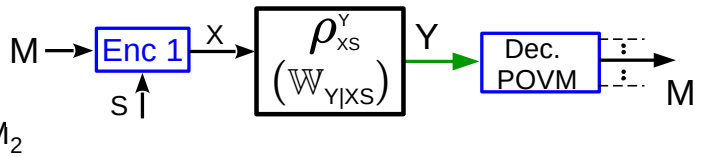

Fig. 2. A general CQ PTP channel with Tx state information (QSTx). When the collection $(\rho_{x,s} : (x,s) \in \mathcal{X} \times \mathcal{S})$ of quantum states are commuting, this channel reduces to a classical PTP channel with Tx state information (CSTx) characterized via a stochastic matrix $\mathbb{W}_{Y|XS}$ (denoted in braces). Gel'fand and Pinsker [2] focused their study on the CSTx.

The focus of our article is to step beyond this conventional use of unstructured codes and design a new coding scheme for the QMSTx based on *union coset codes* (UCC) - an ensemble of *structured* codes possessing algebraic closure properties. We analyze the information theoretic performance of the proposed structured coding scheme to derive a new inner bound (Thms. 2, 4) to the capacity region of the QMSTx. The inner bound we characterize in Thm. 4 subsumes the unstructured coding inner bound. In Sec. III-B and IV-E, we identify non-commutative examples for which the structured coding scheme strictly outperforms the IID random coding scheme. Specifically, we demonstrate that the derived inner bound for these examples is strictly larger than that achievable using unstructured IID random codes. These findings build on our earlier work [13], [14] and maybe viewed as another step [15], [16] in our pursuit of designing coding schemes based on coset codes for network CQ communication.

We now highlight the import and significance of our contributions. Beginning from Shannon's work and through most of the ensuing six decades, information theoretic study has been largely restricted to analyzing unstructured IID coding schemes. Inspired by an ingenious work of Körner and Marton [17], structured coding schemes have been designed for several classical multi-terminal channels [14]–[16], [18]–[22] in the last two decades and have been proven to strictly outperform [4], [18], [22], [23] conventional unstructured IID random coding schemes. Owing to the dominant influence of IID random codes and the difficulty of performance analysis in the quantum setting, coding schemes for multi-terminal quantum channels are largely based on unstructured IID random codes. Our work contributes to the development of structured quantum coding strategies in the CQ setting.

Secondly, the use of coset codes and the role of algebraic closure properties in a QMSTx is unique. Coset codes have facilitated higher rates in communication scenarios wherein a compressive bi-variate function of the messages or codewords have to be decoded. For instance, on both the $3-$user interference [15], [23] and broadcast channels [22], coset codes enable efficient decoding of the bi-variate interference. QMSTx is a CQ MAC wherein both messages need to be decoded and decoding a compressive bi-variate function of either the codewords or the messages can lead to obfuscation of the messages. Indeed, coset codes have no role in communication over a CQ-MAC without Tx states. It is therefore natural to question the utility of structured codes in communicating over a QMSTx. As we illustrate through a self-contained discussion in Sec. III, our findings demonstrate how algebraic closure properties can be exploited to *efficiently sieve relevant information* and thereby facilitate enhanced communication over a QMSTx - a utility that can potentially be exploited in other quantum communication scenarios.

Thirdly, this study enables us to enrich the family of coset codes for CQ communication beyond nested coset codes (NCC) [15], [16] and partitioned coset codes (PCC) [24] studied recently. As elaborated in [14] and recent works [25], [26], NCC or PCC based coding schemes for a classical analogue of a QMSTx, i.e., a classical MAC with states, can be strictly inferior to a UCC based coding scheme. We have taken this cue to propose and design UCC based coding schemes for communication over CQ channels. As an auxiliary result, we also prove (Thm. 5) that UCCs achieve capacity of a single Tx version of the QMSTx, henceforth referred to as the QSTx channel. Lastly, our findings maybe viewed as developing new coding schemes to handle diverse CQ network scenarios arising in an eventual quantum communication network.

Since the early work of Shannon [1], the study of channels with Tx state information continues [3], [27]–[29] to evince interest. Recently Anshu, Hayashi and Warsi [30] have studied the problem of secure communication over fully quantum wiretap channel with Tx state information and obtain error exponents via the technique of simultaneous pinching. See [31]–[33] for analogous works on classical channels. Using the method of types and tools developed by Nötzel [34], Boche, Cai and Nötzel [5] have proved achievability of the Gel'fand-Pinsker inner bound for the QSTx (Fig. 2). More importantly, their work [5] highlights the difference between the causal and non-causal availability of state information at the Tx in regards to the single-letterization of the capacity. Our focus is on designing a new coding scheme and characterizing new single-letter inner bounds. We do not comment on the optimality of the inner bounds derived herein.

Our presentation is pedagogical. We begin with preliminaries - notation and problem statement - in Sec. II. Through a self-contained discussion in the context of a carefully chosen example, Sec. III illustrates the main ideas of our work and the role of algebraic closure in the proposed UCC coding scheme. A general coding scheme for a QMSTx consists of two layers - unstructured codes and UCC. We first present a simplified coding scheme involving only the UCC layer in Sec. IV and derive a new inner bound. In Sec. V, we present a larger inner bound that subsumes the former and comprises of both unstructured and UCC layers. Leveraging techniques in Sec. IV and V, we prove that UCCs achieve the Gel'fand-Pinsker inner bound in Sec. VI. Being a stand-alone section, Sec. VI assists a reader interested only in the latter and demonstrates the versatility of our proof techniques.

II. Preliminaries and Problem Statement

We supplement standard notation in QIT, for example as in [35], with the following. For $n \in \mathbb{N}$, $[n] \triangleq \{1, \cdots, n\}$. $\mathcal{F}_q$ denotes a generic finite field of size q, where q is a prime power and $\oplus$ denotes addition within the finite field in context. For $a, b \in \mathcal{F}_q$, $a \ominus b \triangleq a \oplus (-b)$, where $(-b) \in \mathcal{F}_q$ is the additive inverse of $b \in \mathcal{F}_q$. Existence of finite fields for any prime power can be verified via standard books in Algebra such as [36, Lemma 7.1.4]. For a Hilbert space $\mathcal{H}$, $\mathcal{L}(\mathcal{H}), \mathcal{P}(\mathcal{H})$ and $\mathcal{D}(\mathcal{H})$ denote the collection of linear, positive and density operators acting on $\mathcal{H}$ respectively. We let an underline denote an appropriate aggregation of pairs of objects. For example, $\underline{\mathcal{U}} \triangleq \mathcal{U}_1 \times \mathcal{U}_2$ denotes the Cartesian product for sets, $\underline{x} \triangleq (x_1, x_2) \in \underline{\mathcal{X}}$ and $\underline{x}^n \triangleq (x_1^n, x_2^n)$. The specific aggregation will be clear from context. For $j \in \{1, 2\}$, we let $\dot{j}$ denote the complement index, i.e., $\{j, \dot{j}\} = \{1, 2\}$. For an event (set) $\mathcal{A} \subseteq \Omega$, we let $\overline{\mathcal{A}} = \Omega \setminus \mathcal{A}$ denote its complement. We abbreviate classical-quantum, point-to-point, independent and identically distributed, probability mass function, orthonormal basis, spectral decomposition as CQ, PTP, IID, PMF, ONB, SCD respectively. For a quantum state $\theta^{XY} \in \mathcal{D}\left(\mathcal{H}_{\mathcal{X}} \otimes \mathcal{H}_{\mathcal{Y}}\right)$, $\theta^X \in \mathcal{D}(\mathcal{H}_{\mathcal{X}})$ and $\theta^Y \in \mathcal{D}(\mathcal{H}_{\mathcal{Y}})$ denote the component states, i.e, $\theta^X = \mathrm{tr}_Y(\theta^{XY})$ and $\theta^Y = \mathrm{tr}_X(\theta^{XY})$. We let $H(X,Y)_\theta \triangleq -\mathrm{tr}\big(\theta^{XY} \log\big(\theta^{XY}\big)\big)$, $H(X)_\theta \triangleq -\mathrm{tr}\big(\theta^X \log\big(\theta^X\big)\big)$ denote the Von Neumann entropy of the joint and component quantum states respectively. We let $H(Y|X)_\theta \triangleq H(X,Y)_\theta - H(X)_\theta$ and $I(X;Y)_\theta \triangleq H(X)_\theta + H(Y)_\theta - H(X,Y)_\theta$ denote the conditional quantum entropy and quantum mutual information respectively. The notions of typicality, typical projectors and the associated facts are stated in Appendix A.

Consider a (generic) *QMSTx* specified through (i) two finite input sets $\mathcal{X}_1, \mathcal{X}_2$, (ii) two finite sets $\mathcal{S}_1, \mathcal{S}_2$ of states, (iii) a PMF $\mathtt{p}_{\underline{S}}(\cdot)$ on $\underline{\mathcal{S}}$, (iii) a collection $(\rho_{\underline{xs}} \triangleq \rho_{x_1 x_2 s_1 s_2} \in \mathcal{D}(\mathcal{H}_Y) : (\underline{x}, \underline{s}) \in \underline{\mathcal{X}} \times \underline{\mathcal{S}})$ of density operators and (iv) cost functions $\kappa_j : \mathcal{X}_j \times \mathcal{S}_j \to [0, \infty)$ for $j \in [2]$. The cost function is additive, i.e., having observed the state sequence s_j^n the cost incurred by sender j in preparing the state $\otimes_{t=1}^n \rho_{\underline{x}_t \underline{s}_t}$ is $\overline{\kappa}_j(x_j^n, s_j^n) \triangleq \frac{1}{n}\sum_{t=1}^n \kappa_j(x_{jt}, s_{jt})$. Reliable communication on a QMSTx entails identifying a code. Throughout, except for the examples, no assumption is made on the structure of the collection $(\rho_{\underline{xs}} \in \mathcal{D}(\mathcal{H}_Y) : (\underline{x}, \underline{s}) \in \underline{\mathcal{X}} \times \underline{\mathcal{S}})$ of density operators.

Definition 1. *An $(n, \underline{\mathcal{M}}, \underline{e}, \lambda)$ QMSTx code consists of two message index sets $\mathcal{M}_j : j \in [2]$, two encoder maps $e_j : [\mathcal{M}_j] \times \mathcal{S}_j^n \to \mathcal{X}_j^n$ and a decoder POVM $\lambda \triangleq \{\lambda_{\underline{m}} = \lambda_{m_1, m_2} \in \mathcal{P}(\mathcal{H}^{\otimes n}) : \underline{m} \in \underline{\mathcal{M}}\}$. The average error probability of the code is*

$$\overline{\xi}(\underline{e}, \lambda) \triangleq 1 - \frac{1}{|\underline{\mathcal{M}}|} \sum_{\underline{m} \in \underline{\mathcal{M}}} \sum_{\underline{s}^n \in \underline{\mathcal{S}}^n} \mathtt{p}_{\underline{S}}^n(\underline{s}^n)\, \mathrm{tr}\big(\lambda_{\underline{m}} \rho_{\underline{m}, \underline{s}^n}\big).$$

where $\rho_{\underline{m}, \underline{s}^n} \triangleq \otimes_{t=1}^n \rho_{\underline{x}_t \underline{s}_t}$ and $(x_{j1}, \cdots, x_{jn}) = e_j(m_j, s_j^n)$. Average cost incurred by Tx j in transmitting m_j is $\tau_j(e_j|m_j) \triangleq \sum_{s_j^n} \mathtt{p}_{S_j}^n(s_j^n) \kappa_j(e_j(m_j, s_j^n), s_j^n)$ and the average cost incurred by Tx j is $\tau_j(e_j) \triangleq \frac{1}{|\mathcal{M}_j|}\sum_{m_j} \tau_j(e_j|m_j)$.

The object of interest is the capacity region of a QMSTx defined below. In this article, we focus on characterizing inner bounds to the capacity region of a QMSTx.

Definition 2. *A rate-cost quadruple $(\underline{R}, \underline{\tau}) \in [0, \infty)^4$ is achievable if there exists a sequence of QMSTx codes $(n, \underline{\mathcal{M}}^{(n)}, \underline{e}^{(n)}, \lambda^{(n)})$ for which $\lim_{n \to \infty} \overline{\xi}(\underline{e}^{(n)}, \lambda^{(n)}) = 0$,*

$$\lim_{n \to \infty} n^{-1} \log \mathcal{M}_j^{(n)} = R_j, \text{ and } \lim_{n \to \infty} \tau_j(e_j^{(n)}) \leq \tau_j.$$

The capacity region $\mathscr{C}$ of the QMSTx is the set of all achievable rate-cost vectors and $\mathscr{C}(\underline{\tau}) \triangleq \{\underline{R} : (\underline{R}, \underline{\tau}) \in \mathscr{C}\}$.

III. Role of Algebraic Closure

In this section, we explain *how* and *why* structured codes can facilitate enhanced communication over a QMSTx. We begin by reviewing the best known unstructured coding scheme.

A. Joint Decoding of Unstructured Codes

A QMSTx being a 'MAC extension' of a QSTx (Fig. 2) [5], a coding scheme for a QMSTx can be obtained by combining the Gel'fand-Pinsker technique [2] with a simultaneous decoding POVM of a MAC [37, Thm. 2]. Specifically, Tx j builds a U_j-code (Fig. 4) on an auxiliary set $\mathcal{U}_j$. The U_j-code comprising of $2^{n(R_j + B_j)}$ codewords is partitioned into 2^{nR_j} bins. Message $m_j \in [2^{nR_j}]$ indexes a bin and Tx j looks for a codeword within this bin that is jointly typical with the state sequence s_j^n. The chosen codeword, denoted as $u_j^n(m_j, s_j^n)$, and the state sequence s_j^n are mapped to an input sequence in $\mathcal{X}_j^n$. The latter mapping is deterministic, i.e., each code comes

w	t	$\gamma(w,t)$
0	0	$\|0\rangle\langle 0\|$
0	1	$\|v_\theta^\perp\rangle\langle v_\theta^\perp\|$
1	0	$\|1\rangle\langle 1\|$
1	1	$\|v_\theta\rangle\langle v_\theta\|$

Fig. 3. $\rho_{x_1x_2s_1s_2} = \gamma(x_1 \oplus x_2, s_1 \oplus s_2)$

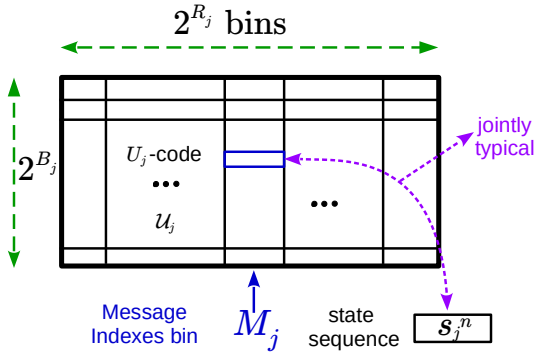

Fig. 4. Encoding rule for sender j.

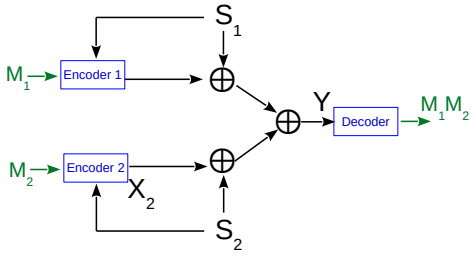

Fig. 5. S_1, S_2 are mutually independent and uniformly distributed. Input X_j is constrained to average Hamming weight of $\tau < \frac{1}{2}$.

with a deterministic mapping that maps pairs of U_j-codewords and state sequences to input 'codewords' on $\mathcal{X}_j^n$. The decoder POVM performs simultaneous decoding on the U_1, U_2-codebooks. Adopting the decoding POVM proposed in proof of [37, Thm. 2], analyzing the error probability one can derive the inner bound characterized in Thm. 1. The latter is the largest known inner bound achievable via any unstructured coding scheme. For a proof of Thm. 1, the reader is referred to proof of Thm. 4 provided in Sec. V, wherein a larger inner bound that subsumes the one characterized below is proven achievable. In Thm. 1 below, component Y is the received quantum state.

Theorem 1. *A rate-cost quadruple $(\underline{R}, \underline{\tau}) \in \mathcal{A}_u \subseteq [0,\infty)^4$ is achievable if there exists finite sets $\mathcal{Q}, \mathcal{U}_1, \mathcal{U}_2$, PMF p_Q on $\mathcal{Q}$, conditional distributions $p_{X_j,U_j|S_jQ}$ on $\mathcal{X}_j \times \mathcal{U}_j$ for $j \in [2]$ such that $p_{\underline{S}Q\underline{U}\underline{X}}(\underline{s},\underline{u},\underline{x}) = \mathrm{p}_{\underline{S}}(\underline{s}) p_Q(q) \prod_{j=1}^2 p_{X_jU_j|S_jQ}(x_j,u_j|s_j,q)$ with respect to which*

$$R_j < I(U_j;Y,U_{\bar{j}}|Q)_\Upsilon - I(U_j;S_j|Q)_\Upsilon, \mathbb{E}\{\kappa_j(X_j,S_j)\} \leq \tau_j, \;\; R_1+R_2 < I(\underline{U};Y|Q)_\Upsilon - I(U_1,U_2;S_1,S_2|Q)_\Upsilon, \quad (1)$$

for $j \in [2]$, where all entropies are computed with respect to the state

$$\Upsilon^{Y\underline{X}SUQ} \triangleq \sum_{\underline{s},\underline{x},\underline{u},q} p_{\underline{S}U\underline{X}Q}(\underline{s},\underline{u},\underline{x},q) \rho_{\underline{x}\underline{s}} \otimes |\underline{x}\;\underline{s}\;\underline{u}\;q\rangle\langle \underline{x}\;\underline{s}\;\underline{u}\;q| \,. \quad (2)$$

Remark 1. *We highlight two aspects of the above coding strategy, specifically in regards to the decoding. The strategy of decoding the pair $U_1, U_2-$of codewords implies that effective communication is happening over the CQMAC $(U_1,U_2) - Y$ channel specified through the collection $(\delta_{u_1u_2} \in \mathcal{D}(\mathcal{H}_Y) : (u_1,u_2) \in \mathcal{U}_1 \times \mathcal{U}_2)$ where $\delta_{u_1,u_2} = \sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}|\underline{U}}(\underline{x},\underline{s}|\underline{u}) \rho_{\underline{x},\underline{s}}$. In contrast to a 'plain' QMAC without states, the presence of states S_1, S_2 implies that we reserve multiple pairs $(U_1,U_2)-$of codewords for a single message pair $\underline{m} = (m_1,m_2)$. Indeed, any one of $2^{n[I(U_1;S_1)+I(U_2;S_2)]}$ codeword pairs can be used to communicate the message pair $\underline{m} = (m_1,m_2)$. The second aspect relates to what 'chunk of the output space' gets reserved for a message pair $\underline{m} = (m_1,m_2)$, colloquially referred to as the 'fan-out' of $\underline{m}$. Suppose $c_j(m_j) : j = 1,2$ is the bin of codewords at Tx j associated with message m_j. The subspace corresponding to the span of the union of the conditional typical projectors $\pi_{p_{\underline{U}},\eta}^{\delta_{\underline{u}^n}}$ of $\delta_{\underline{u}^n} : \underline{u}^n \in c_1(m_1) \times c_2(m_2)$, i.e.*

$$\mathscr{S}_{\underline{m}} \triangleq \{|g\rangle \in \mathcal{H}_{\mathcal{Y}}^{\otimes n} : \langle g|\pi_{p_{\underline{U}},\eta}^{\delta_{\underline{u}^n}}|g\rangle > 0\},$$

has been reserved for a single message pair $\underline{m}$. In order to enlarge the capacity region, it is desirable to keep both the number of codeword pairs reserved for any message pair $\underline{m}$ small and the dimension of $\mathscr{S}_{\underline{m}}$ as least as possible. Indeed, this would enable pack a larger number of orthonormal fan-outs in the output space $\mathcal{H}_{\mathcal{Y}}^{\otimes n}$.

B. Binary Double Dirty MAC

Our discussion for the following example portrays the deficiency of unstructured codes and the role of structure.

Example 1. *Let $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{S}_1 = \mathcal{S}_2 = \{0,1\}$, $\mathrm{p}_{\underline{S}}(\underline{s}) = \frac{1}{4}$ for every $\underline{s} \in \underline{\mathcal{S}}$, $\mathcal{H}_Y \triangleq \mathbb{C}^2$ denote the qubit space, $|v_\theta\rangle = [\cos\theta \;\; \sin\theta]^T \in \mathbb{C}^2$ and $|v_\theta^\perp\rangle = [\sin\theta \;\; -\cos\theta]^T \in \mathbb{C}^2$. For $(\underline{x},\underline{s}) \in \{0,1\}^4$, let $\rho_{x_1x_2s_1s_2} = \gamma(x_1 \oplus x_2, s_1 \oplus s_2) \in \mathcal{D}(\mathbb{C}^2)$, where $\gamma(\cdot,\cdot) \in \mathcal{D}(\mathbb{C}^2)$ is provided in Figure 3, $\oplus$ denotes addition in the binary field $\mathcal{F}_2$ and the cost function $\kappa_j(x_j,s_j) = \mathbb{1}_{\{x_j=1\}}$ is the Hamming weight function. For a $\tau \in (0,\frac{1}{2})$, what is $\mathscr{C}(\tau,\tau)$?*

We begin with the $\theta = 0$ case before discussing the non-commuting $\theta \in (0,\frac{\pi}{2})$ case. $\theta = 0$ case corresponds to the classical channel first studied by Philosof and Zamir [4]. The following discussion describes their findings.

Case $\theta = 0$: Since the collection $\left(\rho_{\underline{x}\underline{s}} : (\underline{x}, \underline{s}) \in \{0,1\}^4\right)$ is commuting, we identify this as a classical MAC with distributed states whose output $Y \in \{0,1\}$, inputs $X_1, X_2 \in \{0,1\}$ and states $S_1, S_2 \in \{0,1\}$ are related as $Y = X_1 \oplus S_1 \oplus X_2 \oplus S_2$. See Figure 5. S_1, S_2 are uniformly distributed and the average Hamming weight of the inputs is constrained to $\tau < \frac{1}{2}$. This implies that, having observed the uniformly distributed state sequence $S_j^n \in \{0,1\}^n$, the encoder can input a sequence $X_j^n \in \{0,1\}^n$ of average Hamming weight at most $\tau < \frac{1}{2}$. The latter constraint precludes the Txs from negating the effect of the state. What rate pairs are then achievable?

We first study the best unstructured coding scheme and characterize the corresponding largest known inner bound. Towards that end, we are required to identify an optimal choice of parameters in Thm. 1 for Ex. 1. Observe that the effective classical channel of Ex. 1 is a 'MAC extension' of a CPSTx whose output $Y \in \{0,1\}$, Hamming cost-constrained input $X \in \{0,1\}$ and uniformly distributed state $S \in \{0,1\}$ are related as $Y = X \oplus S$. Philosof and Zamir [4] proved that the best unstructured coding scheme for Ex. 1 is obtained by replicating, at both the Txs, the capacity achieving scheme for the CPSTx. Specifically, they prove the optimal choice of parameters in Thm. 1 for Ex. 1 to be binary auxiliary sets $\mathcal{U}_1 = \mathcal{U}_2 = \{0,1\}$, $p_{U_j|S_j}(1|0) = p_{U_j|S_j}(0|1) = \tau = 1 - p_{U_j|S_j}(0|0) = 1 - p_{U_j|S_j}(1|1)$ and $X_j = U_j \oplus S_j$ for $j \in [2]$.

We now detail the coding scheme corresponding to the above choice to shed light on its deficiency. See Fig. 6. To communicate at rate $R_j < h_b(\tau)$, Tx j randomly partitions the entire set of 2^n sequences into 2^{nR_j} bins. The message m_j indexes the bin within which the sender looks for a codeword that is within an average Hamming distance of τ from the observed state sequence. Since each bin contains $2^{n(1-R_j)} > 2^{n(1-h_b(\tau))}$ sequences chosen randomly, the sender finds such a codeword with probability that approaches 1 exponentially. Indeed, this can be proved via a simple second moment method as done in [13, Upper Bound on ϵ_{2j} in Appendix B]. Let U_j^n denote the chosen codeword and S_j^n the observed state sequence. Tx j inputs $X_j^n = U_j^n \oplus S_j^n$ on the channel. The choice of the U_j-codeword guarantees that the Hamming weight constraint is met.

What is the maximum sum rate the above unstructured coding scheme can achieve? Recall that each message m_j of sender j is assigned a bin of U_j-codewords, with at least $2^{n(1-R_j)} > 2^{n(1-h_b(\tau))}$ codewords. Observe that the channel relationship $Y = X_1 \oplus S_1 \oplus X_2 \oplus S_2$ implies that the received vector is $Y^n = U_1^n \oplus U_2^n$. Fan-out - the space of received sequences occupied by a *single message pair* $\underline{m} = (m_1, m_2)$ - is therefore got by adding all possible codeword pairs in the two bins indexed by $\underline{m}$. Since the codewords in each bin is picked uniformly and independently without any joint structure, every pair yields with high probability a distinct sum, resulting in the range of this addition to be of size $2^{n(2-R_1-R_2)} > 2^{2n(1-h_b(\tau))}$. Since the 'fan-out' of every message pair is of size at least $2^{2n(1-h_b(\tau))}$, we cannot hope to pack more than $\frac{2^n}{2^{2n(1-h_b(\tau))}}$ fan-outs in the binary output space resulting in the following fact.

Fact 1. *Consider Ex. 1 with average Hamming cost constraint $\tau < \frac{1}{2}$. Any rate pair (R_1, R_2) achievable by unstructured coding schemes satisfies $R_1 + R_2 < \mathit{uce}\{\max\{0, 2h_b(\tau) - 1\}\}$ where $\mathit{uce}\{f(\tau)\}$ denotes the upper convex envelope of the function $f(\tau)$. See [4] for a proof.*

We now present a linear coding scheme that can achieve any rate pair (R_1, R_2) satisfying $R_1 + R_2 < h_b(\tau)$. In the sequel, we provide a more descriptive presentation. In Appendix B, we provide a formal illustration. See Fig. 6. For simplicity, we describe achievability of the rate pair $(h_b(\tau), 0)$. Our coding scheme is identical to the unstructured coding scheme with two key differences. The first key difference is that the bins of each sender's codebook are chosen to be *cosets of a common linear code*. Let λ_2 denote a linear code of rate $1 - h_b(\tau)$ whose cosets can quantize a uniform source to with an average Hamming distortion of τ. In other words, a uniformly and randomly chosen coset of λ_2 contains a codeword within an average Hamming distance of τ of the observed state sequence with high probability. See [38] or [14] wherein, the existence of linear codes of rate $1 - h_b(\tau)$ whose cosets can quantize any uniformly distributed source within a Hamming distance of τ, is proven. We employ cosets of such a linear code to quantize the two state sequences that are both uniformly distributed. Since sender 2 has no message to transmit, it is provided with just λ_2 that serves as its only bin. Sender 1 is provided with all of the $2^{nh_b(\tau)}$ cosets of λ_2, each of which serves as its bins. The encoding is identical to that for unstructured coding.

We shall now not decode the pair $U_1, U_2-$of codewords as done with unstructured IID coding. From the received vector $Y^n = U_1^n \oplus U_2^n$, the decoder has to only figure out which coset of λ_2 did Tx 1 choose its codeword from. Indeed, the index of the bin or coset Tx 1 chooses its codeword from is user 1's message. The bins of user 1's code being cosets of linear code λ_2, from which user 2 picks its codewords, the received sequence $Y^n = U_1^n \oplus U_2^n$ is found in exactly the same coset (or bin) from which user 1 picked its codeword. The Rx can therefore call out the

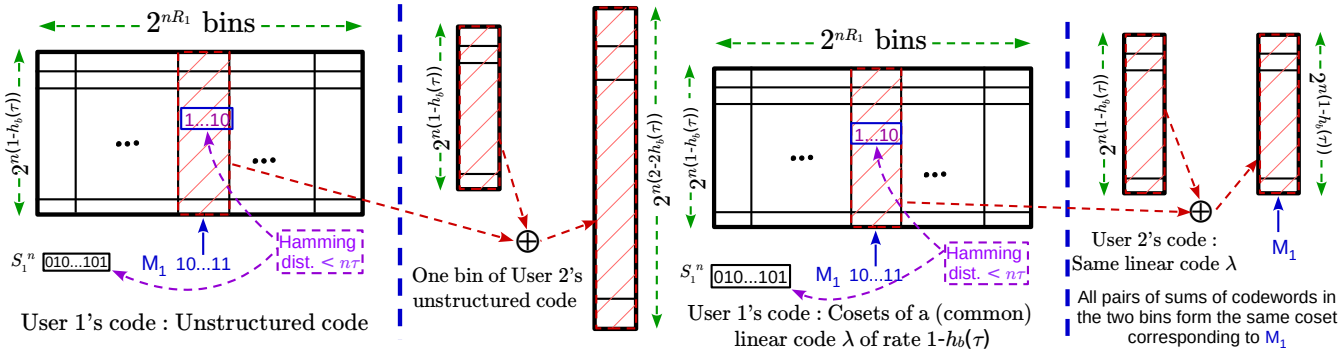

Fig. 6. Code on the left depicts User 1's code and the bin on the right depicts one bin of User 2's code. When a bin of User 1's code is added to a bin of User 2's code, the rate of resulting collection of vectors doubles to $2(1-h_b(\tau))$ as depicted by the long bin on the right.

Fig. 7. User 2 employs a linear code λ as its only bin and user 1 employs cosets of λ as the bins of its code. When a user 1's bin is added to user 2's only bin, the resulting collection of codewords is the same coset of the same rate $1-h_b(\tau)$.

index of the coset in which it observes the received sequence Y^n as user 1's message. Let us analyze the achieved rate. Since the channel is noiseless, sender 1 may employ all cosets of λ_2 and therefore communicate at rate $h_b(\tau)$ which is larger than $2h_b(\tau)-1$ for all $\tau \in (0, \frac{1}{2})$.

Fact 2. *Consider Ex. 1 with average Hamming cost constraint $\tau < \frac{1}{2}$. Any rate pair (R_1, R_2) satisfying $R_1 + R_2 < \mathit{uce}\{\max\{0, h_b(\tau)\}\}$ where $\mathit{uce}\{f(\tau)\}$ denotes the upper convex envelope of the function $f(\tau)$ is achievable via the above coset coding strategy.*

How is it that we are able to achieve a higher rate? First, our approach of decoding by identifying the coset in which the received vector $U_1^n \oplus U_2^n$ resides is altering the effective CQMAC on which we are communicating. We are decoding the sum of the codewords chosen by the two Txs, not the pair. Effective communication is therefore happening over the induced CQMAC $(\Xi_w \in \mathcal{D}(\mathcal{H}_\mathcal{Y}) : w \in \mathcal{W})$ wherein

$$\Xi_w = \sum_{u_1,x_1,s_1} \sum_{u_2,x_2,s_2} \frac{1}{4} p_{U_1X_1|S_1}(u_1,x_1|s_1) p_{U_2X_2|S_2}(u_2,x_2|s_2) \gamma(x_1 \oplus x_2, s_1 \oplus s_2) \mathbb{1}_{\{w=u_1 \oplus u_2\}},$$

which, in the $\theta = 0$ case with the above choice of parameters is the CQMAC $(\Xi_0 = |0\rangle\langle 0|, \Xi_1 = |1\rangle\langle 1|)$. In essence, we are communicating over the effective channel $U_1 \oplus U_2 - Y$ channel. Secondly, note that the number of codeword pairs associated with any message pair $\underline{m}$ remains $2^{2n(1-h_b(\tau))}$. Indeed, the rate of each bin - both in this and the previous unstructured coding schemes - are identical. The key difference is however the size of the chunk of the output space allocated to any message pair $\underline{m}$. Algebraic closure - the property that two cosets of a linear code of rate $(1-h_b(\tau))$ when added yields another coset with the same number $2^{n(1-h_b(\tau))}$ of codewords - ensures that every message pair has a fan-out of size $2^{n(1-h_b(\tau))}$ sequences within the $\{0,1\}^n-$space.

If we contrast the fan-outs of the linear coding scheme - $2^{n(1-h_b(\tau))}$ - and the unstructured IID coding scheme - $2^{n(2-2h_b(\tau))}$, one can account for the difference in the achievable sum rate. Indeed, the linear coding scheme achieves a rate $h_b(\tau)$ which exceeds the unstructured IID coding sum rate $2h_b(\tau)-1$ by $1-h_b(\tau)$ - the difference in the rates of the two fan-outs.

Going further, observe that exponentially many pairs of codewords from λ_2 and the coset chosen by sender 1 have the same sum, the Rx cannot disambiguate the pair of codewords chosen by the Txs. It can only disambiguate the sum $U_1^n \oplus U_2^n$ and not the pair. This implies that the structured coding scheme is forcing the Rx to forgo certain information that it was able to decode in the unstructured coding scheme. However, the structured coding is cleverly designed so as to ensure that this forgone information is not of the messages, but of the states S_1, S_2. Attempting to decode the pair of messages by decoding a compressive bivariate function - mod-2 sum - of the chosen codewords, boost the information rate of the messages while suppressing the amount of information it gathers of the states, and exploiting algebraic closure to ensure this are therefore the central aspects of the structured coding scheme.

Case $\theta \in (0, \frac{\pi}{2})$: The arguments in [4] can be used to prove that the optimal choice of parameters in Thm. 1 for this case too is $\mathcal{U}_1 = \mathcal{U}_2 = \{0,1\}$, $p_{U_j|S_j}(1|0) = p_{U_j|S_j}(0|1) = \tau = 1 - p_{U_j|S_j}(0|0) = 1 - p_{U_j|S_j}(1|1)$ and

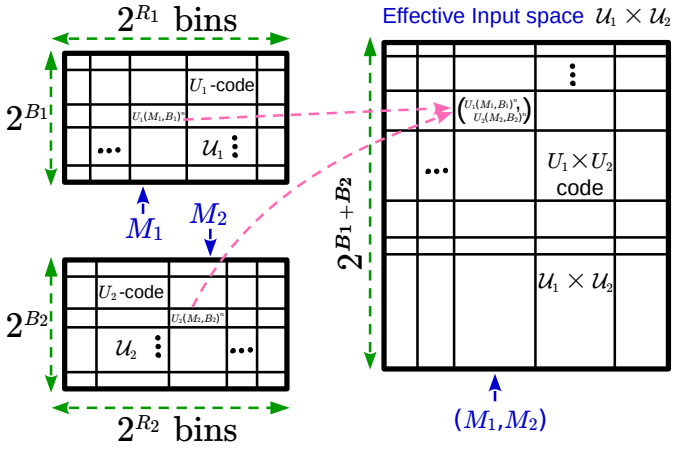

Fig. 8. Unstructured IID coding strategy of decoding the pair of U_1, U_2-codewords implies communication effectively happens via $U_1 \times U_2-$codebook. $2^{n(B_1+B_2)}$ codeword pairs are allocated for each message pair.

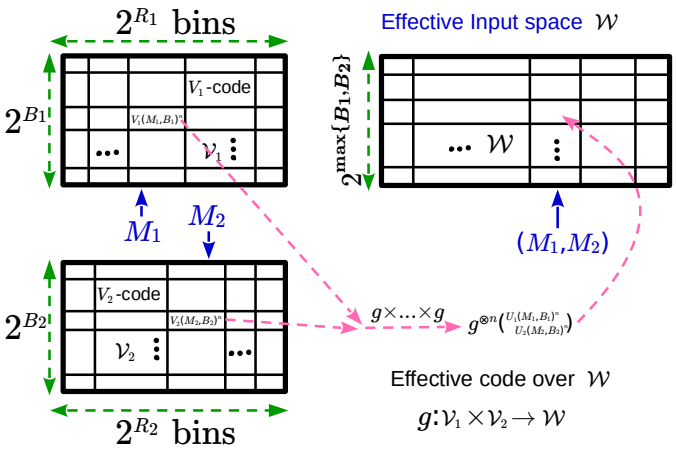

Fig. 9. Structured coding strategy of decoding into the sum codebook implies effective communication via the $U_1 \oplus U_2 = W-$codebook. Algebraic closure implies only $2^{n \max\{B_1,B_2\}}$ codewords allocated for each message pair.

$X_j = U_j \oplus S_j$ where $\oplus$ denotes addition mod-2. This implies the quantum state corresponding to which we compute our information quantities is

$$\sigma^{YS_1S_2X_1X_2U_1U_2} = \sum_{s_1,s_2} \frac{\tau(1-\tau)}{4} \begin{bmatrix} \mathbb{1}_{\{\substack{s_1 \oplus s_2 \\ =0}\}} |1\rangle\langle 1| + \\ \mathbb{1}_{\{\substack{s_1 \oplus s_2 \\ =1}\}} |v_\theta\rangle\langle v_\theta| \end{bmatrix} \otimes |s_1\ s_2\rangle\langle s_1\ s_2| \otimes \begin{bmatrix} |0\ 1\ s_1\ 1 \oplus s_2\rangle\langle 0\ 1\ s_1\ 1 \oplus s_2| + \\ |1\ 0\ 1 \oplus s_1\ s_2\rangle\langle 1\ 0\ 1 \oplus s_1\ s_2| \end{bmatrix}$$

$$+ \sum_{s_1,s_2} \left[\mathbb{1}_{\{\substack{s_1 \oplus s_2 \\ =0}\}} |0\rangle\langle 0| + \mathbb{1}_{\{\substack{s_1 \oplus s_2 \\ =1}\}} \left| v_\theta^\perp \right\rangle\left\langle v_\theta^\perp \right| \right] \otimes |s_1\ s_2\rangle\langle s_1\ s_2| \otimes \begin{bmatrix} \frac{(1-\tau)^2}{4} |0\ 0\ s_1\ s_2\rangle\langle 0\ 0\ s_1\ s_2| + \\ \frac{\tau^2}{4} |1\ 1\ 1 \oplus s_1\ 1 \oplus s_2\rangle\langle 1\ 1\ 1 \oplus s_1\ 1 \oplus s_2| \end{bmatrix}.$$

The bound on the sum rate achievable using IID random codes as stated in Thm. 1 is $I(U_1U_2;Y)_\sigma - I(U_1;S_1)_\sigma - I(U_2;S_2)_\sigma$. In Appendix C, we have provided characterization of the component quantum states with respect to which the above information quantities have to be computed. Referring to the same, it can be verified that $I(U_1U_2;Y)_\sigma - I(U_1;S_1)_\sigma - I(U_2;S_2)_\sigma = \alpha - 2 + 2h_b(\tau)$ where

$$\alpha = \tilde{h}_b((1-2\tau)^2 \sin\theta) - \tilde{h}_b(\sqrt{1 - 4\epsilon(1-\epsilon)\sin^2\theta}),\ \tilde{h}_b(x) \triangleq h_b\left(\frac{1}{2} + \frac{x}{2}\right) \text{ and } \epsilon = 2\tau(1-\tau). \tag{3}$$

It maybe verified that $\alpha = 1$ if $\theta = 0$ indicating the maximum sum rate achievable is a continuous function of θ as one expects. In Prop. 5, we verify that the linear coding scheme achieves any rate pair satisfying $R_1 + R_2 < uce\{\max\{0, \alpha - 1 + h_b(\tau)\}\}$ which strictly subsumes that achievable above.

C. Sieving Relevant Information via Algebraic Closure

The key difference between the structured and unstructured coding scheme is the decoding rule. While the former pins down the pair, the latter only decodes the sum, leaving uncertainty in the pair. Note that, the codeword $u_j^n(m_j, s_j^n)$ chosen by sender j contains, in addition to the message, information about s_j^n. By requiring the receiver to pin down the pair $(u_j^n(m_j, s_j^n) : j \in [2])$ of chosen codewords, the unstructured coding scheme is forcing the receiver to gather information of the state sequences that is not of value to it. Is there a function of $(u_j^n(m_j, s_j^n) : j \in [2])$ that, while containing information of the pair m_1, m_2 of messages can also suppress the amount of information of the pair s_1^n, s_2^n and can the coding scheme enable the Rx decode this function efficiently? The structured coding scheme is enabling the Rx do this via the mod-2 function. This key difference of decoding the sum of chosen codewords is coupled with the algebraic closure property of coset codes, thereby compressing the fan-outs of every message pair and enabling more efficient packing.

From our discussion thus far, we conclude the following. Embedding the desired information - the message pair $\underline{m}$ - in a specific compressive bivariate function - the mod-2 sum in the case of Ex. 1 - of the chosen codewords and building codes, which when operated through this function do not explode in the range of outcomes are the key reasons we are able to achieve strictly higher throughput. As we discuss in the next section, this phenomenon can be exploited in a broader class of CQMACs.

u_1	u_2	$u_1 \vee u_2$
0	0	0
0	1	1
1	0	1
1	1	1

Fig. 10. Logical OR

u_1	u_2	$u_1 \oplus_3 u_2$	u_1	u_2	$u_1 \oplus_3 u_2$
0	0	0	1	2	0
0	1	1	2	0	2
0	2	2	2	1	0
1	0	1	2	2	1
1	1	2			

Fig. 11. Mod-3 addition, i.e, field addition in $\mathcal{F}_3$

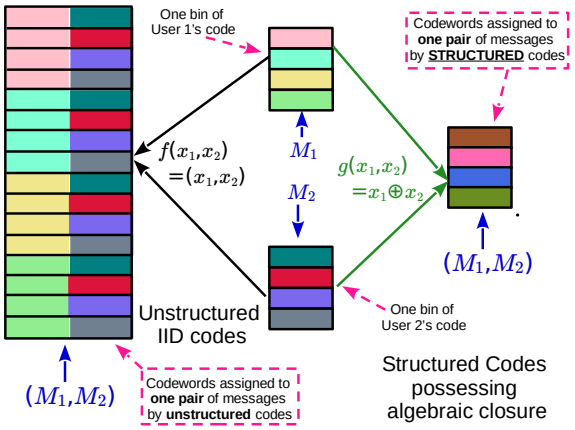

Fig. 12. The collections of four codewords each in the center depict one bin each of the two user's codes. The unstructured IID coding strategy of decoding into $U_1 \times U_2$codebook and disambiguate the pair is depicted on the left and the structured coding strategy of disambiguating only the mod-2 sum by attempting to decode into the sum codebook is depicted on the right.

D. A General Coding Strategy Exploiting the Broader Underlying Theme

The above discussions clearly illustrate the utility of algebraic closure in communication over a QMSTx. For both the commutative and non-commutative cases of Ex. 1, rates corresponding to uniform distributions on U_1, U_2 were sufficient to achieve higher sum rates using coset codes. For a general QMSTx, it is necessary to achieve rates corresponding to non-uniform distributions on $\mathcal{U}_1, \mathcal{U}_2$. Codewords of a random linear code are jointly related and the natural approach of picking its generator matrix uniformly at random will ensure that the codewords are uniformly distributed. How does one achieve rates corresponding to non-uniform distributions via linear codes? And are their examples of QMSTx for which such non-uniform distributions can enable communication at strictly larger rates via coset/linear codes? The rest of our article provides a definitive affirmative answer to this question. We design *union coset codes* in Sec. IV specifically aimed at achieving rates corresponding to non-uniform distributions and characterize a general inner bound in Thms. 2 and 4 that achieve rates corresponding to arbitrary distributions. Following this, we identify a non-commutative Ex. 2 in Sec. IV-E for which rates achievable via UCCs corresponding to a non-uniform distributions yield strictly larger rates than those achievable via unstructured IID random codes.

Are the above ideas restricted to finite field additions exploited via coset codes possessing algebraic closure property? Absolutely not. In fact, on the contrary these point to a richer classical and quantum Shannon theory. Consider $\Upsilon^{YU_1U_2} = \sum_{\underline{u}} p_{U_1U_2}(u_1,u_2)\delta_{u_1u_2} \otimes |u_1\ u_2\rangle\langle u_1\ u_2|$, where δ_{u_1,u_2} is as in Remark 1. Suppose there exists a finite set $\mathcal{W}$, a collection $(\gamma_w \in \mathcal{D}(\mathcal{H}_Y) : w \in \mathcal{W})$ and a map $f : \mathcal{U}_1 \times \mathcal{U}_2 \to \mathcal{W}$ such that $\delta_{u_1u_2} = \gamma_{f(u_1,u_2)}$, then we could extract the message pair (m_1, m_2) by decoding into the effective $\mathcal{W}-$codebook. We would then need to build codes on $\mathcal{U}_1^n, \mathcal{U}_2^n$ that are '$f-$closed'. In other words, when we compute the range of application of $f^{\otimes n}$ on all pairs of codewords in a pair of bins, we must be able to non-trivially contain the outcome range. See Fig. 9 and 12. This leads us to the code construction challenge of designing $f-$closed codes.

On the other hand, our current algebraically closed coset codes can serve a broader range of non-linear scenarios. Consider the logical OR truth table in Fig. 10 and suppose the induced collection $(\delta_{u_1,u_2} : (u_1,u_2) \in \{0,1\}^2)$ of density operators satisfies $\delta_{u_1,u_2} = \delta_{\tilde{u}_1\tilde{u}_2}$ whenever $u_1 \vee u_2 = \tilde{u}_1 \vee \tilde{u}_2$. Can we construct codes that are logical OR $\vee-$closed? If one views u_1, u_2 to live on $\mathcal{F}_3$ instead of $\{0,1\}$ by letting $p_{U_1}(2) = p_{U_2}(2) = 0$, then observe that one can recover $u_1 \vee u_2$ if one recovers $u_1 \oplus_3 u_2$. In other words, $(u_1,u_2) - u_1 \oplus_3 u_3 - u_1 \vee u_2$ is a Markov chain. If we can therefore build codes over $\mathcal{F}_3$ that are algebraically closed with respect to $\oplus_3$ and let the Rx recover ternary addition of the pair of chosen codewords, one could potentially outperform IID random coding strategy. We refer the reader to [23, Ex. 2] where this has been demonstrated albeit in a different communication scenario.

The above discussion and the gains we are able to glean at through simulations for non-linear examples such as Ex. 2 suggest that we design a broader general coding strategy based on algebraically closed coset codes. This motivates our findings in Sec. IV and V. Going beyond field addition, one can design coding strategies based on codes closed under other well structured operations such as group and ring additions to leverage rate gains. We refer the interested reader to [13] and [14] for pursuits along these paths.

IV. Inner Bound based on Union Coset Codes

Building on our discussions in the previous section and particularly Sec. III-D, we present our first main result - a new inner bound to the capacity of the QMSTx based on *union coset codes* and provide a proof of achievability. The coding scheme we propose to prove achievability is a generalization of the one we presented for Ex. 1. Ex. 1 with $\theta = 0$ and $\theta \in (0, \frac{\pi}{2})$ serve as commutative and non-commutative examples respectively, for which the inner bound in Thm. 2 is strictly larger than that proven in Thm. 1. In Sec. IV-E, we present a second non-commutative example for which the characterized inner bound in Thm. 2 is computationally verified to be strictly larger than that achievable via unstructured IID random codes (Thm. 1).

Theorem 2. *A rate-cost quadruple $(\underline{R}, \underline{\tau}) \in \mathcal{A}_c \subseteq [0,\infty)^4$ is achievable if there exists a finite field $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{W} = \mathcal{F}_q$ and conditional PMFs $p_{X_jV_j|S_j}$ on $\mathcal{X}_j \times \mathcal{V}_j$ for $j \in [2]$ with respect to which*

$$R_1 + R_2 < \min\{H(V_j|S_j)_\Upsilon : j \in [2]\} - H(V_1 \oplus V_2|Y)_\Upsilon = \min \left\{ \begin{matrix} I(U_j; Y, U_{\bar{j}})_\Upsilon - I(U_1 \oplus U_2; U_{\bar{j}}|Y)_\Upsilon \\ -I(U_j; S_j)_\Upsilon \end{matrix} : j \in [2] \right\} \quad (4)$$

where all mutual information quantities are computed with respect to the state

$$\Upsilon^{Y\underline{X}\underline{V}W\underline{S}} \triangleq \sum_{\underline{s},\underline{v},w,\underline{x}} p_{\underline{S}\underline{V}W\underline{X}}(\underline{s},\underline{v},w,\underline{x}) \rho_{\underline{x}\underline{s}} \otimes |\underline{x}\ \underline{v}\ w\ \underline{s}\rangle\langle \underline{x}\ \underline{v}\ w\ \underline{s}| \ \textit{where}$$

$$p_{\underline{S}\underline{V}W\underline{X}}(\underline{s},\underline{v},w,\underline{x}) = \mathrm{p}_{\underline{S}}(\underline{s}) \prod_{j=1}^{2} p_{X_jV_j|S_j}(x_j, v_j|s_j) \mathbb{1}_{\{w = v_1 \oplus v_2\}} \textit{ for all } (\underline{s},\underline{v},w,\underline{x}) \in \underline{\mathcal{S}} \times \underline{\mathcal{V}} \times \mathcal{W} \times \underline{\mathcal{X}}.$$

Proof. We begin by outlining our techniques and identifying the new elements. The main novelty is in the code structure we design and the decoding POVM we propose. In Sec. IV-A, we characterize a UCC and describe our codes. The Gel'fand-Pinsker encoding (Sec. IV-B) is employed by both senders. We decode only the sum codeword and hence employ a single user decoding POVM (Sec. IV-C). Since we decode into a UCC obtained by adding two statistically correlated UCCs, our analysis is not a standard one and detailed in Sec. IV-D.

A. Code Structure

The gain in rates for Ex. 1 crucially relied on the bins of both codes being coset shifts of a common linear code, thereby ensuring that the size of the sum of any pair of bins was contained. We observe that the shifts can be arbitrary and there are no structural requirement on the union of these cosets. We are thus led to a UCC.

Definition 3. *A UCC built over $\mathcal{F}_q$ is specified through a generator matrix $g \in \mathcal{F}_q^{k\times n}$ and a map $\iota : \mathcal{F}_q^l \to \mathcal{F}_q^n$ of coset shifts. The collection*

$$c(m) \triangleq \{v^n(a,m) = ag \oplus \iota(m) : a \in \mathcal{F}_q^k\}$$

forms the bin corresponding to message $m \in \mathcal{F}_q^l$ and the union $\cup_m c(m)$ of bins forms the code. We refer to this code of block-length n and rate $\frac{l}{n}$ as an (n,k,l,g,ι) UCC.

We employ UCCs as the codebook for both senders. The symmetry in Ex. 1 permitted us to design codes of the same rate for both senders. In general, to enable codes of different rates, we propose a 'nesting' of the two UCCs. Without loss of generality, assume the size of sender 1's bins is the smaller of the two. We equip user j with UCC $(n, k_j, l_j, g_j, \iota_j)$ and enforce $g_2 = \begin{bmatrix} g_1^T & g_{2/1}^T \end{bmatrix}^T$. See Fig. 13. This ensures that the bins of user 1's code are sub-cosets of the bins of user 2's code, thus guaranteeing the desirable property mentioned prior to Defn. 3. Let $\lambda_j \triangleq (v_j^n(a_j, m_j) \triangleq a_j g_j \oplus \iota_j(m_j) : (a_j, m_j) \in \mathcal{V}^{k_j} \times \mathcal{V}^{l_j})$ denote the codebook of sender j

B. Encoding

Our encoding is identical to that described for unstructured codes in Sec. III-A. See Fig. 4. On observing message $m_j \in [q^{l_j}]$ and state sequence s_j^n, sender j looks for a codeword in $c_j(m_j)$ that is jointly typical with s_j^n. If it finds at least one, one among these is chosen and denoted $v_j^n(m_j, s_j^n)$. If it finds none, $v_j^n(m_j, s_j^n)$ is set to a default codeword in $c_j(m_j)$. The pair $(s_j^n, v_j^n(m_j, s_j^n))$ is mapped to an input sequence via a 'fusion map' $f_j : \mathcal{S}_j^n \times \mathcal{V}_j^n \to \mathcal{X}_j^n$. For the sake of the ensuing analysis, we formalize this encoding with some notation.

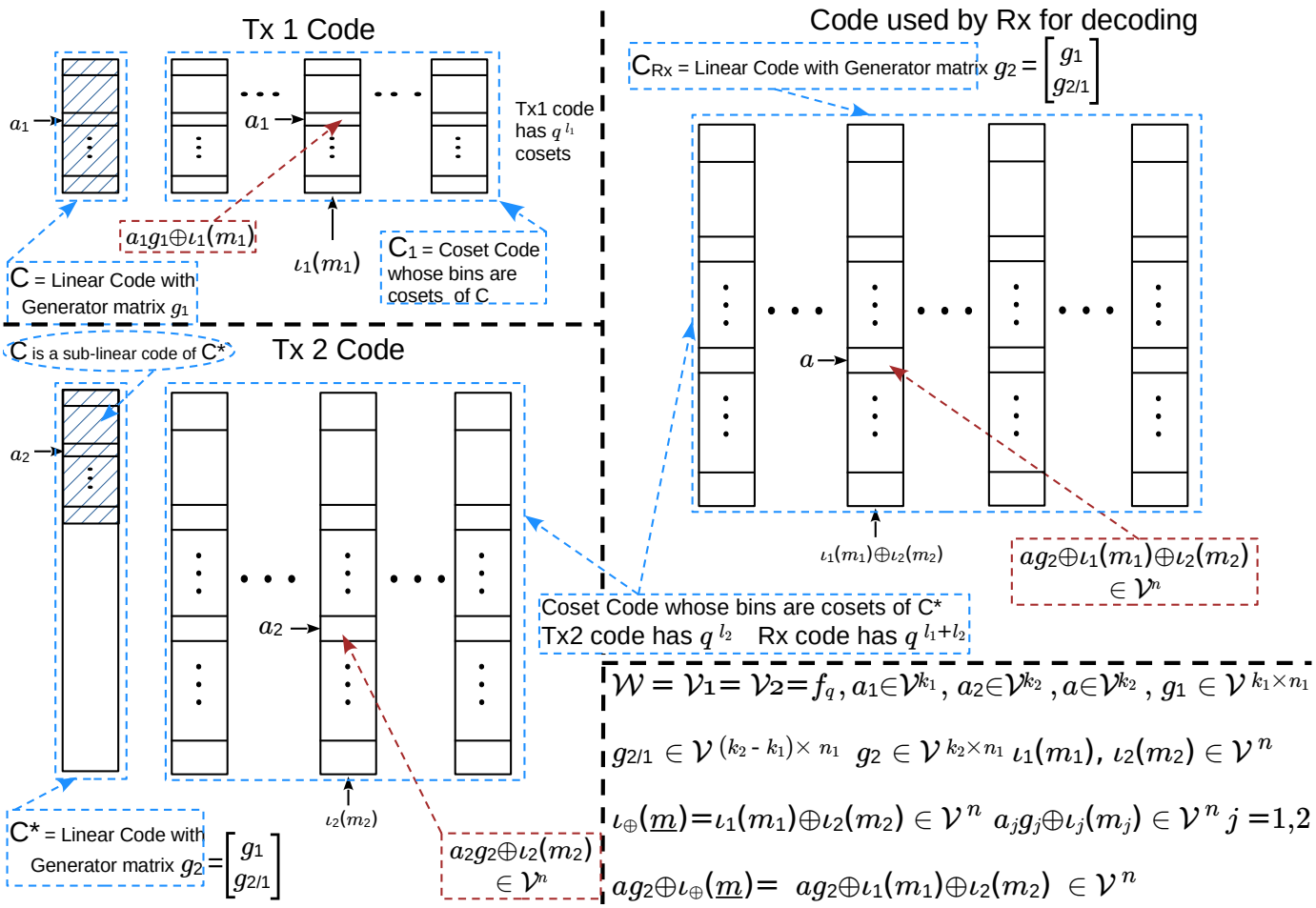

Fig. 13. The entire code structure comprising of two codes employed by the Txs and code employed by the Rx is depicted with legends for the associated terms.

Let $\alpha_j(m_j, s_j^n) \triangleq \sum_{a_j} \mathbb{1}_{\{(v_j^n(a_j, m_j), s_j^n) \in T_{\eta_3}(p_{S_j V_j})\}}$ be the number of available jointly typical codewords and let

$$\mathcal{L}_j(m_j, s_j^n) \triangleq \begin{cases} \{a_j : (v_j^n(a_j, m_j), s_j^n) \in T_{\eta_3}(p_{V_j S_j})\} & \text{if } \alpha_j(m_j, s_j^n) \geq 1 \\ \{0^{k_j}\} & \text{otherwise} \end{cases} \tag{5}$$

For every pair (m_j, s_j^n), $a_j(m_j, s_j^n)$ is an element chosen from $\mathcal{L}_j(m_j, s_j^n)$. We define $v_j^n(m_j, s_j^n) \triangleq v_j^n(a_j(m_j, s_j^n), m_j)$. A predefined 'fusion map' $f_j : \mathcal{S}_j^n \times \mathcal{V}_j^n \to \mathcal{X}_j^n$ is used to map the pair $s_j^n, v_j^n(m_j, s_j^n)$ to an input sequence in $\mathcal{X}_j^n$ henceforth denoted $x_j^n(m_j, s_j^n)$. A remark on our notation is in order. Our notation involves multiple objects referenced via $a_j(\cdot)$ and/or $\alpha_j(\cdot)$. This choice is motivated to ensure related objects have similar notation. We admit this causes some confusion/difficulty. To alleviate this, we have identified the corresponding rows in Table I with a double $*$ to direct the reader's attention to the same.

C. Decoding POVM

Consider the UCC $(n, k_2, l_1 + l_2, g_2, \iota_\oplus)$ depicted on the top right side of Fig. 13 where $\iota_\oplus(\underline{m}) = \iota_1(m_1) \oplus \iota_2(m_2)$ for $\underline{m} = (m_1, m_2) \in \mathcal{F}_q^{l_1} \times \mathcal{F}_q^{l_2}$. Let $w^n(a, \underline{m}) \triangleq ag_2 \oplus \iota_1(m_1) \oplus \iota_2(m_2)$ denote a generic codeword and let λ_R denote this UCC, i.e. the collection $\left(w^n(a, \underline{m}) : (a, m_1, m_2) \in [\mathcal{F}_q^{k_2}] \times [\mathcal{F}_q^{l_1}] \times [\mathcal{F}_q^{l_2}]\right)$. Suppose for each message pair $\underline{m}$, the collection $\left(w^n(a, \underline{m}) : a \in \mathcal{F}_q^{k_2}\right)$ is a distinct coset. In other words, suppose there is a 1:1 correspondence between $\{\underline{m} : \underline{m} \in [\mathcal{F}_q^{l_1}] \times [\mathcal{F}_q^{l_2}]\}$ and the collection of cosets of λ_R. Then, observe that when codewords from every distinct pair of cosets are added, the sum falls in a unique coset of λ_R. If the Rx correctly identifies in which coset of this UCC, lies the sum of the two codewords chosen by the Txs, then it can recover the pair of messages. This motivates our decoding POVM. Since we only need to identify the coset in which the sum of the chosen codewords lie, we let $\pi_{p_W, \eta_2}^{\sigma_{w^n}(a, \underline{m})}$ be the η_2-conditional typical projector of $\otimes_{t=1}^n \sigma_{w_t(a, \underline{m})}$ with respect to the PMF p_W, where $\sigma_w \triangleq \sum_{\underline{x}, \underline{s}} p_{\underline{X}\underline{S}|W}(\underline{x}, \underline{s}|w) \rho_{\underline{x}\underline{s}} : w \in \mathcal{W}$, $p_{\underline{S}\underline{X}W}$ is as defined in the Thm. statement. As stated in Appendix A-B,

Symbol	Description	Comment
$\frac{k_j}{n}\log q$ and $\frac{l_j}{n}\log q$	User j's binning rate and information rate respectively.	WLOG assume $k_2 > k_1$.
$g_1 \in \mathcal{F}_q^{k_1\times n}, g_2 \in \mathcal{F}_q^{k_2\times n}$ $g_2 = [g_1^T \; g_{2/1}^T]^T$	g_1, g_2 are the generator matrices of user 1,2's UCCs. We have assumed WLOG that $k_2 > k_1$.	The generator matrices are nested to ensure containment of the sum codewords when bins of users 1 and 2 are added..
$\iota_j : \mathcal{F}_q^{l_j} \to \mathcal{F}_q^n$	For message $m_j \in \mathcal{F}_q^{l_j}$, $\iota_j(m_j) \in \mathcal{F}_q^n$ is the dither (bias) vector of the coset corresponding to message m_j.	Recall that messages indexes bins and each bin of a user's code are (random) coset shifts of a common linear code.
$v_j^n(a_j, m_j) \triangleq a_j g_j \oplus \iota_j(m_j)$	A generic codeword in bin/coset indexed by message m_j	A generic codeword in bin/coset $c_j(m_j)$ defined next.
$c_j(m_j)$	$\{v_j^n(a_j, m_j) : a_j \in \mathcal{F}_q^{k_j}\}$	Bin/Coset corresponding to message m_j
** $\alpha_j(m_j, s_j^n)$ **	# of m_j-bin codewords that are jointly typical with s_j^n	
$\mathcal{L}_j(m_j, s_j^n)$	Indexes of codewords in $c_j(m_j)$ jointly typical with s_j^n	If this list is empty, then $\mathcal{L}_j(m_j, s_j^n) \triangleq \{0^{k_j}\}$
** $a_j(m_j, s_j^n)$ **	Index chosen from $\mathcal{L}_j(m_j, s_j^n)$ to communicate message m_j when state sequence is s_j^n	
$v_j^n(m_j, s_j^n)$	$v_j^n(a_j(m_j, s_j^n), m_j)$	V_j-codeword chosen to communicate message m_j when state sequence is s_j^n
$f_j : \mathcal{S}_j^n \times \mathcal{V}_j^n \to \mathcal{X}_j^n$	'Fusion map'	Maps chosen V_j-codeword and observed state sequence to X_j-codeword
$x_j^n(m_j, s_j^n)$	$f_j(v_j^n(m_j, s_j^n), s_j^n) = f_j(v_j^n(m_j, s_j^n), m_j)$	X_j-codeword chosen to communicate message m_j when state sequence is s_j^n
$p_{\underline{S}XVW} \triangleq p_{S_1S_1X_2X_2V_1V_2W}$	Chosen 'test channel' satisfies $p_{\underline{S}} = \mathrm{p}_{\underline{S}}$ and $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{W} = \mathcal{F}_q$ is finite field of size q	Additionally $p_{\underline{S}XVW}$ satisfies $p_{\underline{S}XVW} = p_{\underline{S}XV}\mathbb{1}_{W=V_1\oplus V_2}$
$\sigma_w : w \in \mathcal{W}$	$\sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}\vert W}(\underline{x},\underline{s}\vert w)\rho_{\underline{x}\underline{s}} : w \in \mathcal{W}$	
μ	$\sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}}(\underline{x},\underline{s})\rho_{\underline{x}\underline{s}}$	
$(\lambda_{\underline{m}} : \underline{m} \in [\underline{\mathcal{M}}])$	Decoding POVM as defined in (32)	
(a) $\gamma_{a,\underline{m}}$, (b) $\pi_{\eta_1}^{\mu}$,	(a) $\pi_{\eta_1}^{\mu}\pi_{a,\underline{m}}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu}$, (b) η_1-typical projector of the state μ	
$\pi_{a,\underline{m}}^{\sigma,\eta_2} = \pi_{p_W,\eta_2}^{\sigma_{w^n(a,\underline{m})}}$	η_2-conditional typical projector of $\sigma_{w^n(a,\underline{m})}$ with respect to p_W	

TABLE I

DESCRIPTION OF ELEMENTS THAT CONSTITUTE THE CODING SCHEME

we henceforth let $\pi_{a,\underline{m}}^{\sigma,\eta_2} = \pi_{p_W,\eta_2}^{\sigma_{w^n(a,\underline{m})}}$. We elaborate for clarity. Recall our chosen state $\Upsilon^{Y\underline{X}W\underline{S}} = \mathrm{tr}_{\underline{V}}\{\Upsilon^{Y\underline{X}\underline{V}W\underline{S}}\}$ and let

$$\Upsilon^{Y\underline{X}W\underline{S}} = \sum_{w\in\mathcal{W}} p_W(w)\sigma_w \otimes |w\rangle\langle w|\,, \text{ where } \sigma_w = \sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}\vert W}(\underline{x},\underline{s}\vert w)\rho_{\underline{x}\underline{s}} \text{ has SCD } \sigma_w = \sum_{y\in\mathcal{Y}} r_{Y\vert W}(y\vert w)\,\big|h_{y|w}\big\rangle\big\langle h_{y|w}\big|.$$

$$\text{We then have } \pi_{a,\underline{m}}^{\sigma,\eta_2} = \pi_{p_W,\eta_2}^{\sigma_{w^n(a,\underline{m})}} = \sum_{y^n\in\mathcal{Y}^n} \bigotimes_{t=1}^{n} \big|h_{y_t|w_t(a,\underline{m})}\big\rangle\big\langle h_{y_t|w_t(a,\underline{m})}\big|\,\mathbb{1}_{\{(w^n(a,\underline{m}),y^n)\in T_\eta^n(p_W\otimes r_{Y|W})\}}$$

where $p_W \otimes r_{Y|W}$ is the joint PMF with marginal p_W and conditional PMF $r_{Y|W}$. We define $\gamma_{a,\underline{m}} \triangleq \pi_{\eta_1}^{\mu}\pi_{a,\underline{m}}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu}$ where $\pi_{\eta_1}^{\mu}$ is the η_1-typical projector of the state $\mu \triangleq \sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}}(\underline{x},\underline{s})\rho_{\underline{x}\underline{s}}$. The decoding POVM is

$$\lambda_{\underline{m}} \triangleq \left(\sum_{\hat{a},\hat{m}_1,\hat{m}_2}\gamma_{\hat{a},\hat{m}_1,\hat{m}_2}\right)^{-\frac{1}{2}} \sum_{a}\gamma_{a,\underline{m}}\left(\sum_{\hat{a},\hat{m}_1,\hat{m}_2}\gamma_{\hat{a},\hat{m}_1,\hat{m}_2}\right)^{-\frac{1}{2}} \text{ and } \lambda_{-1} \triangleq \boldsymbol{I} - \sum_{\underline{m}}\lambda_{\underline{m}}. \tag{6}$$

Clearly, the decoding POVM has $q^{l_1+l_2}+1$ outcomes $\{-1, \underline{m} \in [q^{l_1}]\times[q^{l_2}]\}$. The Rx declares error if outcome -1 is observed and declares $\underline{\hat{m}}_1, \underline{\hat{m}}_2$ if outcome $\underline{\hat{m}} = (\underline{\hat{m}}_1, \underline{\hat{m}}_2) \in [q^{l_1}]\times[q^{l_2}]$ is observed.

D. Probability of Error Analysis

We employ the random coding technique to prove the existence of a code with the promised rates for which the error probability falls to 0 exponentially in the block-length n. Towards that end, observe that our code and the coding scheme is completely characterized via the following objects : the generator matrices $g_1, g_{2/1}$, the collection $(\iota_j(m_j) : m_j \in [q^{l_j}])$ of dither/bias vectors specifying the coset shifts, the indices $(a_j(m_j, s_j^n) : (m_j, s_j^n) \in [q^{l_j}]\times\mathcal{S}_j^n$, and the final codeword choices $(x_j^n(m_j, s_j^n) : (m_j, s_j^n) \in [q^{l_j}]\times\mathcal{S}_j^n)$. Our first step is to characterize the error probability for a generic choice of these objects. In particular, we characterize an upper bound on this error probability composed of multiple terms. Our second step is to specify a probability distribution on the collection of codes by specifying a distribution on the aforementioned objects. In our third step, we prove that the expectation of each of the above mentioned terms falls to 0 exponentially if the rates of the code satisfy (4).

An upper bound on the error probability for a generic code : For a generic choice of the aforementioned objects, the average error probability is

$$\xi(\underline{e},\lambda)=\sum_{\underline{m}}\frac{\hat{\zeta}(\underline{m})}{|\underline{\mathcal{M}}|}\text{ where }\hat{\zeta}(\underline{m})\triangleq\sum_{\underline{s}^n}\mathtt{p}^n_{\underline{S}}(\underline{s}^n)\hat{\zeta}(\underline{m}|\underline{s}^n),\rho_{\underline{m},\underline{s}^n}\triangleq\bigotimes_{t=1}^{n}\rho_{x_1(m_1,s_1^n)_t x_2(m_2,s_2^n)_t \underline{s}_t}\tag{7}$$

$\hat{\zeta}(\underline{m}|s^n)\triangleq\text{tr}\{(\boldsymbol{I}-\lambda_{\underline{m}})\rho_{\underline{m},\underline{s}^n}\}$, $\boldsymbol{I}=I^{\otimes n}$, $\mathcal{M}_j=[q^{l_j}]$ and hence $|\underline{\mathcal{M}}|=q^{l_1+l_2}$. We consider an arbitrary pair $\underline{m}=(m_1,m_2)$ and henceforth focus our study on $\hat{\zeta}(\underline{m})$. Throughout the rest of our study of $\hat{\zeta}(\underline{m})$, we let $a_\oplus\triangleq a_1(m_1,s_1^n)\;0^{k_2-k_1}\oplus a_2(m_2,s_2^n)$. With this definition and (32), note that

$$\lambda_{\underline{m}}\geq(S+T)^{-\frac{1}{2}}\,S\,(S+T)^{-\frac{1}{2}}\text{ where }S=\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{a_\oplus,\underline{m}}\pi^{\mu}_{\eta_1}\text{ and }T=\sum_{\hat{a}\neq a_\oplus}\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{\hat{a},\underline{m}}\pi^{\mu}_{\eta_1}+\sum_{a\in\mathcal{F}_q^{k_2}}\sum_{\underline{\hat{m}}\neq\underline{m}}\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{a,\underline{\hat{m}}}\pi^{\mu}_{\eta_1}.\tag{8}$$

$$\text{and hence, }\hat{\zeta}(\underline{m}|s^n)\leq\zeta(\underline{m}|\underline{s}^n),\text{ where }\zeta(\underline{m}|\underline{s}^n)\triangleq\text{tr}\left(\left[\boldsymbol{I}-(S+T)^{-\frac{1}{2}}\,S\,(S+T)^{-\frac{1}{2}}\right]\rho_{\underline{m},\underline{s}^n}\right).\tag{9}$$

We shall henceforth focus our study on $\zeta(\underline{m})\triangleq\sum_{\underline{s}^n}\mathtt{p}^n_{\underline{S}}(\underline{s}^n)\zeta(\underline{m}|\underline{s}^n)$ which serves as an upper bound on $\hat{\zeta}(\underline{m})$ in (7). Towards that end, we split the event corresponding to $\zeta(\underline{m})$ into two parts - $\mathcal{E}$ and $\overline{\mathcal{E}}$ - and analyze the event corresponding to the two parts separately. Towards defining $\mathcal{E}$, let

$$\mathcal{E}_{j1}\triangleq\left\{\underline{s}^n\in T_{\frac{\eta_3}{2}}(\mathtt{p}_{\underline{S}})\right\},\mathcal{E}_{j2}\triangleq\left\{\begin{array}{c}|\mathcal{L}_j(m_j,s_j^n)|\geq L_j,\\(s_j^n,v_j^n(m_j,s_j^n))\in T_{\eta_3}(p_{S_jV_j})\end{array}\right\},\mathcal{E}_{121}\triangleq\left\{\begin{pmatrix}s_j^n,v_j^n(m_j,s_j^n)\\:j\in[2]\end{pmatrix}\in T_{2\eta_3}(p_{\underline{SV}})\right\}\tag{10}$$

$\mathcal{E}_{122}\triangleq\left\{\left(s_j^n,v_j^n(m_j,s_j^n),x_j^n(m_j,s_j^n):j\in[2],w^n(a_\oplus,\underline{m})\right)\in T_{4\eta_3}(p_{\underline{SV}X\underline{W}})\right\}$ and finally $\mathcal{E}\triangleq\bigcap_{j=1}^{2}\mathcal{E}_{j1}\cap\mathcal{E}_{j2}\bigcap\mathcal{E}_{121}\cap\mathcal{E}_{122}$,

where $L_j\triangleq\frac{1}{2}\exp\{k_j\log q-n\log q+nH(V_j|S_j)\Upsilon-3n\eta_3\}$ is a threshold chosen to ensure that there is at least one jointly typical sequence. Instead of choosing $L_j=1$, choosing it as above aids our error analysis. This is evident in Appendix J where we derive an upper bound on $\overline{\xi_3}(\underline{m})$ found in (40). We remark that all Von Neumann entropies in this proof are evaluated with respect to the joint state $\Upsilon^{Y\underline{SV}W\underline{X}}$ specified in (5). Since

$$\overline{\mathcal{E}}=\bigcup_{j=1}^{2}\overline{\mathcal{E}}_{j1}\cup\overline{\mathcal{E}}_{j2}\bigcup\overline{\mathcal{E}}_{121}\cup\overline{\mathcal{E}}_{122}=\bigcup_{j=1}^{2}\mathcal{F}_{j1}\cup\mathcal{F}_{j2}\bigcup\mathcal{F}_{121}\cup\mathcal{F}_{122}\text{ where }\mathcal{F}_{j1}\triangleq\overline{\mathcal{E}}_{j1},\mathcal{F}_{j2}\triangleq\mathcal{E}_{j1}\cap\overline{\mathcal{E}}_{j2},\tag{11}$$

$$\mathcal{F}_{121}\triangleq\bigcap_{j=1}^{2}\mathcal{E}_{j1}\cap\mathcal{E}_{j2}\cap\overline{\mathcal{E}}_{121},\mathcal{F}_{122}\triangleq\bigcap_{j=1}^{2}\mathcal{E}_{j1}\cap\mathcal{E}_{j2}\cap\mathcal{E}_{121}\cap\overline{\mathcal{E}}_{122}\text{ we have }1=\mathbb{1}_{\overline{\mathcal{E}}}+\mathbb{1}_{\mathcal{E}}\leq\sum_{j=1}^{2}\sum_{i=1}^{2}\left(\mathbb{1}_{\mathcal{F}_{ji}}+\frac{\mathbb{1}_{\mathcal{F}_{12i}}}{2}\right)+\mathbb{1}_{\mathcal{E}}\tag{12}$$

With these definitions, we have

$$\zeta(\underline{m})\leq\sum_{j=1}^{2}\zeta_j(\underline{m})+\tilde{\zeta}_2(\underline{m}),\text{ where }\zeta_j(\underline{m})\triangleq\sum_{\underline{s}^n}\mathtt{p}^n_{\underline{S}}(\underline{s}^n)\zeta(\underline{m}|\underline{s}^n)\sum_{i=1}^{2}\left(\mathbb{1}_{\mathcal{F}_{ji}}+\frac{\mathbb{1}_{\mathcal{F}_{12i}}}{2}\right)\text{ and }\tilde{\zeta}_2(\underline{m})\triangleq\sum_{\underline{s}^n}\mathtt{p}^n_{\underline{S}}(\underline{s}^n)\zeta(\underline{m}|\underline{s}^n)\mathbb{1}_{\mathcal{E}}.\tag{13}$$

Next, in regards to $\tilde{\zeta}_2(\underline{m})$, consider $\zeta(\underline{m}|s^n)$ defined in (9). Note that since $\pi^{\sigma,\eta_2}_{a,\underline{m}}\geq 0$ for every $a\in\mathcal{F}_q^{k_2}$, we have $S\geq 0$, $T\geq 0$ are PSD. Moreover $S=\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{a_\oplus,\underline{m}}\pi^{\mu}_{\eta_1}\leq\pi^{\mu}_{\eta_1}\boldsymbol{I}\pi^{\mu}_{\eta_1}=\pi^{\mu}_{\eta_1}\leq\boldsymbol{I}$ implying $\boldsymbol{I}-S$ is PSD. From the Hayashi Nagaoka inequality [39], we have

$$\tilde{\zeta}_2(\underline{m})\leq\zeta_3(\underline{m})+\zeta_4(\underline{m})+\zeta_5(\underline{m})\text{ where }\zeta_3(\underline{m})\triangleq 2\sum_{\underline{s}^n}\mathtt{p}^n_{\underline{S}}(\underline{s}^n)\,\text{tr}\left\{\left[\boldsymbol{I}-\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{a_\oplus,\underline{m}}\pi^{\mu}_{\eta_1}\right]\rho_{\underline{m},\underline{s}^n}\right\}\mathbb{1}_{\mathcal{E}},\tag{14}$$

$$\zeta_4(\underline{m})\triangleq 4\sum_{\underline{s}^n,\hat{a}\neq a_\oplus}\mathtt{p}^n_{\underline{S}}(\underline{s}^n)\,\text{tr}\left\{\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{\hat{a},\underline{m}}\pi^{\mu}_{\eta_1}\rho_{\underline{m},\underline{s}^n}\right\}\mathbb{1}_{\mathcal{E}}\text{ and }\zeta_5(\underline{m})\triangleq 4\sum_{\underline{s}^n,\hat{a}\in\mathcal{F}_q^{k_2}}\sum_{\underline{\hat{m}}\neq\underline{m}}\mathtt{p}_{\underline{s}^n}(\underline{s}^n)\,\text{tr}\left\{\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{\hat{a},\underline{\hat{m}}}\pi^{\mu}_{\eta_1}\rho_{\underline{m},\underline{s}^n}\right\}\mathbb{1}_{\mathcal{E}}\tag{15}$$

Collating through (13), (14), (15), we have $\zeta(\underline{m})\leq\sum_{i=1}^{5}\zeta_i(\underline{m})$ where the terms in the latter sum are defined through (13), (14) and (15). We now employ the random coding technique and prove that the average of these terms, evaluated over the ensemble of codes, falls exponentially to 0 if the rate conditions stated in the theorem hold. Towards that end, we now specify the distribution on the ensemble of codes.

Distribution of the random code : We now specify the probability distribution of the random code with respect to which we compute the expectation of the five terms mentioned above. Recall that our codes and the coding scheme are completely specified via the objects : $g_1 \in \mathcal{F}_q^{k_1\times n}$, $g_{2/1} \in \mathcal{F}_q^{(k_2-k_1)\times n}$, $\iota_j(m_j) \in \mathcal{F}_q^n : m_j \in [\mathcal{F}_q^{l_j}]$, $a_j(m_j, s_j^n) \in \mathcal{F}_q^{l_j} : m_j \in [\mathcal{F}_q^{l_j}], s_j^n \in \mathcal{S}_j^n$ and the collection of final codewords $x_j(m_j, s_j^n) : m_j \in [\mathcal{F}_q^{l_j}], s_j^n \in \mathcal{S}_j^n$. It therefore suffices to specify a joint distribution of these objects. The generator matrices $G_1 \in \mathcal{F}_q^{k_1\times n}, G_{2/1} \in \mathcal{F}_q^{(k_2-k_1)\times n}$, and the collection $(\iota_j(m_j) \in \mathcal{F}_q^n : m_j \in [q^{l_j}])$ of dither/bias vectors specifying the coset shifts are mutually independent and uniformly distributed on the respective range spaces. Given $G_1, G_{2/1}$ and the collection $(\iota_j(m_j) \in \mathcal{F}_q^n : m_j \in [q^{l_j}])$, each of $(A_j(m_j, s_j^n) : (m_j, s_j^n) \in [q^{l_j}]\times\mathcal{S}_j^n)$ are mutually independent and uniformly distributed in $\mathcal{L}_j(m_j, s_j^n)$ as defined in (5). Let $V_j^n(m_j, s_j^n) \triangleq V_j^n(A_j(m_j, s_j^n), m_j^n)$ as defined earlier for a generic code. Next, given $G_1, G_{2/1}$, the collections $(\iota_j(m_j) \in \mathcal{F}_q^n : m_j \in [q^{l_j}])$, $(A_j(m_j, s_j^n) : (m_j, s_j^n) \in [q^{l_j}] \times \mathcal{S}_j^n)$ and the event $V_j^n(m_j, s_j^n) = v_j^n(m_j, s_j^n) = (v_j(m_j, s_j^n)_1, \cdots, v_j(m_j, s_j^n)_n) : m_j \in [q^{l_j}], s_j^n \in \mathcal{S}_j^n$, the final codewords $(X_j^n(m_j, s_j^n) : m_j \in [q^{l_j}], s_j^n \in \mathcal{S}_j^n)$ are mutually independent and the probability that the final codeword $X_j^n(m_j, s_j^n) = x_j^n(m_j, s_j^n) = (x_j(m_j, s_j)_1, \cdots x_j(m_j, s_j)_n)$ is $\prod_{t=1}^n p_{X_j|V_jS_j}(x_j(m_j, s_j)_t|v_j(m_j, s_j^n)_t, s_{jt})$. Mathematically stated,

$$P\begin{pmatrix} G_1 = g_1, G_{2/1} = g_{2/1}, \iota_j(m_j) = d_j^n(m_j) : m_j \in [\mathcal{M}_j], \\ A_j(m_j, s_j^n) = a_j(m_j, s_j^n) : (m_j, s_j^n) \in [\mathcal{M}_j] \times \mathcal{S}_j^n \\ V_j(A_j(m_j, s_j^n), s_j^n) = v_j^n(m_j, s_j^n) : (m_j, s_j^n) \in [\mathcal{M}_j] \times \mathcal{S}_j^n \\ X_j(m_j, s_j^n) = x_j^n(m_j, s_j^n) : (m_j, s_j^n) \in [\mathcal{M}_j] \times \mathcal{S}_j^n, j \in [2] \end{pmatrix} = \frac{1}{q^{k_1 l}} \frac{1}{q^{(k_2-k_1)l}} \left(\frac{1}{q^n}\right)^{q^{l_1}} \left(\frac{1}{q^n}\right)^{q^{l_2}} \times$$

$$\left[\prod_{j=1}^{2} \prod_{s_j^n} \prod_{m_j} \frac{\mathbb{1}_{\{a_j(m_j,s_j^n)g_j \oplus d_j^n(m_j) = v_j^n(a_j(m_j,s_j^n))\}}}{|\mathcal{L}_j(m_j, s_j^n)|}\right] \left[\prod_{j=1}^{2} \prod_{s_j^n} \prod_{m_j} \prod_{t=1}^{n} p_{X_j|V_jS_j}(x_j(m_j, s_j^n)_t|v_j(m_j, s_j^n)_t, s_{jt})\right], \quad (16)$$

where $\mathcal{L}_j(m_j, s_j^n)$ is as defined in (5). We make the following remarks for ease of reference at a later point.

Remark 2. *Given $G_1, G_{2/1}$ and the collection $\iota_j(m_j) : m_j \in [q^{l_j}]$, $A_j(m_j, s_j^n)$ is uniformly distributed in $\mathcal{L}_j(m_j, s_j^n)$. Specifically, $A_j(m_j, s_j^n)$ is conditionally independent of $G_2, G_{2/1}$ and the collection $\iota_j(m_j) : m_j \in [q^{l_j}]$ given the lists $\mathcal{L}_j(m_j, s_j^n)$. As a consequence, $A_j(m_j, s_j^n)$ is conditionally independent of $V_j^n(a_j, m_j)$ for any a_j given $\mathcal{L}_j(m_j, s_j^n)$, and moreover, $A_j(m_j, s_j^n)$ is uniformly distributed in the latter list.*

In the rest of our proof, we derive upper bounds on $\overline{\zeta}_j(\underline{m}) \triangleq \mathbb{E}\{\zeta_j(\underline{m})\}$ for $j \in [5]$ that decay exponentially to 0, where the expectations in question are with respect to the distribution of the random code.

Upper bound on $\overline{\zeta}_1(\underline{m}), \overline{\zeta}_2(\underline{m})$: The analysis of both these terms is identical and we let $j \in [2]$ denote a generic term $\overline{\zeta}_j(\underline{m})$. From (13), (9) and the definition of S in (8), we note that $S \geq 0$ is PSD and hence $\boldsymbol{I} - (S+T)^{-\frac{1}{2}} S (S+T)^{-\frac{1}{2}} \leq \boldsymbol{I}$ implying $\zeta(\underline{m}|\underline{s}^n) \leq \mathrm{tr}\big(\boldsymbol{I} \cdot \rho_{\underline{m},\underline{s}^n}\big)$. Substituting this in the definition of $\zeta_j(\underline{m})$ in (13), we obtain $\zeta_j(\underline{m}) \leq \sum_{\underline{s}^n} \mathrm{p}_{\underline{S}}(\underline{s}^n) \sum_{i=1}^2 \left(\mathbb{1}_{\mathcal{F}_{ji}} + \frac{\mathbb{1}_{\mathcal{F}_{12i}}}{2}\right)$. This involves only classical probabilities and our study of $\overline{\zeta}_j(\underline{m})$ will therefore closely mimic [13, Upper Bound on ϵ_{2j} in Appendix B]. The proof of the following propositions are provided in Appendix E.

Proposition 1. *If $\frac{k_j \log q}{n} > \log q - H(V_j|S_j)_\Upsilon + 3\eta_3$ for $j \in [2]$, then there exists a strictly positive $\kappa > 0$ such that for all n sufficiently large $\overline{\zeta}_1(\underline{m}) + \overline{\zeta}_2(\underline{m}) \leq \exp\{-n\kappa\eta_3^2\}$*

To comprehend the above bound, note that codewords of a random UCC are uniformly distributed. In a coset with q^k random codewords that are uniformly distributed (Lemma 5), the expected number of codewords that will be jointly typical with the observed typical state sequence s_j^n is $|T_{\eta_3}(V_j|s_j^n)|q^{k-n}$ whose exponent is lower bounded by $k\log q - n\log q + nH(V_j|S_j)_\Upsilon - 4n\eta$. See Lemma 5. The condition in Proposition 1 guarantees the latter exponent is positive. This implies that $|\mathcal{L}_j(m_j, s_j^n)|$ will concentrate around $\exp\{k_j \log q - n\log q + nH(V_j|S_j)_\Upsilon\} = 2L_j$ and hence the probability that $|\mathcal{L}_j(m_j, s_j^n)| < L_j$ falls exponentially in n.

Upper bound on $\overline{\zeta}_3(\underline{m})$: In our analysis steps, we have adopted the convention that, if unspecified, the summation is over the entire range of the summand.[1] With this convention, we have

$$\zeta_3(\underline{m}) = 2\sum_{\underline{s}^n} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left\{\rho_{\underline{m},\underline{s}^n} - \pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{a_\oplus,\underline{m}}\pi^{\mu}_{\eta_1}\rho_{\underline{m},\underline{s}^n}\right\}\mathbb{1}_{\mathcal{E}} = 2\sum_{\underline{s}^n} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left\{\rho_{\underline{m},\underline{s}^n} - \pi^{\sigma,\eta_2}_{a_\oplus,\underline{m}}\left[\pi^{\mu}_{\eta_1}\rho_{\underline{m},\underline{s}^n}\pi^{\mu}_{\eta_1}\right]\right\}\mathbb{1}_{\mathcal{E}} \quad (17)$$

$$\leq 2\sum_{\underline{s}^n} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left\{\rho_{\underline{m},\underline{s}^n} - \pi^{\sigma,\eta_2}_{a_\oplus,\underline{m}}\rho_{\underline{m},\underline{s}^n}\right\}\mathbb{1}_{\mathcal{E}} + \sum_{\underline{s}^n} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\left\|\rho_{\underline{m},\underline{s}^n} - \pi^{\mu}_{\eta_1}\rho_{\underline{m},\underline{s}^n}\pi^{\mu}_{\eta_1}\right\|_1\mathbb{1}_{\mathcal{E}} \leq \zeta_{31}(\underline{m}) + \zeta_{32}(\underline{m}), \quad (18)$$

where $\zeta_{31}(\underline{m}) \triangleq 2\sum_{\underline{s}^n} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left\{\left[\boldsymbol{I} - \pi^{\sigma,\eta_2}_{a_\oplus,\underline{m}}\right]\rho_{\underline{m},\underline{s}^n}\right\}\mathbb{1}_{\mathcal{E}}$ and $\zeta_{32}(\underline{m}) \triangleq 2\sum_{\underline{s}^n} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\sqrt{\mathrm{tr}\left\{\left[\boldsymbol{I} - \pi^{\mu}_{\eta_1}\right]\rho_{\underline{m},\underline{s}^n}\right\}}\mathbb{1}_{\mathcal{E}}$. (19)

In the above, (17) follows from cyclicity of trace, (18) follows from 'measurement on close states' [35, Exercise 9.1.8] and (19) follows from [35, Chain of Inequalities 9.205 through to 9.209]. As an informed reader might have guessed, our analysis of $\zeta_{32}(\underline{m})$ is via an analysis analogous to the pinching lemma [35, Property 15.2.7]. In Appendix F, we have detailed the steps where we have proved that if $\eta_1 > 4\eta_3$, for all $n \in \mathbb{N}$ sufficiently large, we have

$$\zeta_{32}(\underline{m}) \leq 2|\mathcal{Y}||\underline{\mathcal{X}}||\underline{\mathcal{S}}|\exp\left\{-n(\eta_1 - 4\eta_3)^2\delta(r_{Y\underline{X}\underline{S}}, 4\eta_3, |\mathcal{Y}||\underline{\mathcal{X}}||\underline{\mathcal{S}}|)\right\}.$$

This leaves us with $\overline{\zeta}_3(\underline{m}) \leq \overline{\zeta}_{31}(\underline{m}) + 2|\mathcal{Y}||\underline{\mathcal{X}}||\underline{\mathcal{S}}|\exp\left\{-n(\eta_1 - 4\eta_3)^2\delta(r_{Y\underline{X}\underline{S}}, 4\eta_3, |\mathcal{Y}||\underline{\mathcal{X}}||\underline{\mathcal{S}}|)\right\}$. Towards analyzing $\overline{\zeta}_{31}(\underline{m})$, note that

$$\zeta_{31}(\underline{m}) \leq 2\sum_{\substack{\underline{s}^n,\underline{v}^n\\ w^n,\underline{x}^n}}\sum_{\substack{a_1\in\mathcal{F}_q^{k_1}\\ a_2,a\in\mathcal{F}_q^{k_2}}} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\mathrm{tr}\left\{\left[\boldsymbol{I} - \pi^{\sigma,\eta_2}_{a_\oplus,\underline{m}}\right]\rho_{\underline{m},\underline{s}^n}\right\}\mathbb{1}_{E_1\cap E_2\cap F_3} = 2\sum_{\substack{\underline{s}^n,\underline{v}^n\\ w^n,\underline{x}^n}}\sum_{\substack{a_1\in\mathcal{F}_q^{k_1}\\ a_2,a\in\mathcal{F}_q^{k_2}}} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\mathrm{tr}\left\{\left[\boldsymbol{I} - \pi^{\sigma,\eta_2}_{a,\underline{m}}\right]\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathbb{1}_{E_1\cap E_2\cap F_3}$$

$$= 2\sum_{\substack{\underline{s}^n,\underline{v}^n\\ w^n,\underline{x}^n}}\sum_{\substack{a_1\in\mathcal{F}_q^{k_1}\\ a_2,a\in\mathcal{F}_q^{k_2}}} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left\{\left[\boldsymbol{I} - \pi^{\sigma,\eta_2}_{w^n}\right]\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathbb{1}_{E_1\cap E_2\cap F_3}, \text{ where } E_j \triangleq \left\{\begin{array}{c}|\mathcal{L}_j(m_j,s_j^n)|\geq L_j, a_j(m_j,s_j^n)=a_j\\ v_j^n(a_j,m_j)=v_j^n, x_j^n(m_j,s_j^n)=x_j^n,\\ (s_j^n,v_j^n)\in T_{\eta_3}(p_{S_jV_j}), s_j^n\in T_{\frac{\eta_3}{2}}(\mathrm{p}_{s_j}),\end{array}\right\}, \quad (20)$$

$$\mathcal{F}_3 \triangleq \left\{\begin{array}{c}w^n(a,\underline{m})=w^n, a=a_1\,0^{k_2-k_1}\oplus a_2\\ (\underline{s}^n,\underline{x}^n\underline{v}^n,w^n)\in T_{4\eta_3}(p_{\underline{S}X\underline{V}W})\end{array}\right\} \quad (21)$$

The difficulty in analyzing $\zeta_{31}(\underline{m})$ is the fact that the typical projectors $\pi^{\sigma,\eta_2}_{a,\underline{m}}$ are with respect to state $\sigma_w : w \in \mathcal{W}$ and not $\sigma_{\underline{m},\underline{s}^n}$ on which the decoding POVM is applied. Computing the expectation enables us to average over the choice of $x_j(m_j, s_j^n)$ and thereby perform the state. In Appendix G, we prove the following proposition.

Proposition 2. *For every* $n \in \mathbb{N}$ *sufficiently large, we have* $\mathbb{E}\left\{\zeta_{31}(\underline{m})\right\} = \overline{\zeta}_{31}(\underline{m}) \leq \exp\left\{-n\left[(\eta_2 - 4\eta_3)^2 - 9\eta_3\right]\right\}$.

Upper bound on $\zeta_4(\underline{m})$ *:* Referring to $\zeta_4(\underline{m})$ in (15) and leveraging the definition of E_1, E_2 in (20), we have

$$\zeta_4(\underline{m}) \leq 2\sum_{\substack{\underline{s}^n,\underline{v}^n\\ \underline{x}^n,\hat{w}^n}}\sum_{a_1\in\mathcal{F}_q^{k_1}}\sum_{\substack{a_2,\hat{a}\\ \in\mathcal{F}_q^{k_2}}} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left\{\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{\hat{a},\underline{m}}\pi^{\mu}_{\eta_1}\rho_{\underline{m},\underline{s}^n}\right\}\mathbb{1}_{E_1\cap E_2\cap F_4}$$

$$= 2\sum_{\substack{\underline{s}^n,\underline{v}^n\\ \underline{x}^n,\hat{w}^n}}\sum_{a_1\in\mathcal{F}_q^{k_1}}\sum_{\substack{a_2,\hat{a}\\ \in\mathcal{F}_q^{k_2}}} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left\{\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{\hat{w}^n}\pi^{\mu}_{\eta_1}\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathbb{1}_{E_1\cap E_2\cap F_4} \text{ where } F_4 \triangleq \left\{\begin{array}{c}\hat{a}\neq a_1\,0^{k_2-k_1}\oplus a_2\\ w^n(\hat{a},\underline{m})=\hat{w}^n\end{array}\right\} \quad (22)$$

The key challenge in deriving an upper bound on $\overline{\zeta}_4(\underline{m})$ is the fact that, $W^n(\hat{a}, \underline{m})$ is not statistically independent of the choice $(X_j^n(m_j, s_j^n) : j = 1, 2)$. Indeed, $V_j(m_j, s_j^n)$ is a function of the whole bin $c_j(m_j) = (V_j(a_j, m_j) : a_j \in [q^{k_j}])$, and $W^n(\hat{a}, \underline{m})$ is an addition of codewords in $c_1(m_1)$ and $c_2(m_2)$. The standard proof technique crucially relies on this statistical independence which does not hold in this case. In Appendix H, we put forth a new sequence of steps to overcome this challenge and thereby prove the following proposition. As we have discussed in Sec. I, this new sequence of steps can be adopted in scenarios channel codes with bins such as the CQ broadcast channel [7], [8] among others [23].

[1]For certain sums wherein the summands range over the entire range we have explicitly stated this range for the sake of clarity/to remind the reader.

Proposition 3. *For every $n \in \mathbb{N}$ sufficiently large, we have*

$$\mathbb{E}\left\{\zeta_4(\underline{m})\right\} = \overline{\zeta}_4(\underline{m}) \leq 8\exp\left\{-n\left(H(Y)_\Upsilon - H(Y|W)_\Upsilon + \log q - H(W)_\Upsilon - 9\eta_3 - \eta_1 - 2\eta_2 - \frac{k_2}{n}\log q\right)\right\}.$$

Upper bound on $\zeta_5(\underline{m})$: Referring to $\zeta_5(\underline{m})$ in (15) and leveraging the definition of E_1, E_2 in (20), we have

$$\begin{aligned}\zeta_5(\underline{m}) &\leq 2\sum_{\substack{\underline{s}^n,\underline{v}^n\\ \underline{x}^n,\hat{\underline{w}}^n}}\sum_{\substack{a_1\in\mathcal{F}_q^{k_1}\\ \underline{\hat{m}}\neq\underline{m}}}\sum_{\substack{a_2,\hat{a}\\ \in\mathcal{F}_q^{k_2}}} \mathrm{p}_{\underline{S}}^n(\underline{s}^n)\,\mathrm{tr}\left\{\pi_{\eta_1}^{\mu}\pi_{\hat{a},\underline{\hat{m}}}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu}\rho_{\underline{m},\underline{s}^n}\right\}\mathbb{1}_{E_1\cap E_2\cap F_5}\\ &= 2\sum_{\substack{\underline{s}^n,\underline{v}^n\\ \underline{x}^n,\hat{\underline{w}}^n}}\sum_{\substack{a_1\in\mathcal{F}_q^{k_1}\\ \underline{\hat{m}}\neq\underline{m}}}\sum_{\substack{a_2,\hat{a}\\ \in\mathcal{F}_q^{k_2}}} \mathrm{p}_{\underline{S}}^n(\underline{s}^n)\,\mathrm{tr}\left\{\pi_{\eta_1}^{\mu}\pi_{\hat{w}^n}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu}\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathbb{1}_{E_1\cap E_2\cap F_5}, \text{ where } F_5 \triangleq \{w^n(\hat{a},\underline{\hat{m}}) = \hat{w}^n\} \end{aligned} \tag{23}$$

The analysis of $\overline{\zeta}_5(\underline{m})$, owing to the statistical independence of $W^n(\hat{m},\underline{\hat{m}})$ and the pair $(X_j^n(m_j,s_j^n) : j = 1,2)$ whenever $\underline{\hat{m}} \neq (m_1,m_2)$, is more straightforward. As we observe in the following proposition, exponent in the bound on $\overline{\zeta}_5(\underline{m})$ is smaller than the exponent in the bound on $\overline{\zeta}_4(\underline{m})$. Therefore the bound in the following proposition and influences the rate of the code. The proof of the following proposition in provided in Appendix I.

Proposition 4. *For every $n \in \mathbb{N}$ sufficiently large, we have*

$$\mathbb{E}\left\{\zeta_5(\underline{m})\right\} = \overline{\zeta}_5(\underline{m}) \leq 8\exp\left\{-n\left(H(Y)_\Upsilon - H(Y,W)_\Upsilon + \log q - 9\eta_3 - \eta_1 - 2\eta_2 - \frac{k_2+l_1+l_2}{n}\log q\right)\right\} \tag{24}$$

It must be noted that the exponent in the bound (24) features $\frac{k_2}{n}\log q$ because we have assumed $k_2 \geq k_1$. In general, the last term in the above stated exponent would be $\frac{\max\{k_1,k_2\}+l_1+l_2}{n}\log q$. Replacing the last term in the above exponent, we now collate the upper bounds we have derived on $\overline{\zeta}_i(\underline{m}) : i \in [5]$. Substituting (81), (92), (83), (101) and (110) and setting $\eta_1 = 5\eta_3$, $\eta_2 = 4\eta_3 + \sqrt{\frac{10\eta_3}{\delta_{\mathrm{q}}(\sigma,p_W,\eta_1)}}$, there exists a strictly positive $\kappa > 0$ such that for all $n \in \mathbb{N}$ sufficiently large,

$$\begin{aligned}\overline{\zeta}(\underline{m}) \leq \sum_{i=1}^{5}\overline{\zeta}_i(\underline{m}) &\leq 16\exp\left\{-n\left(\log q - H(W|Y)_\Upsilon - 23\eta_3 - \sqrt{\frac{10\eta_3}{\delta_{\mathrm{q}}(\sigma,p_W,\eta_1)}} - \frac{\max\{k_1,k_2\}+l_1+l_2}{n}\log q\right)\right\}\\ &\quad + \sum_{j=1}^{2}\exp\left\{-n\left(\frac{k_j}{n}\log q - [\log q - H(V_j|S_j)_\Upsilon + 3\eta_3]\right)\right\}\exp\left\{-n\kappa\eta_3^2\right\}. \end{aligned} \tag{25}$$

Performing a Fourier-Motzkin elimination on the four bounds

$$\frac{k_j+l_1+l_2}{n}\log q < \log q - H(W|Y)_\Upsilon - 24\eta_3 - \sqrt{\frac{10\eta_3}{\delta_{\mathrm{q}}(\sigma,p_W,\eta_1)}}, \frac{k_j}{n}\log q > \log q - H(V_j|S_j)_\Upsilon + 4\eta_3 : j = 1,2, \tag{26}$$

we obtain the achievability of the rate pair (R_1,R_2) satisfying $R_1+R_2 = \frac{l_1+l_2}{n}\log q < \min\{H(V_j|S_j)_\Upsilon : j \in [2]\} - H(W|Y)_\Upsilon - 24\eta_3 - \sqrt{\frac{10\eta_3}{\delta_{\mathrm{q}}(\sigma,p_W,\eta_1)}}$. Since $\eta_3 > 0$ is arbitrary and $\delta_{\mathrm{q}}(\sigma,p_W,\eta_1)$ is a positive constant, we can choose η_3 arbitrarily small. This completes proof of achievability of the rate region stated in the theorem. □

E. Non-Commutative Examples

We now identify examples of non-commutative QMSTx for which the inner bound characterized in Thm. 2 is strictly larger than that achievable via unstructured IID codes (Thm. 1).

Theorem 3. *Consider Ex. 1 and refer to $\mathcal{A}_u, \mathcal{A}_c$ defined in Thms. 1, 2 respectively. There exists $\theta \in (0,\frac{\pi}{2})$ for which $\mathcal{A}_u \subsetneq \mathcal{A}_c$.*

Proof. Our proof relies on two facts that can be easily verified. Firstly, the inner bounds $\mathcal{A}_u$ and $\mathcal{A}_c$ are continuous functions of the underlying space of QMSTxs when viewed as functions of the QMSTxs in question. Secondly, let us recall the inner bounds achievable via unstructured IID random codes and structured coset codes. Specifically, $\mathcal{R}$ defined in [4, Eqn. (30)] or equivalently $\mathcal{R}(\boldsymbol{\tau})$ defined in [13, Defn. 10] is the inner bounds achievable via unstructured IID random codes and henceforth denoted $\mathcal{B}_u(\boldsymbol{\tau})$ in this article. Similarly, let $\mathcal{B}_c(\boldsymbol{\tau})$, characterized

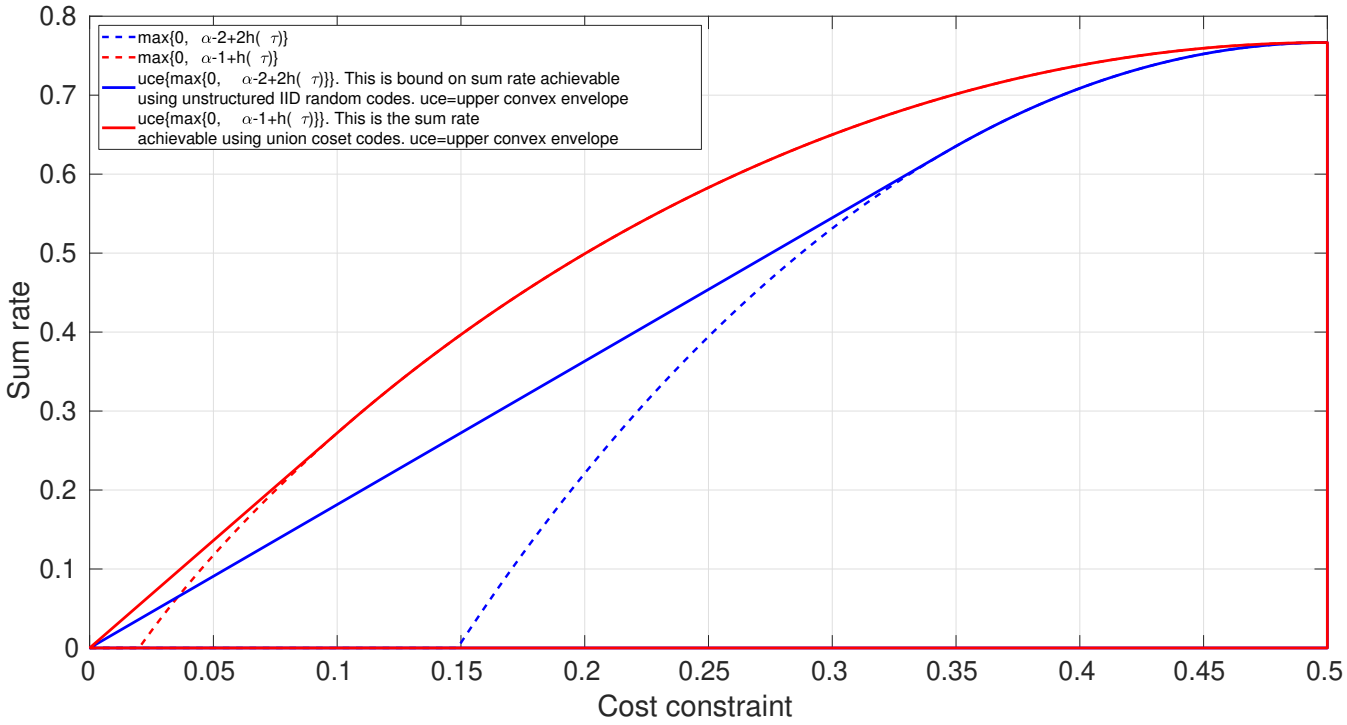

Fig. 14. Bound $uce\{\max\{0, \alpha - 2 + 2h_b(\tau)\}\}$ on the sum rate achievable via IID random codes is plotted in blue and the sum rate $uce\{\max\{0, \alpha - 1 + h_b(\tau)\}\}$ achievable via UCC is plotted in red.

as $\mathcal{R}_f(\boldsymbol{\tau})$, denote the inner bound achievable via structured coset codes for a commutative QMSTx. Through the characterizations of $\mathcal{A}_c$, $\mathcal{B}_c(\boldsymbol{\tau}), \mathcal{A}_u, \mathcal{B}_u(\boldsymbol{\tau})$, it is straight forward to verify that $\mathcal{A}_c$ reduces to $\mathcal{B}_c(\boldsymbol{\tau})$ and $\mathcal{A}_u$ reduces to $\mathcal{B}_u(\boldsymbol{\tau})$ when the QMSTx is commutative. We shall now leverage these two facts in the context of Ex. 1.

For Ex. 1, $\mathcal{B}_u(\boldsymbol{\tau}) = uce\{\max\{0, 2h_b(\tau) - 1\}\}$ and $\mathcal{B}_c(\boldsymbol{\tau}) = uce\{\max\{0, h_b(\tau)\}\}$. Since $\mathcal{B}_u(\boldsymbol{\tau}) \subsetneq \mathcal{B}_c(\boldsymbol{\tau})$, from the two facts stated above $\mathcal{A}_u \subsetneq \mathcal{A}_c$ for sufficiently small $\theta \in (0, \frac{\pi}{2})$. This completes our proof. □

While the above Thm. 3 and proof establishes the sub-optimality of unstructured IID random codes and the strict improvement of coset codes, the proof relies on a continuity argument. Philosof and Zamir's proof of sub-optimality of the unstructured IID random coding strategy is based on an stand-alone proof without appealing to continuity. Can we identify a non-commutative example and provide another such definitive proof? Unfortunately, this is involved as one must identify an optimal choice for parameters that saturate $\mathcal{A}_u$ for a chosen non-commutative QMSTx. This requires an ingenious argument and clever identification of a non-commutative QMSTx. Instead, in the following we provide a partial solution and prove that for a specific choice of parameters the corresponding inner bound achievable via structured coset codes is strictly larger than that achievable via unstructured IID random codes.

Proposition 5. *Consider Ex.1 for $\tau \in (0, \frac{1}{2})$ and $\theta = \frac{\pi}{8}$. There exists a choice of parameters, for which the inner bound achievable via UCCs is strictly larger than that achievable via unstructured codes.*

Proof. By choosing $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{F}_2$ the binary field and $p_{X_jV_j|S_j}(1, 1 \oplus s_j|s_j) = \tau = 1 - p_{X_jV_j|S_j}(0, s_j|s_j)$ for $s_j \in \{0, 1\}$ and $j \in [2]$ and evaluating the inner bound in Thm. 2, it can be verified that any rate pair (R_1, R_2) satisfying $R_1 + R_2 < uce\{\max\{0, \alpha - 1 + h_b(\tau)\}\}$ is achievable where α is as defined in (3). See Fig. 14 for plots of the rate regions $R_1 + R_2 < uce\{\max\{0, \alpha - 2 + 2h_b(\tau)\}\}$ and $R_1 + R_2 < uce\{\max\{0, \alpha - 1 + h_b(\tau)\}\}$ achievable via IID and structured codes respectively to verify the latter is strictly larger. □

Example 2. *Let $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{S}_1 = \mathcal{S}_2 = \{0, 1\}$, $\mathtt{p}_{\underline{S}}(\underline{s}) = \frac{1}{4}$ for every $\underline{s} \in \underline{\mathcal{S}}$, $\mu(0) \triangleq |0\rangle\langle 0|$ and $\mu(1) \triangleq |v_\theta\rangle\langle v_\theta|$, where $|v_\theta\rangle \triangleq [\cos\theta \ \ \sin\theta]^T$. For $(\underline{x}, \underline{s}) \in \{0, 1\}^4$, let $\rho_{x_1x_2s_1s_2} = \mu([x_1 \vee x_2] \oplus [x_2 \vee s_2])$, where $\vee$ denotes (binary) logical OR, $\oplus$ denotes addition in the binary field $\mathcal{F}_2$ and the cost function $\kappa_j(x_j, s_j) = \mathbb{1}_{\{x_j=1\}}$ is the Hamming weight function. For a $\tau \in (0, \frac{1}{2})$, what are the sum rates achievable via unstructured IID and union coset codes?*

Owing to the 'non-linear' relationship between the input symbols $\underline{s}, \underline{s}$ and the index of μ, it is difficult to analytically pin down the test channel maximizing the sum rate bound (1) via unstructured IID codes. We therefore

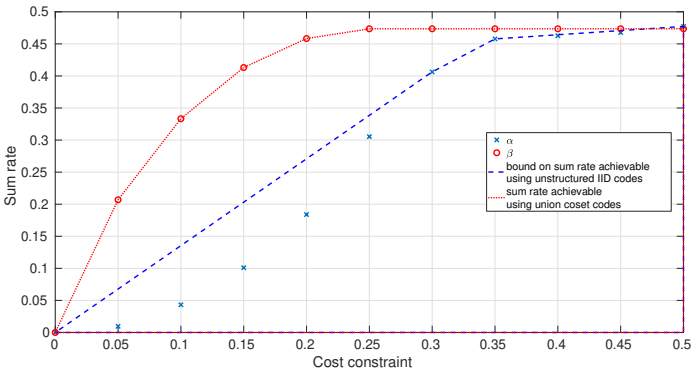

Fig. 15. Computed sum rates achievable using unstructured IID and union coset codes are plotted for Ex. 2 for a choice of $\theta = \frac{7\pi}{16}$. α and β denote the corresponding sum rates achievable in two cases considered before convexification.

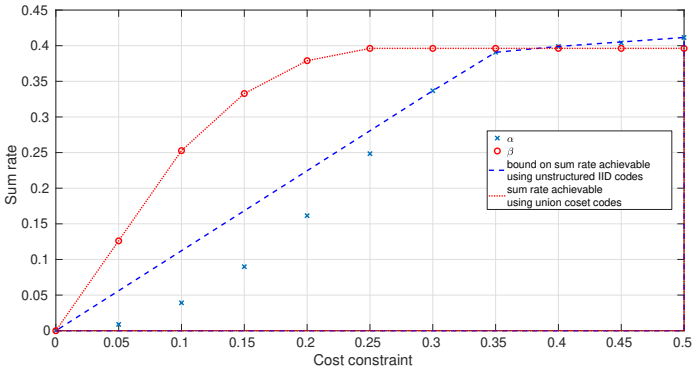

Fig. 16. Computed sum rates achievable using unstructured IID and union coset codes are plotted for Example 2 for a choice of $\theta = \frac{3\pi}{8}$. α and β denote the corresponding sum rates achievable in two cases considered before convexification.

resort to computation. In Fig. 15 we have plotted the sum rates achievable via union coset codes and unstructured codes for a choice of $\theta = \frac{7\pi}{16}$. In Fig. 16 we have plotted the sum rates achievable via union coset codes and unstructured codes for a choice of $\theta = \frac{3\pi}{8}$. The following remark summarizes three important observations.

Remark 3. *Firstly, our plots illustrate that union coset codes can enable communication at significantly higher rates in comparison to rates achievable via unstructured IID codes. Secondly, our computation reveals that Ex. 2 is a channel wherein the maximizing test channel distribution is non-uniform. Specifically, for a symmetric Hamming cost constraint of* $\tau \in [0, \frac{1}{4}]$*, the test channel that achieves the maximum sum rate using coset codes for the above examples is* $p_{U_jX_j|S_j}(0,0|0) = 1 - p_{U_jX_j|S_j}(1,1|0) = 1 - 2\tau$*,* $p_{U_jX_j|S_j}(1,0|1) = 1$*. This illustrates the imports of our findings in Thm. 2 and 4 that characterize achievable rate regions for all possible test channel distributions. Thirdly, Fig. 16 illustrates that unstructured IID codes can potentially achieve higher sum rate at Hamming cost constraints close to* $\frac{1}{2}$*. A general coding scheme must therefore incorporate both coding strategies to be able to achieve maximum possible rates. Our next section is aimed at designing and analyzing a unified coding scheme that incorporates both UCCs and unstructured IID codes.*

F. The Role of Union Coset Codes

Having detailed the proof, we are at an opportune point to explain the role of UCCs and why we chose the same over nested coset codes (NCCs). In contrast to unstructured IID random codes, when we impose structure, i.e. seek codes possessing structure and performing information-theoretic tasks such as packing and covering, we have to pay a rate penalty. Indeed, if we employed unstructured IID random codes, the covering bound in Prop. 1 would be $\frac{k_j \log q}{n} > I(V_j; S_j)_\Upsilon = H(V_j)_\Upsilon - H(V_j|S_j)_\Upsilon = \log q - H(V_j|S_j)_\Upsilon - [\log q - H(V_j)_\Upsilon]$. The first term being the lower bound in Prop. 1, we denote the excess rate - the term in $[\cdot]$ −parenthesis - required in structured coset coding as $\beta_j \triangleq [\log q - H(V_j)_\Upsilon]$. How are we able to obtain gains over unstructured IID random codes despite of paying this penalty in covering?

A closer look at the packing bound in Prop. 4 answers this question. For a moment, let us say we had employed unstructured IID random codes and had found a way to decode the sum of the chosen codewords. In that case the bound would have been $\frac{k_j+l_1+l_2}{n} \log q < I(W;Y)_\Upsilon = \log q - H(W|Y)_\Upsilon - [\log q - H(W)_\Upsilon]$. In other words, unstructured IID random code rate is more constrained that structured coset codes in regards to packing. Some or whole of the advantage that unstructured IID random codes accrue in covering is lost in packing. Precisely, $\max\{H(W)_\Upsilon - H(V_j)_\Upsilon : j \in [2]\}$ is the advantage enjoyed by unstructured IID random codes for not imposing structure. However, this competes with the losses it suffers due to its insistence of decoding the pair and the associated effect of not optimally allocating the available limited output space.

The above discussion also alerts the reader to the rate penalties one has to pay for imposing structure. One should therefore impose structure only in those codes - packing or covering - that are exploited in encoding or decoding. Here, we are only exploiting the fact that all of the coarser codes are coset shifts of a particular linear code. We are

not needing that these coset shifts be structured. In other words, we are not requiring that the different coset shifts, when viewed in totality, form a coset of a linear code. The finer or outer code need not possess algebraic closure properties. Imposing that would unnecessarily entail a rate penalty. This motivates our choice of UCC over NCC. We refer the reader to Remark 5 and [13] wherein group based codes entail different rate penalties motivating the choice of UCC over NCC even more important.

V. Enhancing IID Coding Schemes via UCCs

The UCC based coding scheme can enable efficient decoding of $V_1 \oplus V_2$. On a QMSTx wherein the latter function contains the information of the pair of messages, the UCC coding scheme can outperform the use of unstructured codes. In general, the information corresponding to the message pair can be embedded in both uni-variate and bi-variate functions of auxiliary RVs. As our computation for Ex. 2 for the case $\theta = \frac{3\pi}{8}$ indicates, while UCC codes outperform IID codes for low cost constraints, the latter can potentially achieve higher rates for large cost constraints. It must also be noted that since structured coding scheme employs statistically correlated codebooks, they cannot be employed to decode the pair of codewords. For example, if one considers a simple classical binary additive MAC $Y = X_1 \oplus X_2$ without states, it is impossible for the Rx to decode both codewords if both Txs employ cosets of a common linear code. A general coding scheme for QMSTx must therefore incorporate both unstructured codes and UCCs. We present the following inner bound that subsumes inner bounds stated in Thms. 1, 2.

Theorem 4. *A rate-cost $(\underline{R}, \underline{\tau}) \in \mathcal{A} \subseteq [0,\infty)^4$ quadruple is achievable if there exists finite sets $\mathcal{U}_1, \mathcal{U}_2$, a finite field $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{W} = \mathcal{F}_q$ of size q and conditional PMFs $p_{U_jV_jX_j|S_j} : j \in [2]$ with respect to which*

$$\begin{aligned} R_j &\leq I(U_j; U_{\bar{j}}Y)_\Upsilon - I(U_j; S_j)_\Upsilon + \min\{I(V_j; V_{\bar{j}}, \underline{U}, Y) - I(V_1 \oplus V_2; V_{\bar{j}}|\underline{U}Y) - I(V_j; U_j, S_j) : j \in [2]\} \\ R_1 + R_2 &\leq I(\underline{U}; Y)_\Upsilon - I(\underline{U}; \underline{S})_\Upsilon + \min\{I(V_j; V_{\bar{j}}, \underline{U}, Y) - I(V_1 \oplus V_2; V_{\bar{j}}|\underline{U}Y) - I(V_j; U_j, S_j) : j \in [2]\} \end{aligned} \tag{27}$$

where the above entropies are evaluated with respect to the state

$$\Upsilon^{Y\underline{X}\underline{U}\underline{V}W\underline{S}} \triangleq \sum_{\underline{s},\underline{u},\underline{v},w,\underline{x}} p_{\underline{S}\underline{U}\underline{V}W\underline{X}}(\underline{s},\underline{u},\underline{v},w,\underline{x}) \rho_{\underline{x}\underline{s}} \otimes |\underline{x}\ \underline{u}\ \underline{v}\ w\ \underline{s}\rangle\langle \underline{x}\ \underline{u}\ \underline{v}\ w\ \underline{s}|\,, \tag{28}$$

$$p_{\underline{S}\underline{U}\underline{V}W\underline{X}}(\underline{s},\underline{u},\underline{v},w,\underline{x}) = \mathtt{p}_{\underline{S}}(\underline{s}) \prod_{j=1}^{2} p_{X_jV_jU_j|S_j}(x_j, v_j, u_j|s_j) \mathbb{1}_{\{v_1 \oplus v_2 = w\}}. \tag{29}$$

for all $(\underline{s},\underline{v},w,\underline{x}) \in \underline{\mathcal{S}} \times \underline{\mathcal{V}} \times \mathcal{W} \times \underline{\mathcal{X}}$.

By choosing $\mathcal{V}_1 = \mathcal{V}_2 = \phi$, we can recover the inner bound achievable via IID codes in Thm. 1. By choosing $\mathcal{U}_1 = \mathcal{U}_2 = \phi$, we can recover the inner bound in Thm. 2, thus proving that the above inner bound subsumes all known inner bounds for a general QMSTx.

Proof. We begin with an outline of the code structure, decoding POVMs and the tools/techniques we leverage for our error probability analysis. As mentioned earlier, our approach is one of amalgamating the unstructured IID coding scheme with the UCC based coding scheme. A pair of unstructured codes - one for each sender - identical to that used in a proof of Thm. 1 is designed on auxiliary alphabets $\mathcal{U}_1, \mathcal{U}_2$. Each of these codes is partitioned into bins to enable the encoder choose codewords jointly typical/compatible with the observed state sequence. A pair of UCCs - one for each sender - identical in structure to that employed in proof of Thm. 2 is designed on auxiliary alphabet $\mathcal{V} = \mathcal{V}_1 = \mathcal{V}_2 = \mathcal{F}_q$. Sender j's message comprises of two components $m_{j1} \in [\mathcal{M}_{j1}]$ and $m_{j2} \in [\mathcal{M}_{j2}] \triangleq \mathcal{F}_q^{l_j}$ communicated via the $\mathcal{U}_j-$ and $\mathcal{V}_j-$codebooks respectively.

Our decoding will leverage simultaneous (joint) and successive decoding techniques. The first layer decoding will employ a joint decoding POVM to decode into the $\mathcal{U}_1, \mathcal{U}_2-$codebooks. Following this, a second stage POVM will decode into the sum UCC codebook analogous to that in proof of Thm. 2.

Our error probability analysis will leverage techniques developed in prior works [37] to handle the complexities of joint decoding over a $2-$user QMAC, successive decoding [40] and codebook with bins. Specifically, the decoding POVM and the techniques developed in [37] to analyze joint decoding is enhanced with the list decoding technique proposed in Sec. IV-D to analyze joint decoding into the $\mathcal{U}_1 - \mathcal{U}_2-$codebooks comprising of bins. The analysis of the second stage decoding into the sum of the UCCs is identical to that developed in Sec. IV-D.

A. Code Structure

In contrast to Sec. IV, sender j's message comprises of two components $m_{j1} \in [\mathcal{M}_{j1}]$ and $m_{j2} \in [\mathcal{M}_{j2}] \triangleq \mathcal{F}_q^{l_j}$, communicated via the $\mathcal{U}_j-$ and $\mathcal{V}_j-$codebooks respectively. We let $R_{ji} \triangleq \frac{\log \mathcal{M}_{ji}}{n}$ and $R_j = R_{j1} + R_{j2}$ for $j, i \in [2]$. To distinguish from an underline that we have employed to aggregate random variables, messages, sets *across* users, we let a tilde at the bottom to aggregate message components of a single user. We let $[\underline{\mathcal{M}}_j] \triangleq [\underline{\mathcal{M}}_{j1}] \times [\underline{\mathcal{M}}_{j2}]$ denote the aggregation of sender j's message and $\underline{m}_j \triangleq (m_{j1}, m_{j2})$ for $j \in [2]$ denote a generic message of sender j. As stated in the theorem, $\mathcal{U}_1, \mathcal{U}_2$ are finite sets and $\mathcal{V} = \mathcal{V}_1 = \mathcal{V}_2 = \mathcal{F}_q$ is a finite field of size q. For $j \in [2]$, sender j's code comprises of an unstructured (IID random) code and a UCC, henceforth referred to as $\mathcal{U}_j-$code and $\mathcal{V}_j-$code respectively. $\mathcal{U}_j-$code comprises of $\mathcal{M}_{j1}$ bins $\beta_j(m_{j1}) : m_{j1} \in [\mathcal{M}_{j1}]$, each of which consists of $\mathcal{B}_j$ codewords. Specifically, the $\mathcal{U}_j-$ code is $(\beta_j(m_{j1}) = (u_j^n(m_{j1}, b_j) \in \mathcal{U}_j^n : b_j \in [\mathcal{B}_j]) : m_{j1} \in [\mathcal{M}_{j1}])$ with $u_j^n(m_{j1}, b_j)$ denoting a generic codeword in bin $\beta_j(m_{j1})$. Sender j's $\mathcal{V}_j-$code is a UCC code $(n, k_j, l_j, g_j, \iota_j)$ comprising of q^{l_j} cosets $c_j(m_{j2}) : m_{j2} \in [\mathcal{M}_{j2}] = \mathcal{F}_q^{l_j}$. Specifically, the $\mathcal{V}_j-$code is $(c_j(m_{j2}) = (v_j^n(a_j, m_{j2}) \triangleq a_j g_j \oplus \iota_j(m_{j2}) : a_j \in \mathcal{F}_q^{k_j}) : m_{j2} \in [\mathcal{M}_{j2}])$ wherein $v_j^n(a_j, m_{j2})$ denotes a generic codeword in the coset/bin $c_j(m_{j2})$. Just as in proof of Thm. 2, we assume $k_2 \geq k_1$ without loss of generality and enforce $g_2 = \begin{bmatrix} g_1^T & g_{2/1}^T \end{bmatrix}^T$. This ensures that the bins of user 1's UCC $\mathcal{V}_1-$code are sub-cosets of the bins of user 2's UCC $\mathcal{V}_2-$code.

B. Encoding

On observing message $\underline{m}_j = (m_{j1}, m_{j2})$ and the state sequence s_j^n, encoder j first looks for a codeword in $\beta_j(m_{j1})$ that is jointly typical with s_j^n. If it finds at least one, one among these is chosen and denoted $u_j^n(m_{j1}, s_j^n)$. Next, the encoder looks for a codeword in the coset $c_j(m_{j2})$ that is jointly typical with the pair $(s_j^n, u_j^n(m_{j1}, s_j^n))$. If it finds at least one such codeword, one among these is chosen and denoted $v_j^n(\underline{m}_j, s_j^n)$. Moreover, let $b_j(m_{j1}, s_j^n)$ and $a_j(\underline{m}_j, s_j^n)$ denote the index of the chosen codewords $u_j(m_{j1}, s_j^n)$ and $v_j^n(\underline{m}_j, s_j^n)$ respectively. In other words, let $u_j^n(m_{j1}, b_j(m_{j1}, s_j^n)) = u_j(m_{j1}, s_j^n)$ and $v_j^n(a_j(\underline{m}_j, s_j^n), m_{j2}) = v_j^n(\underline{m}_j, s_j^n)$. If any of the above steps returns no choices, a default pair of codewords from the pair $\beta_j(m_{j1})$ and $c_j(m_{j2})$ bins is chosen. For the sake of the ensuing analysis, we formalize this encoding with notation.

Let $\alpha_{j1}(m_{j1}, s_j^n) = \sum_{b_j} \mathbb{1}_{\left\{(u_j^n(m_{j1}, b_j), s_j^n) \in T_{\frac{\eta_3}{2}}(p_{U_j S_j})\right\}}$ be the number of available jointly typical codewords in the $\mathcal{U}_j-$codebook. Let

$$\mathcal{L}_{j1}(m_{j1}, s_j^n) \triangleq \begin{cases} \left\{b_j : (u_j^n(m_{j1}, b_j), s_j^n) \in T_{\frac{\eta_3}{2}}(p_{U_j S_j})\right\} & \text{if } \alpha_{j1}(m_{j1}, s_j^n) \geq 1 \\ \{0\} & \text{otherwise, i.e. } \alpha_{j1}(m_{j1}, s_j^n) = 0. \end{cases} \tag{30}$$

Let $\alpha_{j2}(\underline{m}_j, s_j^n) \triangleq \sum_{a_j} \mathbb{1}_{\{(u_j(m_{j1}, b_j(m_{j1}, s_j^n)), v_j^n(a_j, m_{j2}), s_j^n) \in T_{\eta_3}(p_{U_j V_j S_j})\}}$ be the number of available jointly typical codeword triplets for a chosen $\mathcal{U}_j-$codeword and

$$\mathcal{L}_{j2}(\underline{m}_j, s_j^n) \triangleq \begin{cases} \left\{a_j : (u_j(m_{j1}, b_j(m_{j1}, s_j^n)), v_j^n(a_j, m_{j2}), s_j^n) \in T_{\eta_3}(p_{U_j V_j S_j})\right\} & \text{if } \alpha_{j2}(\underline{m}_j, s_j^n) \geq 1 \\ \{0\} & \text{otherwise, i.e. } \alpha_{j2}(\underline{m}_j, s_j^n) = 0. \end{cases} \tag{31}$$

For every pair $(\underline{m}_j, s_j^n)$, an element is chosen from $\mathcal{L}_{j1}(m_j, s_j^n)$ and denoted/defined $b_j(m_{j1}, s_j^n)$. We define $u_j^n(m_{j1}, s_j^n) \triangleq u_j^n(m_{j1}, b_j(m_{j1}, s_j^n))$. Next, for the pair $(\underline{m}_j, s_j^n))$, an element is chosen from $\mathcal{L}_{j2}(\underline{m}_j, s_j^n)$ and defined/denoted $a_j(\underline{m}_j, s_j^n)$ and[2] we define $v_j^n(\underline{m}_j, s_j^n) \triangleq v_j^n(a_j(\underline{m}_j, s_j^n), m_{j2})$. A predefined 'fusion map' $f_j : \mathcal{S}_j^n \times \mathcal{U}_j^n \times \mathcal{V}_j^n \to \mathcal{X}_j^n$ is used to map the triplet $s_j^n, u_j^n(m_{j1}, s_j^n), v_j^n(\underline{m}_j, s_j^n)$ to an input sequence in $\mathcal{X}_j^n$ henceforth denoted $x_j^n(\underline{m}_j, s_j^n)$.

C. Decoding POVMs

Having let $\underline{m}_j$ denote the two components of sender j's message, we let $\underline{m}_1 \triangleq (m_{11}, m_{21}) \in [\underline{\mathcal{M}}_1] \triangleq [\mathcal{M}_{11}] \times [\mathcal{M}_{21}]$ denote the components of the two senders messages indexing the $\mathcal{U}_1-, \mathcal{U}_2-$codebooks and $\underline{m}_2 \triangleq (m_{12}, m_{22}) \in [\underline{\mathcal{M}}_2] \triangleq [\mathcal{M}_{12}] \times [\mathcal{M}_{22}] = \mathcal{F}_q^{l_1 + l_2}$ denote the components of the two senders messages indexing the $\mathcal{V}_1-, \mathcal{V}_2-$codebooks. As we stated at the beginning of our proof, we employ a simultaneous (joint) decoding POVM to decode the messages $\underline{m}_1$ indexing the $\mathcal{U}_1-, \mathcal{U}_2-$codebooks. The joint decoding POVM is

[2]The dependence of $a_j(\underline{m}_j, s_j^n)$ on the choice $b_j(m_{j1}, s_j^n)$ is implicit and represented through m_{j1}.

designed based on [37, Sec. V.B]. Next, a successive decoding POVM recovers the components of $\underline{m}_2$. Indeed, recall that the pair of messages is recovered in Sec. IV by decoding into a single sum codebook. Our decoding POVM is $\left\{\sqrt{\lambda_{\underline{m}_1}}\theta_{\underline{m}_2|\underline{m}_1}\sqrt{\lambda_{\underline{m}_1}} : (\underline{m}_1,\underline{m}_2)\in[\underline{\mathcal{M}}_1]\times[\underline{\mathcal{M}}_2]\right\}\cup\{\lambda_{-1},\sqrt{\lambda_{\underline{m}_1}}\theta_{-1|\underline{m}_1}\sqrt{\lambda_{\underline{m}_1}} : \underline{m}_1\in[\underline{\mathcal{M}}_1]\}$. We begin by specifying the POVM $\left\{\lambda_{\underline{m}_1} : \underline{m}_1\in[\underline{\mathcal{M}}_1],\lambda_{-1}\right\}$ of the first stage joint decoder.

Let $\mu \triangleq \sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}}(\underline{x},\underline{s})\rho_{\underline{x},\underline{s}}$, $\sigma_{\underline{u}} \triangleq \sigma_{u_1u_2} \triangleq \sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}|\underline{U}}(\underline{x},\underline{s}|\underline{u})\rho_{\underline{x},\underline{s}}$ and $\sigma^j_{u_j} \triangleq \sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}|U_j}(\underline{x},\underline{s}|u_j)\rho_{\underline{x},\underline{s}}$ for $j\in[2]$. We let π^{μ}_{η} denote the $\eta-$unconditional typical projector of $\mu^{\otimes n}$, $\pi^{\sigma,\eta}_{\underline{u}^n}$ denote the $\eta-$conditional typical projector of $\sigma_{\underline{u}^n}\triangleq\otimes_{t=1}^n\sigma_{u_{1t},u_{2t}}$, $\pi^{j,\eta}_{u_j^n}$ denote the $\eta-$conditional typical projector of $\sigma^j_{u_j^n}\triangleq\otimes_{t=1}^n\sigma^j_{u_{jt}}$. Furthermore, for $j\in[2]$, we let $\pi^{j,\eta}_{m_{j1},b_j}\triangleq\pi^{\sigma_j,\eta}_{m_{j1},b_j}$ denote the conditional typical projector of $\sigma^j_{u_j^n(m_{j1},b_j)}=\otimes_{t=1}^n\sigma^j_{u_j(m_{j1},b_j)_t}$, $\pi^{\sigma,\eta}_{\underline{m}_1,\underline{b}}$ be the conditional typical projector of $\sigma_{u_1^n(m_{11},b_1)u_2^n(m_{21},b_2)}=\otimes_{t=1}^n\sigma_{u_1(m_{11},b_1)_tu_2(m_{21},b_2)_t}$. With these definitions, we let $\alpha_{\underline{m}_1,\underline{b}}\triangleq\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{m_{11},b_1}\pi^{\sigma,\eta_1}_{\underline{m},\underline{b}}\pi^{1,\eta_2}_{m_{11},b_1}\pi^{\mu}_{\eta_4}$,

$$\lambda^1_{\underline{m}_1}=\left(\sum_{\hat{\underline{m}}_1}\sum_{\underline{b}}\alpha_{\hat{\underline{m}}_1,\underline{b}}\right)^{-\frac{1}{2}}\sum_{\underline{b}}\alpha_{\underline{m}_1,\underline{b}}\left(\sum_{\hat{\underline{m}}_1}\sum_{\underline{b}}\alpha_{\hat{\underline{m}}_1,\underline{b}}\right)^{-\frac{1}{2}}$$

and $\lambda^1_{-1}\triangleq\boldsymbol{I}-\sum_{\underline{m}_1}\lambda^1_{\underline{m}_1}$. Next, we specify the second stage POVM.

Consider the UCC $(n,k_2,l_1+l_2,g_2,\iota_{\oplus})$ where $\iota_{\oplus}(\underline{m}_2)=\iota_1(m_{21})\oplus\iota(m_{22})$ for $\underline{m}_2=(m_{21},m_{22})\in\mathcal{F}_q^{l_1}\times\mathcal{F}_q^{l_2}$ and let $w^n(a,\underline{m}_2)\triangleq ag_2\oplus\iota_1(m_{21})\oplus\iota_2(m_{22})$ denote its codewords. Referring to Appendix A-B, we let $\pi^{\delta_{w^n(a,\underline{m}_2)}}_{p_W,\eta_7}$ be the η_7-conditional typical projector of $\otimes_{t=1}^n\delta_{w_t(a,\underline{m}_2)}$ with respect to the PMF p_W, where $\delta_w\triangleq\sum_{\underline{x},\underline{s}}p_{\underline{X}\underline{S}|W}(\underline{x},\underline{s}|w)\rho_{\underline{x}\underline{s}} : w\in\mathcal{W}$ where $p_{\underline{S}\underline{X}W}$ is the corresponding marginal of $p_{\underline{S}\underline{U}\underline{V}W\underline{X}}$ defined in (28). As stated in Appendix A-B, we henceforth let $\pi^{\delta,\eta_7}_{a,\underline{m}_2}=\pi^{\delta_{w^n(a,\underline{m}_2)}}_{p_W,\eta_2}$. We define $\gamma_{a,\underline{m}_2|\underline{m}_1}\triangleq\pi^{\sigma,\eta_6}_{\underline{m}_1,\underline{b}^*}\pi^{\delta,\eta_7}_{a,\underline{m}_2}\pi^{\sigma,\eta_6}_{\underline{m}_1,\underline{b}^*}$ where $\pi^{\sigma,\eta_6}_{\underline{m}_1,\underline{b}^*}$ is the η_6-conditional typical projector of the state $\sigma_{u_1^n(m_{11},b_1^*)u_2^n(m_{21},b_2^*)}=\otimes_{t=1}^n\sigma_{u_1(m_{11},b_1^*)_tu_2(m_{21},b_2^*)_t}$. With these definitions, we let

$$\theta_{\underline{m}_2|\underline{m}_1}\triangleq\left(\sum_{\hat{a},\hat{m}_{21},\hat{m}_{22}}\gamma_{\hat{a},\hat{\underline{m}}_2|\underline{m}_1}\right)^{-\frac{1}{2}}\sum_{a}\gamma_{a,\underline{m}_2|\underline{m}_1}\left(\sum_{\hat{a},\hat{m}_{21},\hat{m}_{22}}\gamma_{\hat{a},\hat{\underline{m}}_2|\underline{m}_1}\right)^{-\frac{1}{2}}\tag{32}$$

and $\theta_{-1}\triangleq\boldsymbol{I}-\sum_{\underline{m}_2}\theta_{\underline{m}_2|\underline{m}_1}$.

D. Probability of Error Analysis

As in proof of Thm. 2, we employ the random coding technique and begin by identifying the components that make up our coding scheme, followed by characterizing an upper bound on the error probability comprising of multiple terms. The first stage of our coding scheme is completely characterized via the collections $\left(u^n_j(m_{j1},b_j)\in\mathcal{U}^n_j : (m_{j1},b_j)\in[\mathcal{M}_{j1}]\times[\mathcal{B}_j]\right)$ and $\left(b_j(m_{j1},s^n_j)\in[\mathcal{B}_j] : (m_{j1},s^n_j)\in[\mathcal{M}_{j1}]\times\mathcal{S}^n_j\right)$ for $j\in[2]$. The second stage is completely characterized via the generator matrices $g_1,g_{2/1}$, the collection $(\iota_j(m_{j2}) : m_{j2}\in[q^{l_j}])$ of dither/bias vectors specifying the coset shifts, the indices $\left(a_j(\mathfrak{m}_j,s^n_j) : (\mathfrak{m}_j,s^n_j)\in[\mathcal{M}_j]\times\mathcal{S}^n_j\right)$, and the final codeword choices $\left(x^n_j(\mathfrak{m}_j,s^n_j) : (\mathfrak{m}_j,s^n_j)\in[\mathcal{M}_j]\times\mathcal{S}^n_j\right)$ each for $j\in2$. For a generic code specified through these objects, we now characterize an upper bound on the error probability.

An upper bound on the error probability for a generic code : We let $[\underline{\mathcal{M}}]\triangleq[\mathcal{M}_1]\times[\mathcal{M}_2]$ and $\underline{m}\triangleq(\mathfrak{m}_1,\mathfrak{m}_2)\in[\underline{\mathcal{M}}]$ denote a generic message pair of the two senders. We remark that $\underline{m}=(\mathfrak{m}_1,\mathfrak{m}_2)=(\underline{m}_1,\underline{m}_2)$, where the components of the first representation are the two sender's messages and the components of the second representation are the message pairs indexing the unstructured and UCC codebooks. For a generic choice of the aforementioned objects, it can be verified using the cyclicity of the trace that the average error probability is

$$\xi=\sum_{\underline{m}}\frac{\hat{\xi}(\underline{m})+\hat{\zeta}(\underline{m})}{|\underline{\mathcal{M}}|}\text{ where }\hat{\xi}(\underline{m})\triangleq\sum_{\underline{s}^n}\mathrm{p}^n_{\underline{S}}(\underline{s}^n)\hat{\xi}(\underline{m}|\underline{s}^n),\hat{\zeta}(\underline{m})\triangleq\sum_{\underline{s}^n}\mathrm{p}^n_{\underline{S}}(\underline{s}^n)\hat{\zeta}(\underline{m}|\underline{s}^n),\hat{\xi}(\underline{m}|\underline{s}^n)\triangleq\mathrm{tr}\left\{\left(\boldsymbol{I}-\lambda_{\underline{m}_1}\right)\rho_{\underline{m},\underline{s}^n}\right\}\tag{33}$$

$$\hat{\zeta}(\underline{m}|\underline{s}^n)\triangleq\mathrm{tr}\left\{\left(\boldsymbol{I}-\theta_{\underline{m}_2|\underline{m}_1}\right)\sqrt{\lambda_{\underline{m}_1}}\rho_{\underline{m},\underline{s}^n}\sqrt{\lambda_{\underline{m}_1}}\right\},\rho_{\underline{m},\underline{s}^n}\triangleq\bigotimes_{t=1}^n\rho_{x_1(\underline{m}_1,s_1^n)_tx_2(\underline{m}_2,s_2^n)_t\underline{s}_t},\tag{34}$$

$\boldsymbol{I} = I^{\otimes n}$, $|\underline{\mathcal{M}}| = |\mathcal{M}_{11}| \cdot |\mathcal{M}_{21}| \cdot q^{l_1+l_2}$. We consider an arbitrary pair $\underline{m} = (m_1, m_2)$ and henceforth focus our study on $\hat{\xi}(\underline{m})$ and $\hat{\zeta}(\underline{m})$. We begin with $\hat{\xi}(\underline{m})$.

An upper bound on $\hat{\xi}(\underline{m})$: Since we have fixed an arbitrary pair $\underline{m} = (\underline{m}_1, \underline{m}_2)$, we henceforth let $b_j^* \triangleq b_j(m_{j1}, s_j^n)$ and $\underline{b}^* \triangleq (b_1^*, b_2^*)$. Since $\alpha_{\underline{m}_1,\underline{b}} \geq 0$ is PSD for every $\underline{b} \in [\mathcal{B}_1] \times [\mathcal{B}_2]$, we have

$$\lambda_{\underline{m}} \geq (S+T)^{-\frac{1}{2}} S (S+T)^{-\frac{1}{2}} \text{ where } S = \alpha_{\underline{m}_1,\underline{b}^*} = \pi^{\mu}_{\eta_4} \pi^{1,\eta_2}_{m_{11},b_1^*} \pi^{\sigma,\eta_1}_{\underline{m}_1,\underline{b}^*} \pi^{1,\eta_2}_{m_{11},b_1^*} \pi^{\mu}_{\eta_4} \text{ and } T = \sum_{\underline{b},\hat{\underline{m}}_1} \alpha_{\hat{\underline{m}}_1,\underline{b}} \mathbb{1}_{\left\{(\hat{m}_1,\hat{b}) \neq (\underline{m}_1,\underline{b}^*)\right\}},$$

$$\text{and hence, } \hat{\xi}(\underline{m}|s^n) \leq \xi(\underline{m}|\underline{s}^n), \text{ where } \xi(\underline{m}|\underline{s}^n) \triangleq \operatorname{tr}\left(\left[\boldsymbol{I} - (S+T)^{-\frac{1}{2}} S (S+T)^{-\frac{1}{2}}\right] \rho_{\underline{m},\underline{s}^n}\right). \tag{35}$$

Our study of $\hat{\xi}(\underline{m})$ will henceforth focus on $\xi(\underline{m}) \triangleq \sum_{\underline{s}^n} \mathrm{p}_{\underline{S}}(\underline{s}^n) \xi(\underline{m}|\underline{s}^n)$. Analogous to our proof of Thm. 2, we split the event corresponding to $\xi(\underline{m})$ into $\mathcal{E}$ and $\overline{\mathcal{E}}$ and analyze the terms corresponding to the these two events separately. Let $\mathcal{E}_{j1} \triangleq \left\{\underline{s}^n \in T_{\frac{\eta_5}{4}}(\mathrm{p}_{\underline{S}})\right\}$,

$$\mathcal{E}_{j2} \triangleq \left\{ \begin{matrix} |\mathcal{L}_{j1}(m_{j1}, s_j^n)| \geq L_{j1}, |\mathcal{L}_{j2}(m_j, s_j^n)| \geq L_{j2} \\ (s_j^n, u_j^n(m_{j1}, s_j^n), v_j^n(m_j, s_j^n)) \in T_{\eta_5}(p_{S_j U_j V_j}) \end{matrix} \right\}, \mathcal{E}_{121} \triangleq \left\{ \begin{pmatrix} s_j^n, u_j^n(m_{j1}, s_j^n), \\ v_j^n(m_j, s_j^n) : j \in [2] \end{pmatrix} \in T_{2\eta_5}(p_{\underline{SUV}}) \right\}$$

$$\mathcal{E}_{122} \triangleq \left\{ \begin{pmatrix} s_j^n, u_j^n(m_{j1}, s_j^n), v_j^n(m_j, s_j^n), \\ x_j^n(m_j, s_j^n) : j \in [2] \end{pmatrix} \in T_{4\eta_5}(p_{\underline{SUVX}}) \right\} \text{ and finally } \mathcal{E} \triangleq \bigcap_{j=1}^{2} \mathcal{E}_{j1} \cap \mathcal{E}_{j2} \bigcap \mathcal{E}_{121} \cap \mathcal{E}_{122},$$

where $L_{j1} = \frac{1}{2} \exp\left\{n\left(\frac{\log|\mathcal{B}_j|}{n} - I(U_j;S_j)_{\Upsilon} - \frac{3\eta_5}{2}\right)\right\}$, $L_{j2} = \frac{1}{2}\exp\left\{n\left(\frac{\log|\mathcal{B}_j|}{n} - \log q + H(V_j|S_j,U_j)_{\Upsilon} - 3\eta_5\right)\right\}$. We remark that all Von Neumann entropies are evaluated with respect to the state characterized in (28). Since

$$\overline{\mathcal{E}} = \bigcup_{j=1}^{2} \overline{\mathcal{E}}_{j1} \cup \overline{\mathcal{E}}_{j2} \bigcup \overline{\mathcal{E}}_{121} \cup \overline{\mathcal{E}}_{122} = \bigcup_{j=1}^{2} \mathcal{F}_{j1} \cup \mathcal{F}_{j2} \bigcup \mathcal{F}_{121} \cup \mathcal{F}_{122} \text{ where } \mathcal{F}_{j1} \triangleq \overline{\mathcal{E}}_{j1}, \mathcal{F}_{j2} \triangleq \mathcal{E}_{j1} \cap \overline{\mathcal{E}}_{j2}, \tag{36}$$

$$\mathcal{F}_{121} \triangleq \bigcap_{j=1}^{2} \mathcal{E}_{j1} \cap \mathcal{E}_{j2} \cap \overline{\mathcal{E}}_{121}, \mathcal{F}_{122} \triangleq \bigcap_{j=1}^{2} \mathcal{E}_{j1} \cap \mathcal{E}_{j2} \cap \mathcal{E}_{121} \cap \overline{\mathcal{E}}_{122} \text{ we have } 1 = \mathbb{1}_{\overline{\mathcal{E}}} + \mathbb{1}_{\mathcal{E}} \leq \sum_{j=1}^{2} \sum_{i=1}^{2} \left(\mathbb{1}_{\mathcal{F}_{ji}} + \frac{\mathbb{1}_{\mathcal{F}_{12i}}}{2}\right) + \mathbb{1}_{\mathcal{E}} \tag{37}$$

With these definitions, we have

$$\xi(\underline{m}) \leq \sum_{j=1}^{2} \xi_j(\underline{m}) + \tilde{\xi}_2(\underline{m}), \text{ where } \xi_j(\underline{m}) \triangleq \sum_{\underline{s}^n} \mathrm{p}^n_{\underline{S}}(\underline{s}^n) \xi(\underline{m}|\underline{s}^n) \sum_{i=1}^{2} \left(\mathbb{1}_{\mathcal{F}_{ji}} + \frac{\mathbb{1}_{\mathcal{F}_{12i}}}{2}\right), \tilde{\xi}_2(\underline{m}) \triangleq \sum_{\underline{s}^n} \mathrm{p}^n_{\underline{S}}(\underline{s}^n) \xi(\underline{m}|\underline{s}^n) \mathbb{1}_{\mathcal{E}}. \tag{38}$$

Firstly, in regards to $\xi_j(\underline{m}) : j \in [2]$, with $\xi(\underline{m}|s^n)$ defined in (35), observe that $S = \alpha_{\underline{m}_1,\underline{b}^*} = \pi^{\mu}_{\eta_4} \pi^{1,\eta_2}_{m_{11},b_1^*} \pi^{\sigma,\eta_1}_{\underline{m},\underline{b}^*} \pi^{j,\eta_2}_{m_{11},b_1^*} \pi^{\mu}_{\eta_4} \geq 0$ is PSD implying $\left[\boldsymbol{I} - (S+T)^{-\frac{1}{2}} S (S+T)^{-\frac{1}{2}}\right] \leq I$ and hence $\xi(\underline{m}|s^n) \leq \operatorname{tr}\left(\boldsymbol{I} \cdot \rho_{\underline{m},\underline{s}^n}\right) = 1$. We therefore have $\xi_j(\underline{m}) \leq \sum_{\underline{s}^n} \sum_{i=1}^{2} \mathrm{p}^n_{\underline{S}}(\underline{s}^n) \left(\mathbb{1}_{\mathcal{F}_{ji}} + \frac{\mathbb{1}_{\mathcal{F}_{12i}}}{2}\right)$ for $j \in [2]$. Next, in regards to $\tilde{\xi}_2(\underline{m})$ defined through (38) and (35), observe that $S \geq 0, T \geq 0$ and moreover $0 \leq S \leq \boldsymbol{I}$ is dominated by $\boldsymbol{I}$. Leveraging the 'measurement on close states' [35, Exercise 9.1.8] and the Hayashi Nagaoka inequality, $\xi(\underline{m}|\underline{s}^n)$ in (35) satisfies

$$\xi(\underline{m}|\underline{s}^n) \leq \left\| \pi^{2,\eta_3}_{m_{21},b_2^*} \rho_{\underline{m},\underline{s}^n} \pi^{2,\eta_3}_{m_{21},b_2^*} - \rho_{\underline{m},\underline{s}^n} \right\|_1 + \operatorname{tr}\left(\left[\boldsymbol{I} - (S+T)^{-\frac{1}{2}} S (S+T)^{-\frac{1}{2}}\right] \pi^{2,\eta_3}_{m_{21},b_2^*} \rho_{\underline{m},\underline{s}^n} \pi^{2,\eta_3}_{m_{21},b_2^*}\right) \leq \sum_{i=3}^{7} \xi_i(\underline{m}|\underline{s}^n), \tag{39}$$

$$\text{where } \xi_3(\underline{m}|\underline{s}^n) \triangleq 3 \left\| \pi^{2,\eta_3}_{m_{21},b_2^*} \rho_{\underline{m},\underline{s}^n} \pi^{2,\eta_3}_{m_{21},b_2^*} - \rho_{\underline{m},\underline{s}^n} \right\|_1, \xi_4(\underline{m}|\underline{s}^n) \triangleq 2 \operatorname{tr}\left(\left[\boldsymbol{I} - \alpha_{\underline{m}_1,\underline{b}^*}\right] \rho_{\underline{m},\underline{s}^n}\right), \tag{40}$$

$$\xi_5(\underline{m}|\underline{s}^n) \triangleq 4 \sum_{\hat{\underline{m}}_1} \sum_{\hat{\underline{b}}} \operatorname{tr}\left(\alpha_{\hat{\underline{m}}_1,\hat{\underline{b}}} \pi^{2,\eta_3}_{m_{21},b_2^*} \rho_{\underline{m},\underline{s}^n} \pi^{2,\eta_3}_{m_{21},b_2^*}\right) \mathbb{1}_{\left\{(\hat{m}_{11},\hat{b}_1) \neq (m_{11},b_1^*)\right\}}, \tag{41}$$

$$\xi_6(\underline{m}|\underline{s}^n) \triangleq 4 \sum_{\hat{\underline{m}}_1} \sum_{\hat{\underline{b}}} \operatorname{tr}\left(\alpha_{\hat{\underline{m}}_1,\hat{\underline{b}}} \pi^{2,\eta_3}_{m_{21},b_2^*} \rho_{\underline{m},\underline{s}^n} \pi^{2,\eta_3}_{m_{21},b_2^*}\right) \mathbb{1}_{\left\{(\hat{m}_{21},\hat{b}_2) \neq (m_{21},b_2^*)\right\}}, \tag{42}$$

$$\xi_7(\underline{m}|\underline{s}^n) \triangleq 4 \sum_{\hat{\underline{m}}_1} \sum_{\hat{\underline{b}}} \operatorname{tr}\left(\alpha_{\hat{\underline{m}}_1,\hat{\underline{b}}} \pi^{2,\eta_3}_{m_{21},b_2^*} \rho_{\underline{m},\underline{s}^n} \pi^{2,\eta_3}_{m_{21},b_2^*}\right) \mathbb{1}_{\left\{(\hat{m}_{j1},\hat{b}_j) \neq (m_{j1},b_j^*) : j=1,2\right\}} \tag{43}$$

In writing the second inequality in (39), we have leveraged the fact that, when the Hayashi Nagaoka inequality is applied on the second term of the RHS of the first inequality, the resulting first term is dominated by $\xi_4(\underline{m}|\underline{s}^n) + \frac{2}{3}\xi_3(\underline{m}|\underline{s}^n)$ via the measurement on close states [35, Exercise 9.1.8]. This, in addition to the presence of the first term on the RHS of the first inequality explains the factor 3 in the definition of $\xi_3(\underline{m}|\underline{s}^n)$. The rest of the terms $\xi_i(\underline{m}|\underline{s}^n) : i = 5, 6, \cdots, 7$ are the other terms that make up the upper bound in the Hayashi Nagaoka inequality. We therefore have $\tilde{\xi}_2(\underline{m}) \leq \sum_{i=3}^{7} \xi_i(\underline{m})$, where for $i \in \{3, 4, \cdots, 7\}$, we have $\xi_i(\underline{m}) \triangleq \sum_{\underline{s}^n} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)\xi_i(\underline{m}|\underline{s}^n)\mathbb{1}_{\mathcal{E}}$. Collating through (38), (39) through (42), we have $\xi(\underline{m}) \leq \sum_{i=1}^{7} \xi_i(\underline{m})$. We now employ the random coding technique and prove that the average of these terms, evaluated over the ensemble of codes, falls exponentially to 0 if the rate conditions (27) hold. Towards that end, we specify the distribution on the ensemble of codes.

Distribution of the Random Code : We now specify the probability distribution of the random code with respect to which we compute the expectation of the eight terms mentioned above. We refer the reader to Sec. V-D for a list of components in the first and second stage that completely specify our coding scheme. It therefore suffices to specify a joint distribution of the corresponding random components : $\left(U_j^n(m_{j1}, b_j) \in \mathcal{U}_j^n : (m_{j1}, b_j) \in [\mathcal{M}_j] \times [\mathcal{B}_j]\right)$, $\left(B_j(m_{j1}, s_j^n) \in [\mathcal{B}_j] : (m_{j1}, s_j^n) \in [\mathcal{M}_j] \times \mathcal{S}_j^n\right)$ for $j = 1, 2$, the generator matrices $G_1, G_{2/1}$, the collection $(\iota_j(m_{j2}) : m_{j2} \in [q^{l_j}])$ of dither/bias vectors specifying the coset shifts, the indices $\left(A_j(\underline{m}_j, s_j^n) : (\underline{m}_j, s_j^n) \in [\mathcal{M}_j] \times \mathcal{F}_q^{l_j} \times \mathcal{S}_j^n\right)$, and the final codeword choices $\left(X_j^n(\underline{m}_j, s_j^n) : (\underline{m}_j, s_j^n) \in [\mathcal{M}_j] \times \mathcal{F}_q^{l_j} \times \mathcal{S}_j^n\right)$ each for $j = 1, 2$. The collections $\left(U_j^n(m_{j1}, b_j) \in \mathcal{U}_j^n : (m_{j1}, b_j) \in [\mathcal{M}_j] \times [\mathcal{B}_j]\right)$ for $j \in [2]$, the generator matrices $G_1, G_{2/1}$ and the dither/bias vectors $(\iota_j(m_{j2}) : m_{j2} \in [\mathcal{F}_q^{l_j}])$ are mutually independent. $G_1 \in \mathcal{F}_q^{k_1 \times n}, G_{2/1} \in \mathcal{F}_q^{(k_2-k_1)\times n}$ and $\iota_j(m_{j2}) \in \mathcal{F}_q^n : m_{j2} \in [\mathcal{F}_q^{l_j}]$ are uniformly distributed on their respective range spaces. The codewords in the collection $\left(U_j^n(m_{j1}, b_j) \in \mathcal{U}_j^n : (m_{j1}, b_j) \in [\mathcal{M}_j] \times [\mathcal{B}_j]\right)$ for $j \in [2]$ are mutually independent and $U_j^n(m_{j1}, b_j)$ is distributed with PMF $p_{U_j}^n$ for each $(m_{j1}, b_j) \in [\mathcal{M}_j] \times [\mathcal{B}_j]$. Given all of these objects, the collection of chosen indices $(B_j(m_{j1}, s_j^n) : m_{j1} \in [\mathcal{M}_j], s_j^n \in \mathcal{S}_j^n)$ are mutually independent and uniformly distributed in $\mathcal{L}_{j1}(m_{j1}, s_j^n)$. Next, given all of the above objects, the collection of chosen indices $(A_j(\underline{m}_j, s_j^n) : (\underline{m}_j, s_j^n) \in [\underline{\mathcal{M}}_j] \times \mathcal{S}_j^n)$ are mutually independent and uniformly distributed in $\mathcal{L}_{j2}(\underline{m}_j, s_j^n)$.

Remark 4. *Given the entire codebooks $(U_j^n(m_{j1}, b_j) : (m_{j1}, b_j) \in [\mathcal{M}_{j1}] \times \mathcal{B}_j) : j \in [2]$, the collection of indices $(B_j(m_{j1}, s_j^n) : (m_{j1}, s_j^n) \in [\mathcal{M}_{j1}] \times \mathcal{S}_j^n)$ are mutually independent and uniformly distributed in $\mathcal{L}_{j1}(m_{j1}, s_j^n)$. Next, given the entire codebooks $(U_j^n(m_{j1}, b_j) : (m_{j1}, b_j) \in [\mathcal{M}_{j1}] \times \mathcal{B}_j) : j \in [2]$, the generator matrices $G_1, G_{2/1}$, the collection $\iota_j(m_{j2}) : m_{j2} \in [q^{l_j}]$ of dither/bias vectors and the collection $(B_j(m_{j1}, s_j^n) : (m_{j1}, s_j^n) \in [\mathcal{M}_{j1}] \times \mathcal{S}_j^n)$, the indices $(A_j(\underline{m}_j, s_j^n) : (\underline{m}_j, s_j^n) \in [\underline{\mathcal{M}}_j] \times \mathcal{S}_j^n)$ are mutually independent and uniformly distributed in $\mathcal{L}_{j2}(\underline{m}_j, s_j^n)$.*

In the rest of our analysis of the first stage decoding, we derive upper bounds on $\overline{\xi}_i(\underline{m}) \triangleq \mathbb{E}\{\xi_i(\underline{m})\} : i \in [8]$ that decay exponentially to 0, where the expectation in question is with respect to the random code.

Upper bound on $\overline{\xi_1}(\underline{m})$, $\overline{\xi_2}(\underline{m})$: For a generic $j \in [2]$, our discussion following (38) leads us to $\overline{\xi}_j \leq \sum_{i=1}^{2}\left(\overline{\xi}_{ji} + \frac{\overline{\xi}_{12i}}{2}\right)$, where $\overline{\xi}_{ji} = \sum_{\underline{s}^n} \mathrm{p}_S^n(\underline{s}^n)P(\mathcal{F}_{ji})$ and $\overline{\xi}_{12i} = \sum_{\underline{s}^n} \mathrm{p}^n_{\underline{S}}(\underline{s}^n)P(\mathcal{F}_{12i})$ for $i \in [2]$. From classical typicality (Lemma 1), there exists a $\kappa_{j1} > 0$ such that for all n sufficiently large, $\overline{\xi}_{j1}(\underline{m}) \leq \exp\{-n\kappa_{j1}\eta_5^2\}$. Employing standard classical information theoretic analysis as presented in Appendix E, the following proposition can be proved.

Proposition 6. *If $B_j > I(U_j; S_j)_{\Upsilon} + 2\eta_5$, $\frac{k_j}{n}\log q > \log q - H(V_j|U_j, S_j)_{\Upsilon} + 3\eta_5$, then there exists $\kappa_{12} > 0$ such that for all n sufficiently large $\overline{\xi}_1(\underline{m}) + \overline{\xi}_2(\underline{m}) \leq \exp\{-n\kappa_{12}\}$.*

Upper bound on $\overline{\xi_3}(\underline{m})$: Deriving an upper bound on $\overline{\xi_3}(\underline{m})$ will essentially involve using the gentle measurement lemma for ensembles [35, Lemma 9.4.3]. We provide a proof of the following proposition in Appendix J.

Proposition 7. *For every $n \in \mathbb{N}$ sufficiently large, we have $\overline{\xi}_3(\underline{m}) \leq 96\exp\left\{-n\left(\frac{(\eta_3-\eta_5)^2}{2} - \frac{29\eta_5}{4}\right)\right\}$.*

Upper Bound on $\overline{\xi_4}(\underline{m})$: The analysis of $\overline{\xi}_4(\underline{m})$ is analogous to that of $\overline{\xi_4}(\underline{m})$ and leverages the 'measurement on closes states' [35, Exercise 9.1.8] and gentle measurement lemma for ensemble averages [35, Lemma 9.4.3]. The proof of the following proposition is fleshed out in Appendix K.

Proposition 8. *For every* $n \in \mathbb{N}$ *sufficiently large, we have* $\overline{\xi}_4(\underline{m}) \leq \exp\left\{-n\left(\eta_1-\eta_5\right)^2\right\} + 2\exp\left\{-\frac{n}{2}\left(\eta_4-\eta_5\right)^2\right\} + 2\exp\left\{-\frac{n}{2}\left(\eta_2-\eta_5\right)^2\right\}$.

Upper Bound on $\overline{\xi}_5(\underline{m})$: $\xi_5(\underline{m})$ corresponds to the error in the first sender's $\mathcal{U}_1-$message. In Appendix L, we build on the techniques developed in [37, Proof Of Thm. 2] and prove the following proposition.

Proposition 9. *For every* $n \in \mathbb{N}$ *sufficiently large, we have*

$$\overline{\xi}_5(\underline{m}) \leq \exp\left\{-n\left(I(Y;U_1|U_2)_\Upsilon - 2\eta_1 - 9\eta_5 - \eta_3 - R_{11} - B_1\right)\right\} \tag{44}$$

Upper Bound on $\overline{\xi}_6(\underline{m})$: $\xi_6(\underline{m})$ corresponds to the error in the second sender's $\mathcal{U}_2-$codebook message. In Appendix M, we build on the techniques developed in [37, Proof Of Thm. 2] and prove the following proposition.

Proposition 10. *For every* $n \in \mathbb{N}$ *sufficiently large, we have*

$$\overline{\xi}_6(\underline{m}) \leq \exp\left\{-n\left(I(Y;U_2|U_1)_\Upsilon - 2\eta_1 - 9\eta_5 - \eta_3 - R_{11} - B_1\right)\right\} \tag{45}$$

Upper Bound on $\overline{\xi}_7(\underline{m})$: Our last term in our analysis of $\hat{\xi}(\underline{m})$ in (33) is $\overline{\xi}_7(\underline{m})$. We study the same now, following which we proceed to analyzing $\hat{\zeta}(\underline{m})$. $\xi_7(\underline{m})$ corresponds to the error in the both sender's $\mathcal{U}_1-,\mathcal{U}_2-$codebook messages. In Appendix N, we build on the techniques developed in [37, Proof Of Thm. 2] and prove the following proposition.

Proposition 11. *For every* $n \in \mathbb{N}$ *sufficiently large, we have*

$$\overline{\xi}_7(\underline{m})\exp\{-n\left(I(Y;U_1,U_2)_\Upsilon - 8\eta_5 - \eta_4 - 2\eta_1 - R_{11} - R_{21} - B_1 - B_2\right)\} \tag{46}$$

Propositions 6 through 11 have characterized conditions under which $\overline{\xi}(\underline{m})$ falls exponentially in n. This concludes our analysis of $\overline{\xi}(\underline{m})$ and we now proceed to analyzing $\hat{\zeta}(\underline{m})$.

Upper bound on $\hat{\zeta}(\underline{m})$: From [35, Exercise 9.18], i.e 'measurement on close states' and the gentle measurement lemma, specifically [35, Chain of Inequalities 9.205 through to 9.209], $\hat{\zeta}(\underline{m}|\underline{s}^n)$ as defined in (34) is upper bounded via

$$\hat{\zeta}(\underline{m}|\underline{s}^n) \leq \zeta(\underline{m}|\underline{s}^n) + \left\|\rho_{\underline{m},\underline{s}^n} - \sqrt{\lambda_{\underline{m}_1}}\rho_{\underline{m},\underline{s}^n}\sqrt{\lambda_{\underline{m}_1}}\right\|_1 \leq \zeta(\underline{m}|\underline{s}^n) + 2\sqrt{\mathrm{tr}\left\{\left(\boldsymbol{I}-\lambda_{\underline{m}_1}\right)\rho_{\underline{m},\underline{s}^n}\right\}} \leq \zeta(\underline{m}|\underline{s}^n) + 2\sqrt{\hat{\xi}(\underline{m}|\underline{s}^n)}$$

$$\text{implying } \frac{\hat{\xi}(\underline{m})+\hat{\zeta}(\underline{m})}{|\mathcal{M}|} \leq \frac{\hat{\xi}(\underline{m})+\zeta(\underline{m})+2\sqrt{\hat{\xi}(\underline{m})}}{|\mathcal{M}|}$$

$$\text{where } \zeta(\underline{m}|\underline{s}^n) \triangleq \mathrm{tr}\left\{\left(\boldsymbol{I}-\theta_{\underline{m}_2|\underline{m}_1}\right)\rho_{\underline{m},\underline{s}^n}\right\}, \quad \zeta(\underline{m}) \triangleq \sum_{\underline{s}^n} \mathtt{p}_{\underline{S}}(\underline{s}^n)\zeta(\underline{m}|\underline{s}^n)$$

In view of our analysis of $\xi(\underline{m})$ which serves as an upper bound on $\hat{\xi}(\underline{m})$, we are only required to derive an upper bound on $\zeta(\underline{m})$. In view of our detailed proof of Thm. 2, the similarity of the steps herein and in the interest of brevity, we omit a detailed analysis. We put forth the following proposition which is straightforward to prove using the steps developed in the proof of Thm. 2.

Proposition 12. *If* $\frac{k_2+l_1+l_2}{n}\log q < \log q - H(W|Y,U_1,U_2)_\Upsilon$ *and* $\frac{k_j}{n}\log q > \log q - H(V_j|U_j,S_j)_\Upsilon$ *then there exits a* $\kappa_\zeta > 0$ *such that for all* $n \in \mathbb{N}$ *sufficiently large* $\overline{\zeta}(\underline{m}) \leq \exp\{-n\kappa_\zeta\}$

As was the case in proof of Thm. 2, we have assumed $k_2 \geq k_1$. In the general case, the first bound in Proposition 12 has to be replaced by $\frac{\max\{k_1,k_2\}+l_1+l_2}{n}\log q < \log q - H(W|Y,U_1,U_2)_\Upsilon$. With this, we now collate all our

findings in Propositions 6 through 12 to conclude the existence of $\kappa_\xi > 0$ such that for all $n \in \mathbb{N}$ sufficiently large, $\hat{\xi}(\underline{m}) < \exp\{-n\kappa_\xi\}$ if, for $j \in [2]$

$$\begin{array}{c} B_j > I(S_j;U_j)_\Upsilon, \quad \frac{k_j \log q}{n} > \log q - H(V_j|U_j,S_j)_\Upsilon, \sum_{j=1}^2 R_{j1} + B_j < I(U_1,U_2;Y)_\Upsilon + I(U_1;U_2)_\Upsilon, \\ R_{11} + B_1 < I(U_1;U_2,Y)_\Upsilon, \ \ R_{21} + B_2 < I(U_2;U_1,Y)_\Upsilon, \frac{\max\{k_1,k_2\}+l_1+l_2}{n} \log q < \log q - H(W|Y,U_1,U_2)_\Upsilon. \end{array} \tag{47}$$

We substitute $\frac{l_j \log q}{n} = R_j - R_{j1}$ for $j \in [2]$ and add the two non-negative bounds $R_{j1} > 0$ and $R_j - R_{j1} > 0$. Performing a Fourier Motzkin elimination on the resulting set of bounds yields the rate region stated in (27). This completes our proof.

□

VI. Communicating over Classical-Quantum Channel with Random States using UCCs

We now focus on the PTP CQ channel with classical random states available non-causally at the transmitter, abbreviated as a QSTx. See Fig. 2. We shall prove that UCCs achieve the Gel'fand-Pinsker inner bound [2], [5]. The goal of our presentation in this section is three fold. Firstly, our current proof of achievability of the Gel'fand-Pinsker inner bound for the QSTx provided by Boche, Cai and Nötzel [5] leverage tools from representation theory. While their findings are novel, the tools they leverage are unfamiliar to mainstream information theorists. It is therefore of interest to provide simple proofs of these results via conventional information-theoretic tools along the lines of [2]. Secondly, we intend to present a proof technique that works with both unstructured IID random codes and structured coset codes. Lastly, the techniques we employ in this section is identical to those employed for proving Thms. 2 and 4, thereby demonstrating the versatility of the proof techniques developed in this article.

We begin with a formal description of a QSTx. Consider a (generic) *QSTx* specified through (i) a finite input set $\mathcal{X}$, (ii) a finite set $\mathcal{S}$ of states, (iii) a PMF $\mathtt{p}_S(\cdot)$ on $\mathcal{S}$, (iii) a collection $(\rho_{xs} \in \mathcal{D}(\mathcal{H}) : (x,s) \in \mathcal{X} \times \mathcal{S})$ of density operators and (iv) cost function $\kappa : \mathcal{X} \times \mathcal{S} \to [0,\infty)$. The cost function is additive, i.e., having observed the state sequence s^n the cost incurred by the sender in preparing the state $\otimes_{t=1}^n \rho_{x_t s_t}$ is $\overline{\kappa}(x^n,s^n) \triangleq \frac{1}{n}\sum_{t=1}^n \kappa(x_t,s_t)$. Reliable communication on a QSTx entails identifying a code.

Definition 4. *An $(n, \mathcal{M}, e, \lambda)$ QSTx code consists of a message index set $\mathcal{M}$, an encoder map $e : \mathcal{M} \times \mathcal{S}^n \to \mathcal{X}^n$ with codewords denoted $(x^n(m,s^n) = (x(m,s^n)_t : 1 \le t \le n) : (m,s^n) \in \mathcal{M} \times \mathcal{S}^n)$ and a decoder POVM $\lambda \triangleq \{\lambda_m \in \mathcal{P}(\mathcal{H}^{\otimes n}) : m \in \mathcal{M}\}$. The average error probability of the code is*

$$\overline{\xi}(e,\lambda) \triangleq \frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}} \hat{\zeta}(m), \ \textit{where} \ \hat{\zeta}(m) \triangleq \sum_{s^n \in \mathcal{S}^n} \mathtt{p}_S^n(s^n) \operatorname{tr}\big([\boldsymbol{I} - \lambda_m]\, \rho_{x^n(m,s^n),s^n}\big), \rho_{x^n(m,s^n),s^n} = \bigotimes_{t=1}^n \rho_{x(m,s^n)_t,s_t}.$$

Average cost incurred by the sender in transmitting message m is $\tau(e|m) \triangleq \sum_{s^n} \mathtt{p}_S^n(s^n)\kappa(e(m,s^n),s^n)$ and the average cost incurred by the sender is $\tau(e) \triangleq \frac{1}{|\mathcal{M}|}\sum_m \tau(e|m)$.

The object of interest is the capacity region of a QSTx defined below. In this section, we prove achievability of the current known largest single-letter inner bounds to the capacity region of a QSTx.

Definition 5. *A rate-cost quadruple $(R,\tau) \in [0,\infty)^2$ is achievable if there exists a sequence of QSTx codes $(n, \mathcal{M}^{(n)}, e^{(n)}, \lambda^{(n)})$ for which $\lim_{n\to\infty} \overline{\xi}(e^{(n)},\lambda^{(n)}) = 0$,*

$$\lim_{n\to\infty} n^{-1}\log \mathcal{M}^{(n)} = R, \ \textit{and} \ \lim_{n\to\infty} \tau(e^{(n)}) \le \tau.$$

The capacity region $\mathscr{C}$ of the QSTx is the set of all achievable rate-cost vectors and $\mathscr{C}(\tau) \triangleq \{R : (R,\tau) \in \mathscr{C}\}$.

Theorem 5. *Consider a QSTx characterized through a finite set $\mathcal{S}$ of states, a PMF $\mathtt{p}_S$ on $\mathcal{S}$ modeling the distribution of the random state, an input set $\mathcal{X}$ and a collection of density operators $(\rho_{xs} \in \mathcal{D}(\mathcal{H}) : (x,s) \in \mathcal{X} \times \mathcal{S})$. For $\tau > 0$, $R \in \mathscr{C}(\tau)$ if there exists a PMF $\mathtt{p}_S p_{VX|S}$ on $\mathcal{S} \times \mathcal{V} \times \mathcal{X}$ for which $\sum_{x,s} \mathtt{p}_S(s)p_{X|S}(x|s)\kappa(x,s) \le \tau$ and $R < I(V;Y)_\Upsilon - I(V;S)_\Upsilon$ where all information quantities are computed with respect to the quantum state*

$$\sigma_{YSXV} \triangleq \sum_{x,s,v} \mathtt{p}_S(s)p_{VX|S}(v,x|s)\rho_{xs} \otimes |s\ x\ v\rangle\langle s\ x\ v|\,. \tag{48}$$

Symbol	Description	Comment
$\frac{k}{n}\log q$ and $\frac{l}{n}\log q$	Binning rate and information rate respectively.	
$g \in \mathcal{F}_q^{k\times n}$	generator matrix of the cosets that form the bins of the code	
$\mathcal{M} = q^l, [\mathcal{M}] = [q^l]$	Number of messages, message set	
$\iota(m) \in \mathcal{F}_q^n$	dither/bias vector corresponding to message $m \in [\mathcal{M}]$	
$c(m)$	$\left(v^n(a,m) = ag \oplus \iota(m) : m \in [\mathcal{M}]\right)$	Coset/Bin corresponding to message $m \in [\mathcal{M}]$
$v^n(a,m)$	$ag \oplus \iota(m) \in \mathcal{F}_q^n$	a generic codeword in $c(m)$
$\mathcal{L}(m, s^n)$	$\{a \in \mathcal{F}_q^k : (s^n, v^n(a,m)) \in T_{\eta_3}(p_{SV})\}$	List of codewords in $c(m)$ jointly typical with state s^n
$a^* = a^*_{m,s^n}$	Index of the chosen codeword $v^n(a^*_{m,s^n}, s^n)$ used to communicate message m when state sequence is s^n	
$f : \mathcal{V} \times \mathcal{S} \to \mathcal{X}$	'fusion map' to map chosen codeword and state sequence to input sequence	
$x^n(m, s^n)$	Channel Input sequence chosen to communicate message m when state sequence is s^n	
$\rho_{m,s^n} = \rho_{x^n(m,s^n),s^n}$	$\bigotimes_{t=1}^{n} \rho_{x(m,s^n)_t}$	Quantum state used to communicate message m when state sequence is s^n
μ	$\sum_{s,x} p_{SX}(s,x)\rho_{xs}$	Average Density operator
σ_v	$\sum_{s,x} p_{SX\|V}(s,x\|v)\rho_{xs}$	Average Density operator
$\pi^{\mu}_{\eta_1}$	η_1-Typical projector of μ	
$\pi^{\sigma,\eta_2}_{a,m}$	η_2-conditional typical projector of $\sigma_{v^n(a,m)}$	
$\gamma_{a,m}$	$\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{a,m}\pi^{\mu}_{\eta_1}$	
λ_m	$(\sum_{\hat{a}}\sum_{\hat{m}}\gamma_{\hat{a},\hat{m}})^{-\frac{1}{2}}\sum\gamma_{a,m}(\sum_{\hat{a}}\sum_{\hat{m}}\gamma_{\hat{a},\hat{m}})^{-\frac{1}{2}}$	Operators of decoding POVM
$\left(\begin{smallmatrix}\lambda_m : m\in[\mathcal{M}],\\ \lambda_{-1} = \boldsymbol{I} - \sum_m \lambda_m\end{smallmatrix}\right)$	Decoding POVM	
(a) $\mathcal{F}_1$, (b) $\mathcal{F}_2$, and (c) $\mathcal{F}_3$	(a) $\overline{\mathcal{E}}_1$, (b) $\mathcal{E}_1 \cap \overline{\mathcal{E}}_2$, and (c) $\mathcal{E}_1 \cap \mathcal{E}_2 \cap \overline{\mathcal{E}}_3$	Error events at Encoder
$\overline{\zeta}_4(m)$	Error Event corresponding to quantum state ρ_{m,s^n} ***not*** overlapping substantially with γ_m	
$\overline{\zeta}_5(m)$	Error Event corresponding to quantum state ρ_{m,s^n} having substantial overlap with $\gamma_{\hat{m}} : \hat{m} \neq m$	

TABLE II
DESCRIPTION OF ELEMENTS THAT CONSTITUTE THE CODING SCHEME FOR COMMUNICATION OVER QSTX

Proof. The two new elements in our proof are the code structure (Sec. VI-A). Specifically, we build a union coset code to communicate over the QSTx. Since the codewords of a random union coset code are not mutually independent and are uniformly distributed, a standard information theoretic proof is not applicable. We therefore provide detailed steps in the sequel.

A. Code Structure

Let $\mathcal{V} = \mathcal{F}_q$ be a finite field of size q. Consider a (n, k, l, g, ι) UCC whose codewords are $(v^n(a,m) \triangleq ag \oplus \iota(m) : (a,m) \in \mathcal{V}^k \times \mathcal{V}^l)$. The message index set $\mathcal{M} = [q^l]$ and the bin corresponding to message m is the collection $c(m) \triangleq (ag \oplus \iota(m) : a \in \mathcal{V}_k)$. As we describe in the sequel, the encoder observes the state sequence $s^n \in \mathcal{S}^n$ and chooses a codeword in the bin $c(m)$ indexed by the message $m \in \mathcal{M}$.

B. Encoding

For every possible pair (m, s^n) of message and state sequence, let

$$\alpha(m, s^n) \triangleq \sum_{a\in\mathcal{V}^k} \mathbb{1}_{\{(v^n(a,m),s^n)\in T^n_{\eta_3}(p_{VS})\}} \tag{49}$$

be the number of codewords in the bin $c(m)$ indexed that is jointly typical with the observed state sequence $s^n \in \mathcal{S}^n$. Let

$$\mathcal{L}(m, s^n) \triangleq \begin{cases} \{a : (v^n(a,m), s^n) \in T_{\eta_3}(p_{VS})\} & \text{if } \alpha(m, s^n) \geq 1 \\ \{0^k\} & \text{otherwise, i.e. } \alpha(m, s^n) = 0. \end{cases} \tag{50}$$

be a list of candidate code words that is available to the encoder for the message, state sequence pair (m, s^n). Let a^*_{m,s^n} be chosen from $\mathcal{L}(m, s^n)$ and $v^n(m, s^n) \triangleq v^n(a^*_{m,s^n}, m)$. A predefined 'fusion map' $f : \mathcal{S}^n \times \mathcal{V}^n \to \mathcal{X}^n$ is used to map the pair $s^n, v^n(m, s^n)$ to an input sequence in $\mathcal{X}^n$ henceforth denoted $x^n(m, s^n)$. On observing

state sequence s^n and message m, the encoder chooses $x^n(m,s^n) = (x(m,s^n)_t : 1 \le t \le n)$, and we define $\rho_{m,s^n} \triangleq \bigotimes_{t=1}^{n} \rho_{x(m,s^n)_t s_t}$.

C. Decoding POVMs

Consider a PMF $p_{SVX} = \mathtt{p}_S p_{VX|S}$ on $\mathcal{S}\times\mathcal{V}\times\mathcal{X}$. Let

$$\mu \triangleq \sum_{x,s} p_{SX}(s,x)\rho_{xs}, \sigma_v \triangleq \sum_{x,s} p_{XS|V}(x,s|v)\rho_{xs} \text{ have SCD } \mu = \sum_{y\in\mathcal{Y}} q(y)\,|f_y\rangle\langle f_y| \text{ and } \sigma_v = \sum_{y\in\mathcal{Y}} r_{Y|V}(y|v)\,|e_{y|v}\rangle\langle e_{y|v}|$$

respectively, where SCD (as specified in Sec. II) refers to spectral decomposition. Let

$$\pi_{\eta_1}^{\mu} \triangleq \sum_{y^n\in\mathcal{Y}^n} \bigotimes_{t=1}^{n} |f_{y_t}\rangle\langle f_{y_t}| \mathbb{1}_{\{y^n\in T^n_{\eta_1}(q)\}} \text{ and } \pi_{v^n}^{\sigma,\eta_2} \triangleq \begin{cases} 0 & \text{if } v^n \notin T^n_{\eta_2}(p_V) \\ \sum_{y^n\in\mathcal{Y}^n} \bigotimes_{t=1}^{n} |e_{y_t|v_t}\rangle\langle e_{y_t|v_t}| \mathbb{1}_{\{(v^n,y^n)\in T^n_{\eta_2}(p_V r_{Y|V})\}} & \text{otherwise.} \end{cases} \tag{51}$$

be the unconditional and conditional typical projectors. For $(a,m)\in\mathcal{V}^k\times\mathcal{V}^l$, let $\pi_{a,m}^{\sigma,\eta_2} \triangleq \pi_{v^n(a,m)}^{\sigma,\eta_2}$

$$\gamma_{a,m} \triangleq \pi_{\eta_1}^{\mu}\pi_{a,m}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu} \text{ and } \lambda_m \triangleq \left(\sum_{\hat{a},\hat{m}\in\mathcal{V}^k\times\mathcal{V}^l} \gamma_{\hat{a},\hat{m}}\right)^{-\frac{1}{2}} \sum_{a\in\mathcal{V}^k} \gamma_{a,m} \left(\sum_{\hat{a},\hat{m}\in\mathcal{V}^k\times\mathcal{V}^l} \gamma_{\hat{a},\hat{m}}\right)^{-\frac{1}{2}} \text{for } m\in[q^l] \text{ and } \lambda_{-1} = I_{\mathcal{H}}^{\otimes n} - \sum_{m\in\mathcal{M}} \lambda_m \tag{52}$$

and $\{\lambda_m : m\in\mathcal{M} = [q^l], \lambda_{-1}\}$ be the decoding POVM.

D. Error Probability

We employ the random coding technique to prove the existence of a code with the promised rates for which the error probability falls to 0 exponentially in the block-length n. Towards that end, observe that our code and the coding scheme is completely characterized via the following objects (i) the generator matrix $g\in\mathcal{V}^{k\times n}$, (ii) the map $\iota : \mathcal{V}^l \to \mathcal{V}^n$, (iii) the collection $(a^*_{m,s^n}\in\mathcal{V}^k : (m,s^n)\in\mathcal{M}\times\mathcal{S}^n)$ and (iv) the collection $(x^n(m,s^n)\in\mathcal{X}^n : (m,s^n)\in\mathcal{M}\times\mathcal{S}^n)$ of channel input sequences. Our first step is to characterize the error probability for a generic choice of these objects. In particular, we characterize an upper bound on this error probability composed of multiple terms. Our second step is to specify a probability distribution on the collection of codes by specifying a distribution on the aforementioned objects. In our third step, we prove that the expectation of each of the above mentioned terms falls to 0 exponentially if the rate of the code R satisfies $R < I(V;Y)_\Upsilon - I(V;S)_\Upsilon$ where the associated entropic quantities are computed with respect to state (48).

An upper bound on the error probability for a specific code : For a generic choice of the aforementioned objects, the error probability averaged over the messages is

$$\xi(e,\lambda) = \frac{1}{q^l}\sum_m \hat{\zeta}(m) \text{ where } \hat{\zeta}(m) \triangleq \sum_{s^n} \mathtt{p}_S^n(s^n)\hat{\zeta}(m|s^n) \tag{53}$$

$$\hat{\zeta}(m|s^n) \triangleq \mathrm{tr}\{(\boldsymbol{I}-\lambda_m)\rho_{m,s^n}\}, \rho_{m,s^n} \triangleq \bigotimes_{t=1}^{n} \rho_{x(m,s^n)_t s_t}$$

where $\boldsymbol{I} = I^{\otimes n}$, $\mathcal{M} = [q^l]$ and hence $|\mathcal{M}| = q^l$. We consider an arbitrary message $m\in[q^l]$ and henceforth focus our study on $\hat{\zeta}(m)$. Throughout the rest of our study of $\hat{\zeta}(m)$, we let $a^* \triangleq a^*_{m,s^n}$. With this definition and (52), note that

$$\lambda_m \ge (S+T)^{-\frac{1}{2}} S (S+T)^{-\frac{1}{2}} \text{ where } S = \gamma_{a^*,m} = \gamma_{a^*_{m,s^n},m}, \text{ and } T = \sum_{a\neq a^*} \gamma_{a,m} + \sum_{\hat{m}\neq m}\sum_{a} \gamma_{a,\hat{m}} \tag{54}$$

$$\text{and hence, } \hat{\zeta}(m|s^n) \le \zeta(m|s^n), \text{ where } \zeta(m|s^n) \triangleq \mathrm{tr}\left(\left[\boldsymbol{I} - (S+T)^{-\frac{1}{2}} S (S+T)^{-\frac{1}{2}}\right]\rho_{m,s^n}\right). \tag{55}$$

We shall henceforth focus our study on $\zeta(m) \triangleq \sum_{s^n} \mathtt{p}_S^n(s^n)\zeta(m|s^n)$ which serves as an upper bound on $\hat{\zeta}(m)$ in (53). Towards that end, we split the event corresponding to $\zeta(m)$ into two parts - $\mathcal{E}$ and $\overline{\mathcal{E}}$ - and analyze the event corresponding to the two parts separately. Towards defining $\mathcal{E}$, let

$$\mathcal{E}_1 \triangleq \left\{ s^n \in T_{\frac{\eta_3}{2}}(\mathtt{p}_S) \right\}, \mathcal{E}_2 \triangleq \{|\mathcal{L}(m,s^n)| \geq L, (s^n, v^n(m,s^n)) \in T_{\eta_3}(p_{SV})\}$$
$$\mathcal{E}_3 \triangleq \left\{ \left(s^n, v^n(m,s^n), x^n(m,s^n) \right) \in T_{2\eta_3}(p_{SVX}) \right\} \text{ and finally } \mathcal{E} \triangleq \mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3,$$

where $L \triangleq \frac{1}{2}\exp\{k\log q - n\log q + nH(V|S)_{\Upsilon} - 3n\eta_3\}$. We remark that all Von Neumann entropies in this proof are evaluated with respect to the joint state Υ^{YSVWX} specified in (48). Since

$$\overline{\mathcal{E}} = \bigcup_{i=1}^{3} \overline{\mathcal{E}}_i = \bigcup_{i=1}^{3} \mathcal{F}_i \text{ where } \mathcal{F}_1 \triangleq \overline{\mathcal{E}}_1, \mathcal{F}_2 \triangleq \mathcal{E}_1 \cap \overline{\mathcal{E}}_2, \mathcal{F}_3 \triangleq \mathcal{E}_1 \cap \mathcal{E}_2 \cap \overline{\mathcal{E}}_3, \text{ we have } 1 = \mathbb{1}_{\overline{\mathcal{E}}} + \mathbb{1}_{\mathcal{E}} \leq \sum_{i=1}^{3} \mathbb{1}_{\mathcal{F}_i} + \mathbb{1}_{\mathcal{E}} \quad (56)$$

With these definitions, we have

$$\zeta(m) \leq \sum_{i=1}^{3} \zeta_i(m) + \tilde{\zeta}_4(m), \text{ where } \zeta_i(m) \triangleq \sum_{s^n} \mathtt{p}_S^n(s^n)\zeta(m|s^n)\mathbb{1}_{\mathcal{F}_i} : i \in [3] \text{ and } \tilde{\zeta}_4(m) \triangleq \sum_{s^n} \mathtt{p}_S^n(s^n)\zeta(m|s^n)\mathbb{1}_{\mathcal{E}}. \quad (57)$$

Next, in regards to $\tilde{\zeta}_4(m)$, consider $\zeta(m|s^n)$ defined in (55). Note that since $\pi_{a,m}^{\sigma,\eta_2} \geq 0$ for every $a \in \mathcal{F}_q^k$, we have $S = \gamma_{a^*,m} = \pi_{\eta_1}^{\mu}\pi_{a^*,m}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu} \geq 0$, $T \geq 0$ are PSD. Moreover $S = \gamma_{a^*,m} = \pi_{\eta_1}^{\mu}\pi_{a^*,m}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu} \leq \pi_{\eta_1}^{\mu}\boldsymbol{I}\pi_{\eta_1}^{\mu} = \pi_{\eta_1}^{\mu} \leq \boldsymbol{I}$ implying $\boldsymbol{I} - S$ is PSD. From the Hayashi Nagaoka inequality [39], we have

$$\tilde{\zeta}_4(m) \leq \zeta_4(m) + \zeta_5(m) + \zeta_6(m) \text{ where } \zeta_4(m) \triangleq 2\sum_{s^n} \mathtt{p}_S^n(s^n)\,\mathrm{tr}\left\{ \left[\boldsymbol{I} - \pi_{\eta_1}^{\mu}\pi_{a^*,m}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu} \right] \rho_{m,s^n} \right\} \mathbb{1}_{\mathcal{E}}, \quad (58)$$

$$\zeta_5(m) \triangleq 4\sum_{s^n,\hat{a}\neq a^*} \mathtt{p}_S^n(s^n)\mathrm{tr}\left\{\pi_{\eta_1}^{\mu}\pi_{v^n(\hat{a},m)}\pi_{\eta_1}^{\mu}\rho_{m,s^n}\right\}\mathbb{1}_{\mathcal{E}} \text{ and } \zeta_6(m) \triangleq 4\sum_{s^n,\hat{a}\in\mathcal{F}_q^k}\sum_{\hat{m}\neq m} \mathtt{p}_{s^n}(s^n)\mathrm{tr}\left\{\pi_{\eta_1}^{\mu}\pi_{v^n(\hat{a},m)}\pi_{\eta_1}^{\mu}\rho_{m,s^n}\right\}\mathbb{1}_{\mathcal{E}} \quad (59)$$

Collating through (57), (58), (59), we have $\zeta(m) \leq \sum_{i=1}^{6}\zeta_i(m)$ where the terms in the latter sum are defined through (57), (58) and (59). We now employ the random coding technique and prove that the average of these terms, evaluated over the ensemble of codes, falls exponentially to 0 if the rate conditions stated in the theorem hold. Towards that end, we now specify the distribution on the ensemble of codes.

Distribution of the Random Code : The generator matrix G, the map ι and the collection $(A_{m,s^n}^* \in \mathcal{V}^k : (m,s^n) \in \mathcal{M}\times\mathcal{S}^n)$ of a random code are distributed with PMF

$$P\begin{pmatrix} G = g, \iota(\tilde{m}) = d^n(\tilde{m}) : \tilde{m} \in \mathcal{V}^l, a_{m,s^n}^* = a(m,s^n) \\ X^n(m,s^n) = x^n(m,s^n) : (m,s^n) \in \mathcal{M}\times\mathcal{S}^n \end{pmatrix} = \frac{1}{q^{kn}}\left[\prod_{\tilde{m}\in\mathcal{V}^l}\frac{1}{q^n}\right]\cdot\left[\prod_{m\in\mathcal{V}^l}\prod_{s^n\in\mathcal{S}^n}\frac{\mathbb{1}_{\{a(m,s^n)\in\mathcal{L}(m,s^n)\}}}{|\mathcal{L}(m,s^n)|}\right.$$
$$\left.\prod_{t=1}^{n} p_{X|VS}(x(m,s^n)_t|v(a(m,s^n),s^n)_t,s_t)\right]. \quad (60)$$

From (60), it can be verified that the generator matrix G and the range of $(\iota(m) : m \in \mathcal{V}^l)$ are mutually independent and uniformly distributed in the respective range spaces. Moreover, for $(m,s^n) \in \mathcal{M}\times\mathcal{S}^n$ and any $a \in \mathcal{L}(m,s^n)$, we note that

$$P\left(a_{m,s^n}^* = a(m,s^n) \,\middle|\, G = g, (\iota(\tilde{m}) = d^n(\tilde{m}) : \tilde{m} \in \mathcal{V}^l)\right) = \frac{1}{|\mathcal{L}(m,s^n)|}\mathbb{1}_{\{a(m,s^n)\in\mathcal{L}(m,s^n)\}}, \quad (61)$$

a relation we shall have opportunity to use in our analysis.

In the rest of our proof, we derive upper bounds on $\overline{\zeta}_i(m) \triangleq \mathbb{E}\{\zeta_i(m)\}$ for $i \in [6]$ that decay exponentially to 0, where the expectations in question are with respect to the distribution of the random code.

Upper Bound on $\overline{\zeta}_1(m) + \overline{\zeta}_2(m) + \overline{\zeta}_3(m)$: From (57), (55) and the definition of S in (54), we note that $S \geq 0$ is PSD and hence $\boldsymbol{I} - (S+T)^{-\frac{1}{2}}S(S+T)^{-\frac{1}{2}} \leq \boldsymbol{I}$ implying $\zeta(m|s^n) \leq \mathrm{tr}(\boldsymbol{I}\cdot\rho_{m,s^n})$. Substituting this in the definition of $\zeta_i(m)$ in (57), we obtain $\zeta_i(m) \leq \sum_{s^n}\mathtt{p}_S(s^n)\mathbb{1}_{\mathcal{F}_i}$ for $i \in [3]$. This involves only classical probabilities and our study of $\overline{\zeta}_i(m)$ will therefore closely mimic [13, Upper Bound on ϵ_{2j} in Appendix B]. The proof of the following proposition also follows directly from the proof of Proposition 1 provided in Appendix E.

Proposition 13. *If* $\frac{k\log q}{n} > \log q - H(V|S)_\Upsilon + 3\eta_3$*, then there exists a strictly positive* $\kappa_1 > 0$ *such that for all* n *sufficiently large* $\overline{\zeta}_1(m) + \overline{\zeta}_2(m) + \overline{\zeta}_3(m) \leq \exp\{-n\kappa\eta_3^2\}$

Upper bound on $\overline{\zeta}_4(m)$: From (58), we have

$$\zeta_4(m) = 2\sum_{s^n} \mathtt{p}_S^n(s^n) \operatorname{tr}\left\{\rho_{m,s^n} - \pi_{\eta_1}^\mu \pi_{a^*,m}^{\sigma,\eta_2} \pi_{\eta_1}^\mu \rho_{m,s^n}\right\} \mathbb{1}_{\mathcal{E}} = 2\sum_{s^n} \mathtt{p}_S^n(s^n) \operatorname{tr}\left\{\rho_{m,s^n} - \pi_{a^*,m}^{\sigma,\eta_2}\left[\pi_{\eta_1}^\mu \rho_{m,s^n} \pi_{\eta_1}^\mu\right]\right\} \mathbb{1}_{\mathcal{E}} \quad (62)$$

$$\leq 2\sum_{s^n} \mathtt{p}_S^n(s^n) \operatorname{tr}\left\{\rho_{m,s^n} - \pi_{a^*,m}^{\sigma,\eta_2} \rho_{m,s^n}\right\} \mathbb{1}_{\mathcal{E}} + \sum_{s^n} \mathtt{p}_S^n(s^n) \left\|\rho_{m,s^n} - \pi_{\eta_1}^\mu \rho_{m,s^n} \pi_{\eta_1}^\mu\right\|_1 \mathbb{1}_{\mathcal{E}} \leq \zeta_{41}(m) + \zeta_{42}(m), \quad (63)$$

$$\text{where } \zeta_{41}(m) \triangleq 2\sum_{s^n} \mathtt{p}_S^n(s^n) \operatorname{tr}\left\{\left[\boldsymbol{I} - \pi_{a^*,m}^{\sigma,\eta_2}\right] \rho_{m,s^n}\right\} \mathbb{1}_{\mathcal{E}} \text{ and } \zeta_{42}(m) \triangleq 2\sum_{s^n} \mathtt{p}_S^n(s^n) \sqrt{\operatorname{tr}\left\{\left[\boldsymbol{I} - \pi_{\eta_1}^\mu\right] \rho_{m,s^n}\right\}} \mathbb{1}_{\mathcal{E}}. \quad (64)$$

In the above, (62) follows from cyclicity of trace, (63) follows from 'measurement on close states' [35, Exercise 9.1.8] and (64) follows from [35, Chain of Inequalities 9.205 through to 9.209]. As an informed reader might have guessed, our analysis of $\zeta_{42}(m)$ is via an analysis analogous to the pinching lemma [35, Property 15.2.7]. In Appendix F, we have derived an upper bound on the analogous term - $\zeta_{32}(\underline{m})$ - for the QMSTx that involves a pair of encoders informed with a pair of classical states. The derivation therein can be employed here to prove the existence of a $\kappa_2 > 0$ such that, for all $n \in \mathbb{N}$ sufficiently large, $\zeta_{42}(m) \leq \exp\{-n\kappa_2\}$. This implies $\overline{\zeta}_4(m) \leq \overline{\zeta}_{41}(m) + \exp\{-n\kappa_2\}$ and we are left with $\overline{\zeta}_{41}(m), \overline{\zeta}_5(m)$ and $\overline{\zeta}_5(m)$. Referring to (64), (59) and in particular the argument within the trace, we note that the common stumbling block is to characterize the overlap between the $\pi_{v^n}^{\sigma,\eta_2}-$conditional typical projector of σ_{v^n} and ρ_{x^n,s^n}. In Appendix O we develop a sequence of steps to overcome this common stumbling block. Therein, we prove the following proposition.

Proposition 14. *For sufficiently large* $n \in \mathbb{N}$*, we have*

$$\overline{\zeta}_{41}(m) \leq \exp\{-n(2[\eta_2 - \eta_3]^2 - 5\eta_3)\},\ \overline{\zeta}_5(m) \leq \exp\{-n(\log q - H(V|Y)_\Upsilon - \frac{k}{n}\log q)\} \text{ and}$$

$$\overline{\zeta}_6(m) \leq \exp\{-n(\log q - H(V|Y)_\Upsilon - \frac{k+l}{n}\log q)\}$$

We now collate the three bounds on the rates $\frac{k}{n}\log q, \frac{l}{n}\log q$ obtained in Proposition 13 and 14. We have $\overline{\zeta}(m)$ shrinks to 0 exponentially if (i) $\frac{k}{n}\log q > \log q - H(V|S)_\Upsilon$, (ii) $\frac{k}{n}\log q < \log q - H(V|Y)_\Upsilon$ and (iii) $\frac{k+l}{n}\log q < \log q - H(V|Y)_\Upsilon$. The second bound being redundant in the face of the third bound, we have proved that any rate $R < H(V|S)_\Upsilon - H(V|Y)_\Upsilon = I(V;Y)_\Upsilon - I(V;S)_\Upsilon$ is achievable. This completed our proof. □

Remark 5. *Props. 13, 14 specify the covering and packing bounds respectively. Note that if we employ unstructured IID random codes, the covering bound would be* $\frac{k\log q}{n} > I(V;S)_\Upsilon = H(V)_\Upsilon - H(V|S)_\Upsilon = \log q - H(V|S)_\Upsilon - [\log q - H(V)_\Upsilon]$ *The first term being the lower bound in Prop. 13, we denote the excess rate - the term in* $[\cdot]-$*parenthesis - required in structured coset coding as* $\beta \triangleq [\log q - H(V)_\Upsilon]$*. Observe that if we were to employ unstructured IID random codes, the packing bound would be* $\frac{k+l}{n}\log q < I(U;Y)_\Upsilon = H(V)_\Upsilon - H(V|Y)_\Upsilon = \log q - H(V|Y)_\Upsilon - \beta$*. While structured coset codes require an excess rate of* β *for covering, we can pack exactly* $q^{n\beta}$ *times more number of bins and can therefore recover the penalty we pay in covering. Alternatively stated, the excess rate paid in covering is recovered cent-to-cent via enhanced packing rates. In essence, when we build coset codes over finite fields, we get structure for free.*

It is not clear if structure is free when we build codes that are algebraically closed with respect to operations on sets with looser structure [13, Sec. VIII]. For example, if we design codes over groups that are algebraically closed with respect to the group operation, the excess rate we pay in covering is smaller than the rebate we obtain via packing [13]. We therefore have a cost for structure. Whether this is provably inevitable remains open. As elaborated in Sec. IV-F, this motivates our design and study of UCC, as against to NCC. We refer the interested reader to [13], wherein these subtleties are better understood. Specifically [13, Sec. VIII] designs group based coding strategies for the classical MAC with distributed states.

APPENDIX A
CLASSICAL TYPICALITY AND TYPICAL PROJECTORS: DEFINITIONS AND FACTS

Here we list basic facts from classical and quantum typicality that we have used in our study. The proofs of these facts can be found in [41] and [35] respectively.

A. Classical Typicality

Suppose $\mathcal{X}_1, \cdots, \mathcal{X}_K$ are K finite sets, $p_{X_{[K]}} \triangleq p_{X_1 \cdots X_K}$ is a PMF on $\mathcal{X}_{[K]} \triangleq \mathcal{X}_1 \times \cdots \times \mathcal{X}_K$ and $X_{[K]} \triangleq (X_1, \cdots, X_K)$ denotes a random vector taking values in $\mathcal{X}_{[K]}$ with PMF $p_{X_{[K]}}$. For any $S \subseteq [K]$, we let $\mathcal{X}_S \triangleq \bigtimes_{s \in S} \mathcal{X}_s$, p_{X_S} denote the marginal of $X_S \triangleq (X_s : s \in S)$. $a_S \in \mathcal{X}_S$, $a_{[K]} \in \mathcal{X}_{[K]}$, $x_S^n \in \mathcal{X}_S^n$ and $x_{[K]}^n \in \mathcal{X}_{[K]}^n$ denote generic elements. Let $N(a_S|x_S^n) \triangleq \sum_{i=1}^n \mathbb{1}_{\{x_{S,i}=a_S\}}$ and $N(a_{[K]}|x_{[K]}^n) \triangleq \sum_{i=1}^n \mathbb{1}_{\{x_{[K],i}=a_{[K]}\}}$ denote the number of occurrences of a_S and $a_{[K]}$ in the sequences x_S^n and $x_{[K]}^n$ respectively. For any $\eta > 0$ and any $S \subseteq [K]$, we let

$$T_\eta^n(p_{X_S}) \triangleq T_\eta^n(X_S) \triangleq \left\{ x_S^n \in \mathcal{X}_S^n : \left| \frac{N(a_S|x_S^n)}{n} - p_{X_S}(a_S) \right| \leq \frac{\eta p_{X_S}(a_S)}{\log |\mathcal{X}_{[K]}|} \text{ for all } a_S \in \mathcal{X}_S \right\}. \tag{65}$$

For disjoint subsets $S, T \subseteq [K]$, $\eta > 0$ and $x_S^n \in \mathcal{X}_S^n$, we let

$$T_\eta^n(p_{X_S X_T}|x_T^n) \triangleq T_\eta^n(X_S|x_T^n) \triangleq \{x_S^n \in \mathcal{X}_S^n : (x_S^n, x_T^n) \in T_\eta(p_{X_S,X_T})\} \tag{66}$$

and $p_{X_S|X_T}$ denote the conditional PMF of X_S given X_T. Specifically, $p_{X_S|X_T}(a_S|a_T) \triangleq \frac{p_{X_S X_T}(a_S,a_T)}{p_{X_T}(a_T)}$ whenever $p_{X_T}(a_T) > 0$. In order to state the following typicality bounds, we define, for any disjoint subsets $S \subseteq [K]$, $T \subseteq [K]$,

$$\mu_{\mathcal{X}_S} \triangleq \min \{p_{X_S}(a_S) : p_{X_S}(a_S) > 0\}, \mu_{\mathcal{X}_S|\mathcal{X}_T} = \triangleq \min \left\{ \sqrt{p_{X_T}(a_T)} p_{X_S|X_T}(a_S|a_T) : p_{X_S X_T}(a_S, a_T) > 0 \right\}.$$

Remark 6, Lemmas 1 and 2 can be proved using standard typicality arguments in conjunction with the Hoeffding's lemma.

Remark 6. *(i) For any $S \subseteq [K]$, $\eta > 0$, if $x_S^n \in T_\eta^n(X_S)$ and $p_S(a_S) = 0$, then $N(a_S|x_S^n) = 0$.*
(ii) Suppose $S \subseteq R \subseteq [K]$ and $T \triangleq R \setminus S$. For any $\eta > 0$, if $x_R^n = (x_S^n, x_T^n) \in T_\eta^n(X_R) = T_\eta^n(X_S, S_T)$, then $x_S^n \in T_\eta^n(X_S)$ and $x_T^n \in T_\eta^n(X_T)$. In other words, sub-components of typical elements are typical.
(iii) For any $\eta > 0$, disjoint subsets $S \subseteq [K]$, $T \subseteq [K]$ and any $x_T^n \in \mathcal{X}_T^n$, $T_\eta^n(X_S|x_T^n) \subseteq T_\eta^n(X_S)$.
(iv) Suppose $S, T \subseteq [K]$ are disjoint and $x_T^n \notin T_\eta^n(X_T)$, then $T_\eta^n(X_S|x_T^n) = \{\}$ is empty.

Lemma 1. *(i) For any $S \subseteq [K]$ and $\eta > 0$, if $x_S^n \in T_\eta^n(X_S)$, then $\left|\frac{1}{n} \log p_{X_S}(x_S^n) + H(X_S)\right| \leq \eta$, or equivalently $\exp\{-n(H(X_S) + \eta)\} \leq p_{X_S}^n(x_S^n) \leq \exp\{-n(H(X_S) - \eta)\}$.*
(ii) For any $S \subseteq [K]$, $\eta > 0$ and any $n \in \mathbb{N}$, we have

$$P(X_S^n \notin T_\eta^n(X_S)) = \sum_{x_S^n \in \mathcal{X}_S^n} p_{X_S}^n(x_S^n) \mathbb{1}_{\left\{x_S^n \in \mathcal{X}_S^n \setminus T_\eta^n(X_S)\right\}} \leq 2|\mathcal{X}_S| \exp\left\{-\frac{2n\eta^2 \mu_{\mathcal{X}_S}^2)}{(\log |\mathcal{X}_{[K]}|)^2}\right\} \leq 2|\mathcal{X}_{[K]}| \exp\left\{-\frac{2n\eta^2 \mu_{\mathcal{X}_{[K]}}^2)}{(\log |\mathcal{X}_{[K]}|)^2}\right\}$$

(iii) For any $S \subseteq [K]$, $\eta > 0$ and any $n \in \mathbb{N}$, we have

$$(1 - 2|\mathcal{X}_S| \exp\left\{-2n\eta^2 \mu_{\mathcal{X}_S}^2 (\log |X_{[K]}|)^{-2}\right\}) \exp\{n(H(X_S) - \eta)\} \leq |T_\eta^n(X_S)| \leq \exp\{n(H(X_S) + \eta)\}.$$

In particular, if $n \geq \max\{\frac{1}{\eta} \log 2, \frac{(\log |\mathcal{X}_{[K]}|)^2 \log(4|\mathcal{X}_S|)}{2\eta^2 \mu_{\mathcal{X}_{[K]}}^2}\}$ or $n \geq \max\{\frac{1}{\eta} \log 2, \frac{(\log |\mathcal{X}_{[K]}|)^2 \log(4|\mathcal{X}_S|)}{2\eta^2 \mu_{\mathcal{X}_S}^2}\}$, we have

$$\exp\{n(H(X_S) - 2\eta)\} \leq |T_\eta^n(X_S)| \leq \exp\{n(H(X_S) + \eta)\}.$$

Lemma 2. *Suppose $S \subseteq [K]$ and $T \subseteq [K]$ are disjoint subsets, $\eta_2 > \eta_1 > 0$ and $x_T^n \in T_{\eta_1}^n(X_T)$.*
(i) If $(x_S^n, s_T^n) \in T_{\eta_2}^n(X_S, X_T)$. Then $\left|\frac{1}{n} \log p_{X_S|X_T}^n(x_S^n|x_T^n) + H(X_S|X_T)\right| \leq \eta_1 + \eta_2$ or equivalently $\exp\{-n(H(X_S|X_T) + \eta_1 + \eta_2)\} \leq p_{X_S|X_T}^n(x_S^n|x_T^n) \leq \exp\{-n(H(X_S|X_T) - \eta_1 - \eta_2)\}$.
(ii) For any $n \in \mathbb{N}$, we have

$$\begin{aligned} P(X_S^n \notin T_{\eta_2}^n(X_S|x_T^n)|X_T^n = x_t^n) &= \sum_{x_S^n \in \mathcal{X}_S^n} p_{X_S|X_T}(x_S^n|x_T^n) \mathbb{1}_{\left\{x_S^n \in \mathcal{X}_S \setminus T_{\eta_2}^n(X_S|x_T^n)\right\}} \\ &\leq 2|\mathcal{X}_{S \cup T}| \exp\left\{-2n(\eta_2 - \eta_1)^2 \delta(p_{X_S X_T}, \eta_1, |\mathcal{X}_{[K]}|)\right\} \leq 2|\mathcal{X}_{[K]}| \exp\left\{-2n(\eta_2 - \eta_1)^2 \delta(p_{X_S X_T}, \eta_1, |\mathcal{X}_{[K]}|)\right\} \end{aligned}$$

$$\text{where } \delta(p_{X_S X_T}, \eta_1, |\mathcal{X}_{[K]}|) \triangleq \frac{\mu_{\mathcal{X}_S|\mathcal{X}_T}^2 (\log |\mathcal{X}_{[K]}| - \eta_1)^2}{(\log |\mathcal{X}_{[K]}| + \eta_1)^2 (\log |\mathcal{X}_{[K]}|)^2}. \tag{67}$$

(iii) For any $n \in \mathbb{N}$, we have

$$\left(1-2|\mathcal{X}_{[K]}|\exp\left\{-2n(\eta_2-\eta_1)^2\delta(p_{X_SX_T},\eta_1,|\mathcal{X}_{[K]}|)\right\}\right)\exp\left\{n(H(X_S|X_T)-\eta_1-\eta_2)\right\} \leq |T_{\eta_2}(X_S|x_T^n)| \leq \exp\left\{n(H(X_S|X_T)+\eta_1+\eta_2)\right\}.$$

Specifically for $n \geq \max\left\{\frac{1}{\eta_1+\eta_2}\log 2, \frac{\log 4|\mathcal{X}_{[K]}|}{2(\eta_2-\eta_1)^2\delta(p_{X_SX_T},\eta_1,|\mathcal{X}_{[K]}|)}\right\}$, we have

$$\exp\left\{n(H(X_S|X_T)-2\eta_1-2\eta_2)\right\} \leq |T_{\eta_2}(X_S|x_T^n)| \leq \exp\left\{n(H(X_S|X_T)+\eta_1+\eta_2)\right\}.$$

B. Typical Projectors

Suppose $\gamma \in \mathcal{D}(\mathcal{H})$ has a spectral decomposition $\gamma = \sum_{y\in\mathcal{Y}} p_Y(y)\,|e_y\rangle\langle e_y|$ and $\eta > 0$. We define the (unconditional) $\eta-$typical projector of γ as $\pi_\eta^\gamma \triangleq \sum_{y^n\in T_\eta^n(p_Y)} \bigotimes_{t=1}^n |e_{y_i}\rangle\langle e_{y_i}|$. Our notation has suppressed the dependence of π_η^γ on n to reduce clutter. Suppose $\mathcal{W}$ is a finite set and $(\sigma_w \in \mathcal{D}(\mathcal{H}) : w \in \mathcal{W})$ is a collection of density operators, each with a spectral decomposition $\sigma_w = \sum_{y\in\mathcal{Y}} q_{Y|W}(y|w)\left|f_{y|w}\right\rangle\left\langle f_{y|w}\right|$. For a PMF p_W on $\mathcal{W}$, $w^n = (w_1,\cdots,w_n) \in \mathcal{W}^n$, $\eta > 0$, we define the $\eta-$conditional typical projector of σ_{w^n} with respect to p_W as

$$\pi_{p_W,\eta}^{\sigma_{w^n}} \triangleq \sum_{y^n\in\mathcal{Y}^n} \bigotimes_{t=1}^n \left|f_{y_i|w_i}\right\rangle\left\langle f_{y_i|w_i}\right| \mathbb{1}_{\left\{y^n\in T_\eta^n(p_W q_{Y|W}|w^n)\right\}}. \tag{68}$$

In defining $\pi_{p_W,\eta}^{\sigma_{w^n}}$, we have employed the alternate notation for the conditional typical subset as stated in (66). Most often the PMF p_W is fixed and clear from context and w^n is a codeword from a codebook requiring additional indices for its specification as in $w^n(a_j,m_j)$. When p_W is clear from context, in order to reduce clutter we let $\pi_{w^n}^{\sigma,\eta} = \pi_{p_W,\eta}^{\sigma_{w^n}}$ and $\pi_{a_j,m_j}^{\sigma,\eta} = \pi_{p_W,\eta}^{\sigma_{w^n(a_j,m_j)}}$. We define the value of the smallest strictly positive eigen value of γ as μ_γ, i.e,

$$\mu_\gamma \triangleq \min\left\{\lambda : \lambda > 0, \gamma\,|v\rangle = \lambda\,|v\rangle, |v\rangle \neq |0\rangle, \text{ the } 0 \text{ vector in } \mathcal{H}\right\}.$$

For a finite set $\mathcal{W}$, a PMF p_W on $\mathcal{W}$ and a collection $(\sigma_w : w \in \mathcal{W})$ of density operators, we define

$$\mu_{p_W,\sigma} \triangleq \min\left\{\sqrt{p_W(w)}\mu_{\sigma_w} : p_W(w)\mu_{\sigma_w} > 0, w \in \mathcal{W}\right\}.$$

The following can be proved using well established typicality arguments and the Hoeffding's inequality [42, Problem 3.18 b].[3].

Lemma 3. *Suppose $\gamma \in \mathcal{D}(\mathcal{H})$, $\eta > 0$ and $\boldsymbol{I} \in \mathcal{H}^{\otimes n}$ denotes the identity operator, then the following hold.*

(i) $\mathrm{tr}([\boldsymbol{I}-\pi_\eta^\gamma]\rho^{\otimes n}) \leq 2\cdot dim(\mathcal{H})\cdot\exp\left\{-\frac{2n\eta^2\mu_\gamma^2}{\left[\log[dim(\mathcal{H})]\right]^2}\right\}$.
(ii) $\exp\{-n(H(\gamma)+\eta)\}\pi_\eta^\gamma \leq \pi_\eta^\gamma\gamma^{\otimes n}\pi_\eta^\gamma \leq \exp\{-n(H(\gamma)-\eta)\}\pi_\eta^\gamma$.
(iii) For any $n \in \mathbb{N}$,

$$\left(1-2\cdot dim(\mathcal{H})\cdot\exp\left\{-\frac{2n\eta^2\mu_\gamma^2}{\left[\log\left[dim(\mathcal{H})\right]\right]^2}\right\}\right)\exp\left\{n(H(\gamma)-\eta)\right\} \leq \mathrm{tr}\left(\pi_\eta^\gamma\right) \leq \exp\left\{n(H(\gamma)+\eta)\right\}$$

Specifically, for $n \geq \max\{\frac{1}{\eta}\log 2, \frac{\left[\log[dim(\mathcal{H})]\right]^2\log\left(4dim(\mathcal{H})\right)}{2\eta^2\mu_\gamma^2}\}$, we have $\exp\left\{n(H(\gamma)-2\eta)\right\} \leq \mathrm{tr}(\pi_\eta^\gamma) \leq \exp\left\{n(H(\gamma)+\eta)\right\}$.

The above facts can be proven using the same sequence of steps as those used in proving [35, Property 15.1.1 - 15.1.3] and using the Hoeffding inequality for concentration. Analogous classical statements are proven in [41].

Lemma 4. *Suppose p_W is a PMF on $\mathcal{W}$ - a finite set -, $(\sigma_w \in \mathcal{D}(\mathcal{H}) : w \in \mathcal{W})$ is a collection of density operators, $\eta_2 > \eta_1 > 0$ and $\boldsymbol{I} \in \mathcal{H}^{\otimes n}$ denotes the identity operator, then the following hold.*

[3]The bound $2e^{-2\eta^2k}$ stated in [42, Problem 3.18b] is incorrect and must be replaced by $2e^{-\frac{2\eta^2}{k}}$.

(i) If $w^n \notin T^n_{\eta_2}(p_W)$*, then* $\pi^{\sigma,\eta_2}_{w^n} = \pi^{\sigma_{w^n}}_{p_W,\eta_2} = 0$.
(ii) If $w^n \in T^n_{\eta_1}(p_W)$,

$$\begin{aligned}\operatorname{tr}([\boldsymbol{I} - \pi^{\sigma,\eta_2}_{w^n}]\,\sigma_{w^n}) = \operatorname{tr}\big(\big[\boldsymbol{I} - \pi^{\sigma_{w^n}}_{p_W,\eta_2}\big]\,\sigma_{w^n}\big) &\leq 2|\log[\mathit{dim}(\mathcal{H})]\,|\exp\left\{-2n(\eta_2-\eta_1)^2\delta_q(\sigma,p_W,\eta_1)\right\}\\ \text{where } \delta_q(\sigma,p_W,\eta_1) &\triangleq \frac{\mu^2_{p_W,\sigma}(\log[\mathit{dim}(\mathcal{H})]-\eta_1)^2}{(\log[\mathit{dim}(\mathcal{H})]+\eta_1)^2(\log[\mathit{dim}(\mathcal{H})])^2}\end{aligned}$$

(iii) If $w^n \in T^n_{\eta_1}(p_W)$*, we have*

$$\exp\{-n(\sum_{w\in\mathcal{W}} p_W(w)H(\sigma_w)+\eta_1+\eta_2)\}\pi^{\sigma,\eta_2}_{w^n} \leq \pi^{\sigma,\eta_2}_{w^n}\sigma_{w^n}\pi^{\sigma,\eta_2}_{w^n} \leq \exp\{-n(\sum_{w\in\mathcal{W}} p_W(w)H(\sigma_w)-\eta_1-\eta_2)\}\pi^{\sigma,\eta_2}_{w^n}$$

(iv) If $w^n \in T^n_{\eta_1}(p_W)$*, for any* $n \in \mathbb{N}$ *we have*

$$\begin{aligned}&(1-2|\log[\mathit{dim}(\mathcal{H})]\,|\exp\left\{-2n(\eta_2-\eta_1)^2\delta_q(\sigma,p_W,\eta_1)\right\})\exp\left\{n\left(\sum_{w\in\mathcal{W}} p_W(w)H(\sigma_w)-\eta_1-\eta_2\right)\right\}\\ &\qquad\leq \operatorname{tr}(\pi^{\sigma,\eta_2}_{w^n}) = \operatorname{tr}\big(\pi^{\sigma_{w^n}}_{p_W,\eta_2}\big) \leq \exp\left\{n\left(\sum_{w\in\mathcal{W}} p_W(w)H(\sigma_w)+\eta_1+\eta_2\right)\right\}.\end{aligned}$$

In particular, for $n \geq \max\left\{\frac{1}{\eta_1+\eta_2}\log 2, \frac{\log[\mathit{dim}(\mathcal{H})]}{2(\eta_2-\eta_1)^2\delta_q(\sigma,p_W,\eta_1)}\right\}$*, we have*

$$\exp\left\{n\left(\sum_{w\in\mathcal{W}} p_W(w)H(\sigma_w)-2\eta_1-2\eta_2\right)\right\} \leq \operatorname{tr}(\pi^{\sigma,\eta_2}_{w^n}) = \operatorname{tr}\big(\pi^{\sigma_{w^n}}_{p_W,\eta_2}\big) \leq \exp\left\{n\left(\sum_{w\in\mathcal{W}} p_W(w)H(\sigma_w)+\eta_1+\eta_2\right)\right\}.$$

APPENDIX B
FORMAL DESCRIPTION OF LINEAR CODING SCHEME FOR EX. 1

We shall describe the achievability of any rate pair (R_1, R_2) satisfying $R_1 + R_2 < h_b(\tau)$. Our proof relies on the existence of a binary linear code of rate $1 - h_b(\tau)$ whose randomly chosen coset can quantize a uniform binary source within a Hamming distance of τ.

Code Structure : Let $\mathcal{X} = \{0, 1\}$ be the binary field with $\oplus$ denoting mod-2 addition. Let $k, l \in [n]$ be integers with $l \triangleq n - k$. Let $g \in \mathcal{X}^{k\times n}$ and $h \in \mathcal{X}^{l\times n}$ be the generator matrix and parity check matrix respectively, of a linear code λ of rate $\frac{k}{n} = 1 - h_b(\tau)$. We partition the l rows of h so that $h^T = [h_1^T\ h_2^T]$ into two sub-matrices with $h_j \in \mathcal{X}^{nR_j\times n}$. For $j \in [2]$, Tx j holds the collection $\big(c(m_j) : m_j \in \mathcal{X}^{nR_j}\big)$ of 2^{nR_j} cosets with $c(m_j) \triangleq \{ag \oplus m_j h_j : a \in \mathcal{X}^k\}$ for $m_j \in \mathcal{X}^{nR_j}$.

Encoding : Having observed message $M_j \in \mathcal{X}^{nR_j}$ and state sequence S_j, Tx j chooses a codeword within $c(M_j)$ that is within a Hamming distance $n\tau$ from S_j. Let $a_j^* \in \mathcal{X}^k$ be such that $w_H(a_j^* g \oplus M_j h_j \oplus S_j) \leq n\tau$, where $w_H(\cdot)$ denotes Hamming weight. With $X_j \triangleq a_j^* g \oplus M_j h_j \oplus S_j$ meeting the Hamming cost constraint, Tx j inputs the same on the channel.

Decoding : Having observed $Y^n = X_1^n \oplus S_1^n \oplus X_2^n \oplus S_2^n = a_1^* g \oplus M_1 h_1 \oplus a_2^* g \oplus M_2 h_2 = (a_1^* \oplus a_2^*)\, g \oplus Mh$ where $M = (M_1\ M_2) \in \mathcal{X}^{n(R_1+R_2)}$, the Rx declares the coset of λ in which the received vector Y^n lies. Alternatively, the Rx can compute $h^T Y^n = M = (M_1\ M_2)$ since $h^T g = 0$ and $h^T h = I_{l\times l}$.

Error Analysis : Since the channel is noiseless, the only source of error is at the Txs. So long as the there exists $a_j^* \in \mathcal{X}^k$ satisfying $w_H(a_j^* g \oplus M_j h_j \oplus S_j) \leq n\tau$ with arbitrarily high probability, the pair of messages can be communicated to the Rx with arbitrary reliability. This is ensured through the following fact whose proof can be found in [14] or can also be proven with bare hands via a simple second moment method.

Fact 3. *Suppose* $\mathcal{X} = \{0, 1\}$ *is the binary field with* $\oplus$ *denoting mod*-2 *addition,* $\tau \in (0, \frac{1}{2})$ *and* $w_H(x) = x$ *for* $x \in \{0, 1\}$ *is the Hamming weight function. Suppose* $k_n \in \mathbb{N} : n \geq 1$ *is a sequence of integers with* $k_n < n$ *satisfying* $\lim_{n\to\infty} \frac{k_n}{n} > 1 - h_b(\tau)$ *and let* $l_n = n - k_n$*. Let* $S^n \in \mathcal{X}^n$ *and* $M_n \in \mathcal{X}^{l_n}$ *be uniformly distributed and independent random vectors. For any* $\epsilon > 0$*, there exists* $N_\epsilon \in \mathbb{N}$*, such that for all* $n \geq N_\epsilon$ *there exists a linear code of rate at most* $\frac{k_n}{n} + \epsilon$ *with generator matrix* $g \in \mathcal{X}^{k_n\times n}$ *and a parity check matrix* $h_n \in \mathcal{X}^{l_n\times n}$ *such that*

$$P(\{\exists a \in \mathcal{X}^{k_n} : w_H(ag \oplus M_n h_n, S^n) < \tau + \epsilon\}) \geq 1 - \epsilon.$$

APPENDIX C
CHARACTERIZATION OF THE QUANTUM STATES IN EVALUATION OF INFORMATION QUANTITIES FOR EX. 1

Consider Ex. 1 for $\theta \in (0, \frac{\pi}{2})$. In this appendix, we provide characterization of the quantum state in (2) for the choice $\mathcal{U}_1 = \mathcal{U}_2 = \{0,1\}$, $p_{U_j|S_j}(1|0) = p_{U_j|S_j}(0|1) = \tau = 1 - p_{U_j|S_j}(0|0) = 1 - p_{U_j|S_j}(1|1)$ and $X_j = U_j \oplus S_j$ for $j \in [2]$, where $\oplus$ denotes addition mod-2. The characterizations below enable us compute the information quantities and thereby quantify the upper bound on the sum rate achievable via IID random codes. The latter is stated in our discussion prior to Sec. III-C. For the choice of parameters stated earlier, the quantum state in (2) is

$$\sigma^{YS_1S_2X_1X_2U_1U_2} = \sum_{s_1,s_2} \frac{\tau(1-\tau)}{4} \begin{bmatrix} \mathbb{1}_{\{\substack{s_1\oplus s_2\\=0}\}} |1\rangle\langle 1| + \\ \mathbb{1}_{\{\substack{s_1\oplus s_2\\=1}\}} |v_\theta\rangle\langle v_\theta| \end{bmatrix} \otimes |s_1\ s_2\rangle\langle s_1\ s_2| \otimes \begin{bmatrix} |0\ 1\ s_1\ 1\oplus s_2\rangle\langle 0\ 1\ s_1\ 1\oplus s_2| + \\ |1\ 0\ 1\oplus s_1\ s_2\rangle\langle 1\ 0\ 1\oplus s_1\ s_2| \end{bmatrix}$$
$$+ \sum_{s_1,s_2} \left[\mathbb{1}_{\{\substack{s_1\oplus s_2\\=0}\}} |0\rangle\langle 0| + \mathbb{1}_{\{\substack{s_1\oplus s_2\\=1}\}} \left|v_\theta^\perp\right\rangle\left\langle v_\theta^\perp\right| \right] \otimes |s_1\ s_2\rangle\langle s_1\ s_2| \otimes \begin{bmatrix} \frac{(1-\tau)^2}{4} |0\ 0\ s_1\ s_2\rangle\langle 0\ 0\ s_1\ s_2| + \\ \frac{\tau^2}{4} |1\ 1\ 1\oplus s_1\ 1\oplus s_2\rangle\langle 1\ 1\ 1\oplus s_1\ 1\oplus s_2| \end{bmatrix}.$$

Partial tracing over the appropriate component systems, we have

$$\sigma^{S_1S_2U_1U_2} = \sum_{s_1,s_2} \frac{\tau(1-\tau)}{4} \left(|s_1\ s_2\ 1\oplus s_1\ s_2\rangle\langle s_1\ s_2\ 1\oplus s_1\ s_2| + |s_1\ s_2\ s_1\ 1\oplus s_2\rangle\langle s_1\ s_2\ s_1\ 1\oplus s_2| \right)$$
$$+ \sum_{s_1,s_2} \frac{\tau^2}{4} |s_1\ s_2\ 1\otimes s_1\ 1\oplus s_2\rangle\langle s_1\ s_2\ 1\otimes s_1\ 1\oplus s_2| + \sum_{s_1,s_2} \frac{(1-\tau)^2}{4} |s_1\ s_2\ s_1\ s_2\rangle\langle s_1\ s_2\ s_1\ s_2| \text{ implying}$$

$$\sigma^{S_jU_j} = \sum_{s_j} \frac{\tau(1-\tau)+\tau^2}{2} |s_j\ 1\oplus s_j\rangle\langle s_j\ 1\oplus s_j| + \sum_{s_j} \frac{\tau(1-\tau)+(1-\tau)^2}{2} |s_j\ s_j\rangle\langle s_j\ s_j|$$
$$= \frac{\tau}{2} |0\ 1\rangle\langle 0\ 1| + \frac{\tau}{2} |1\ 0\rangle\langle 1\ 0| + \frac{1-\tau}{2} |0\ 0\rangle\langle 0\ 0| + \frac{1-\tau}{2} |1\ 1\rangle\langle 1\ 1| \text{ for } j \in [2] \text{ and}$$

$$\sigma^{YU_1U_2} = \sum_{s_1,s_2} \frac{\tau(1-\tau)}{4} \left[\mathbb{1}_{\{\substack{s_1\oplus s_2\\=0}\}} |1\rangle\langle 1| + \mathbb{1}_{\{\substack{s_1\oplus s_2\\=1}\}} |v_\theta\rangle\langle v_\theta| \right] \otimes \left[|s_1\ 1\oplus s_2\rangle\langle s_1\ 1\oplus s_2| + |1\oplus s_1\ s_2\rangle\langle 1\oplus s_1\ s_2| \right]$$
$$+ \sum_{s_1,s_2} \left[\mathbb{1}_{\{\substack{s_1\oplus s_2\\=0}\}} |0\rangle\langle 0| + \mathbb{1}_{\{\substack{s_1\oplus s_2\\=1}\}} \left|v_\theta^\perp\right\rangle\left\langle v_\theta^\perp\right| \right] \otimes \left[\tfrac{(1-\tau)^2}{4} |s_1\ s_2\rangle\langle s_1\ s_2| + \tfrac{\tau^2}{4} |1\oplus s_1\ 1\oplus s_2\rangle\langle 1\oplus s_1\ 1\oplus s_2| \right]$$
$$= \frac{2\tau(1-\tau)}{4} |1\rangle\langle 1| \otimes (|0\ 1\rangle\langle 0\ 1| + |1\ 0\rangle\langle 1\ 0|) + \frac{2\tau(1-\tau)}{4} |v_\theta\rangle\langle v_\theta| \otimes (|0\ 0\rangle\langle 0\ 0| + |1\ 1\rangle\langle 1\ 1|)$$
$$+ \left[\frac{(1-\tau)^2+\tau^2}{4} \right] \left[|0\rangle\langle 0| \otimes (|0\ 0\rangle\langle 0\ 0| + |1\ 1\rangle\langle 1\ 1|) + \left|v_\theta^\perp\right\rangle\left\langle v_\theta^\perp\right| (|0\ 1\rangle\langle 0\ 1| + |1\ 0\rangle\langle 1\ 0|) \right] \text{ implying}$$

$$= \frac{\left(\epsilon |1\rangle\langle 1| + (1-\epsilon) \left|v_\theta^\perp\right\rangle\left\langle v_\theta^\perp\right|\right)}{4} \otimes \begin{pmatrix} |0\ 1\rangle\langle 0\ 1| + \\ |1\ 0\rangle\langle 1\ 0| \end{pmatrix} + \frac{(\epsilon |v_\theta\rangle\langle v_\theta| + (1-\epsilon) |0\rangle\langle 0|)}{4} \otimes \begin{pmatrix} |0\ 0\rangle\langle 0\ 0| + \\ |1\ 1\rangle\langle 1\ 1| \end{pmatrix} \text{ implying}$$
$$\sigma^Y = \frac{\epsilon}{2} |1\rangle\langle 1| + \frac{(1-\epsilon)}{2} \left|v_\theta^\perp\right\rangle\left\langle v_\theta^\perp\right| + \frac{\epsilon}{2} |v_\theta\rangle\langle v_\theta| + \frac{(1-\epsilon)}{2} |0\rangle\langle 0|, \quad \sigma^{U_1U_2} = \frac{1}{4} \sum_{u_1,u_2} |u_1\ u_2\rangle\langle u_1\ u_2|$$

where $\epsilon = 2\tau(1-\tau)$.

APPENDIX D
DISTRIBUTION OF CODEWORDS IN A UNIFORMLY DISTRIBUTED RANDOM UCC

We recall the distribution of the two random UCCs that make up our coding scheme. The generator matrices $G_1 \in \mathcal{F}_q^{k_1\times n}, G_{2/1} \in \mathcal{F}_q^{(k_2-k_1)\times n}$ and the collection of dither/bias vectors $\iota_1(m_1) : m_1 \in [q^{l_1}], \iota_2(m_2) : m_2 \in [q^{l_2}]$ are mutually independent and uniformly distributed on the respective range spaces. We are led to the following.

Lemma 5. *Suppose the generator matrices $G_1 \in \mathcal{F}_q^{k_1\times n}, G_{2/1} \in \mathcal{F}_q^{(k_2-k_1)\times n}$ and the collection of dither/bias vectors $\iota_1(m_1) : m_1 \in [q^{l_1}], \iota_2(m_2) : m_2 \in [q^{l_2}]$ are mutually independent and uniformly distributed on the*

respective range spaces. Suppose $G_2 = \begin{bmatrix} G_1^T & G_{2/1}^T \end{bmatrix}^T$, $V_j(a_j, m_j) = a_j G_j \oplus i_j(m_j)$ *for* $a_j \in \mathcal{F}_q^{k_j}$ *and* $W^n(a, \underline{m}) \triangleq aG_2 \oplus \iota_1(m_1) \oplus \iota_2(m_2)$ *for* $a \in \mathcal{F}_q^{k_2}$, $m_j \in [q^{l_j}]$. *We have the following.*

1) *For any* $j = 1, 2$, *any choice of* $m_j \in \mathcal{F}_q^{l_j}$, *distinct* $a_j, \hat{a}_j \in \mathcal{F}_q^{k_j}$, *i.e.,* $a_j \neq \hat{a}_j$ *and any* $v_j^n, \hat{v}_j^n \in \mathcal{V}_j^n$, *we have* $P(V_j^n(a_j, m_j) = v_j^n, V_j^n(\hat{a}_j, m_j) = \hat{v}_j^n) = \frac{1}{q^{2n}}$. *In other words, random codewords in any bin/coset* $c_j(m_j)$ *are uniformly distributed and pairwise independent.*
2) *If* $a = a_1\ 0^{k_2-k_1} \oplus a_2$, $(v_1^n, v_2^n, w^n) \in T_\eta^n(p_{\underline{V}W})$ *where* $p_{\underline{V}W}(\underline{v}, w) = p_{\underline{V}}(v_1, v_2)\mathbb{1}_{\{w = v_1 \oplus v_2\}}$, *then* $P(V_j(a_j, m_j) = v_j^n : j \in [2], W^n(a, \underline{m}) = w^n) = \frac{1}{q^{2n}}$,
3) *If* $\hat{a} \in \mathcal{F}_q^{k_2}$ *and* $\hat{a} \neq a_1\ 0^{k_2-k_1} \oplus a_2$, *then* $P(V_j(a_j, m_j) = v_j^n : j \in [2], W^n(\hat{a}, \underline{m}) = \hat{w}^n) = \frac{1}{q^{3n}}$ *for any choice* $v_1^n, v_2^n, \hat{w}^n$ *for any choice of* $v_1^n, v_2^n, w^n \in \mathcal{F}_q^n$,
4) *If* $\underline{\hat{m}} \neq \underline{m}$, *then for any* $a_j \in \mathcal{F}_q^{k_j}$, $\hat{a} \in \mathcal{F}_q^{k_2}$ *and any choice* $v_1^n, v_2^n, w^n \in \mathcal{F}_q^n$, *we have* $P(V_j(a_j, m_j) = v_j^n : j \in [2], W^n(\hat{a}, \underline{\hat{m}}) = \hat{w}^n) = \frac{1}{q^{3n}}$.
5) *For* $j \in [2]$, *suppose* $\mathcal{V}_j = \mathcal{F}_q$ *is the finite field of size* q, $\mathcal{S}_j$ *is a finite set and* $p_{S_jV_j} \triangleq \mathtt{p}_{S_j} p_{V_j|S_j}$ *is a joint PMF on* $\mathcal{S}_j \times \mathcal{V}_j$. *For* $\eta > 0$, *let* $\alpha_j(m_j, s_j^n) \triangleq \sum_{a_j} \mathbb{1}_{\{(s_j^n, V_j^n(a_j, m_j)) \in T_{2\eta}(p_{S_jV_j})\}}$. *Then* $\mathbb{E}\{\alpha_j(m_j, s_j^n)\} \geq \exp\{n\,(k_j \log q - n \log q + H(V_j|S_j) - 4\eta)\}$ *for all* n *sufficiently large if* $s_j^n \in T_\eta^n(p_{S_j}) = T_\eta^n(\mathtt{p}_{S_j})$

Proof. Since all associated objects are uniformly distributed and mutually independent, these statements can be proved via a counting argument. Throughout, we let $a_1 = (a_{1,r} : 1 \leq r \leq k_1)$, $a_2 = (a_{2,s} : 1 \leq s \leq k_2)$, similarly $\hat{a} = (\hat{a}_s : 1 \leq s \leq k_2)$. The k_1 rows of g_1 are $g_{1,r} : 1 \leq r \leq k_1$ and the $k_2 - k_1$ rows of $g_{2/1}$ are $g_{2/1,t} : 1 \leq t \leq k_2 - k_1$. We now prove the first statement. Since $a_j \neq \hat{a}_j$, there exists $i \in \{1, 2, \cdots, k_j\}$ such that $a_{j,i} \neq \hat{a}_{j,i}$. Note that

$$
\begin{aligned}
&P\begin{pmatrix} V_j^n(a_j,m_j)=v_j^n \\ V_j^n(\hat{a}_j,m_j)=\hat{v}_j^n \end{pmatrix} = P\begin{pmatrix} a_jG_j\oplus\iota_j(m_j)=v_j^n \\ \hat{a}_jG_j\oplus\iota_j(m_j)=\hat{v}_j^n \end{pmatrix} = P\begin{pmatrix} (\hat{a}\ominus a)G_j=\hat{v}_j^n\ominus v_j^n \\ \iota_j(m_j)=v_j^n\ominus aG_j \end{pmatrix} = P\begin{pmatrix} (\hat{a}_i-a_i)G_{j,i}=\hat{v}_j^n\ominus v_j^n\ominus\sum_{l\neq i}(\hat{a}_l\ominus a_l)G_{j,l} \\ \iota_j(m_j)=v_j^n\ominus aG_j \end{pmatrix} \\
&= P\begin{pmatrix} G_{j,i}=(\hat{a}_i-a_i)^{-1}\left(\hat{v}_j^n\ominus v_j^n\ominus\sum_{l\neq i}(\hat{a}_l\ominus a_l)G_{j,l}\right) \\ \iota_j(m_j)=v_j^n\ominus aG_j \end{pmatrix} = \sum_{g_{j,l}:l\neq i} \frac{1}{q^{(k_j-1)n}} \sum_{g_{j,i},d_j^n} \frac{1}{q^{2n}} \mathbb{1}_{\left\{\begin{smallmatrix} g_{j,i}=(\hat{a}_i-a_i)^{-1}\left(\hat{v}_j^n\ominus v_j^n\ominus\sum_{l\neq i}(\hat{a}_l\ominus a_l)g_{j,l}\right) \\ \iota_j(m_j)=v_j^n\ominus ag_j \end{smallmatrix}\right\}} \qquad (69) \\
&= \sum_{g_{j,l}:l\neq i} \frac{1}{q^{(k_j-1)n}} \frac{1}{q^{2n}} = \frac{1}{q^{2n}},
\end{aligned}
$$

where the summation over $g_{j,i}, d_j^n$ in (69) vanishes, because for any choice of $g_{j,l} : l = 1, \cdots, i-1, i+1, \cdots, k_j$ the indicator function therein is non-zero for a unique choice of $g_{j,i}$ and d_j^n. This proves the first statement and we now consider the second statement. We have $w^n = v_1^n \oplus v_2^n$ and $a = a_1\ 0^{k_2-k_1} \oplus a_2$. Observe that

$$
\begin{aligned}
&P\begin{pmatrix} V_j(a_j,m_j)=v_j^n : j\in[2] \\ W^n(a,\underline{m})=w^n \end{pmatrix} = P\begin{pmatrix} a_1G_1\oplus\iota_1(m_1)=v_1^n, a_2G_2\oplus\iota_2(m_2)=v_2^n \\ aG_2\oplus\iota_1(m_1)\oplus\iota_2(m_2)=w^n \end{pmatrix} = P\begin{pmatrix} a_1G_1\oplus\iota_1(m_1)=v_1^n, a_2G_2\oplus\iota_2(m_2)=v_2^n \\ [a_1\ 0^{k_2-k_1}\oplus a_2]G_2\oplus\iota_1(m_1)\oplus\iota_2(m_2)=v_1^n\oplus v_2^n \end{pmatrix} \\
&= P\begin{pmatrix} a_1G_1\oplus\iota_1(m_1)=v_1^n, a_2G_2\oplus\iota_2(m_2)=v_2^n \\ a_1G_1\oplus a_2G_2\oplus\iota_1(m_1)\oplus\iota_2(m_2)=v_1^n\oplus v_2^n \end{pmatrix} = P\begin{pmatrix} a_1G_1\oplus\iota_1(m_1)=v_1^n \\ a_2G_2\oplus\iota_2(m_2)=v_2^n \end{pmatrix} = P\begin{pmatrix} \iota_1(m_1)=v_1^n\ominus\sum_{r=1}^{k_1} a_{1,r}G_{1,r}, \\ \iota_2(m_2)=v_2^n\ominus\sum_{r=1}^{k_1} a_{2,r}G_{1,r}\ominus\sum_{s=1}^{k_2-k_1} a_{2,k_1+s}G_{2/1,s} \end{pmatrix} \\
&= \sum_{g_1\in\mathcal{F}_q^{k_1\times n}} \sum_{g_{2/1}\in\mathcal{F}_q^{(k_2-k_1)\times n}} P\begin{pmatrix} G_1=g_1, \\ G_{2/1}=g_{2/1} \end{pmatrix} \sum_{d_1^n,d_2^n} P\begin{pmatrix} \iota_1(m_1)=d_1^n, \\ \iota_2(m_2)=d_2^n \end{pmatrix} \mathbb{1}_{\left\{\begin{smallmatrix} d_1^n=v_1^n-a_1g_1 \\ d_2^n=v_2^n\ominus\sum_{r=1}^{k_1} a_{2,r}g_{1,r}\ominus\sum_{s=1}^{k_2-k_1} a_{2,k_1+s}g_{2/1,s} \end{smallmatrix}\right\}} \\
&= \sum_{g_1\in\mathcal{F}_q^{k_1\times n}} \sum_{g_{2/1}\in\mathcal{F}_q^{(k_2-k_1)\times n}} \frac{1}{q^{k_1n}} \frac{1}{q^{(k_2-k_1)n}} \sum_{d_1^n,d_2^n} \frac{1}{q^n} \frac{1}{q^n} \mathbb{1}_{\left\{\begin{smallmatrix} d_1^n=v_1^n-a_1g_1 \\ d_2^n=v_2^n\ominus\sum_{r=1}^{k_1} a_{2,r}g_{1,r}\ominus\sum_{s=1}^{k_2-k_1} a_{2,k_1+s}g_{2/1,s} \end{smallmatrix}\right\}} \qquad (70) \\
&= \sum_{g_1\in\mathcal{F}_q^{k_1\times n}} \sum_{g_{2/1}\in\mathcal{F}_q^{(k_2-k_1)\times n}} \frac{1}{q^{k_1n}} \frac{1}{q^{(k_2-k_1)n}} \frac{1}{q^{2n}} = \frac{1}{q^{2n}},
\end{aligned}
$$

where the summation over d_1^n, d_2^n in (70) vanishes because for any choice of $g_1, g_{2/1}$ the indicator $\mathbb{1}_\alpha$ is non-zero for a unique choice of d_1^n and d_2^n. We now prove the third statement. We have $\hat{a} \neq a_1\ 0^{k_2-k_1} \oplus a_2$. Either (i) there exists an $i \in \{1, \cdots, k_1\}$ such that $\hat{a}_i \neq a_{1i} \oplus a_{2i}$, or (ii) there exists $i \in \{k_1 + 1, \cdots, k_2\}$ such that $\hat{a}_i \neq a_{2i}$. Suppose (i) holds, then

$$P\begin{pmatrix}V_j(a_j,m_j)=v_j^n:j\in[2]\\ W^n(\hat{a},\underline{m})=\hat{w}^n\end{pmatrix}=P\begin{pmatrix}a_1G_1\oplus\iota_1(m_1)=v_1^n,a_2G_2\oplus\iota_2(m_2)=v_2^n\\ \hat{a}G_2\oplus\iota_1(m_1)\oplus\iota_2(m_2)=\hat{w}^n\end{pmatrix}$$
$$=P\begin{pmatrix}\sum_{r=1}^{k_1}a_{1,r}G_{1,r}\oplus\iota_1(m_1)=v_1^n,\ \sum_{r=1}^{k_1}a_{2,r}G_{1,r}\oplus\sum_{s=1}^{k_2-k_1}a_{2,k_1+s}G_{2/1,s}\oplus\\ \iota_2(m_2)=v_2^n,\sum_{r=1}^{k_1}\hat{a}_rG_{1,r}\oplus\sum_{s=1}^{k_2-k_1}\hat{a}_{k_1+s}G_{2/1,s}\oplus\iota_1(m_1)\oplus\iota_2(m_2)=\hat{w}^n\end{pmatrix}$$
$$=P\begin{pmatrix}(\hat{a}_i\ominus a_{1i}\ominus a_{2i})G_{1i}=\hat{w}^n\ominus v_1^n\ominus v_2^n\ominus\sum_{r\neq i}(\hat{a}_r\ominus a_{1r}\ominus a_{2r})G_{1,r}\ominus\sum_{s=1}^{k_2-k_1}(\hat{a}_{k_1+s}\ominus a_{2,k_1+s})G_{2/1,s},\\ \iota_1(m_1)=v_1^n\ominus\sum_{r=1}^{k_1}a_{1,r}G_{1,r},\ \iota_2(m_2)=v_2^n\ominus\sum_{r=1}^{k_1}a_{2,r}G_{1,r}\ominus\sum_{s=1}^{k_2-k_1}a_{2,k_1+s}G_{2/1,s},\end{pmatrix}$$
$$=\sum_{g_{1r}:r\neq i}\sum_{g_{2/1}}\frac{1}{q^{(k_1-1)n}}\frac{1}{q^{(k_2-k_1)n}}\sum_{d_1^n,d_2^n}\sum_{g_{1,i}}\frac{1}{q^{3n}}\mathbb{1}_\alpha=\sum_{g_{1r}:r\neq i}\sum_{g_{2/1}}\frac{1}{q^{(k_1-1)n}}\frac{1}{q^{(k_2-k_1)n}}\frac{1}{q^{3n}}=\frac{1}{q^{3n}}\ \text{where}\quad(71)$$
$$\alpha\triangleq\left\{\begin{matrix}g_{1i}=(\hat{a}_i\ominus a_{1i}\ominus a_{2i})^{-1}\big(\hat{w}^n\ominus v_1^n\ominus v_2^n\ominus\sum_{r\neq i}(\hat{a}_r\ominus a_{1r}\ominus a_{2r})g_{1,r}\ominus\sum_{s=1}^{k_2-k_1}(\hat{a}_{k_1+s}\ominus a_{2,k_1+s})g_{2/1,s}\big)\\ \iota_1(m_1)=v_1^n\ominus\sum_{r=1}^{k_1}a_{1,r}g_{1,r},\ \iota_2(m_2)=v_2^n\ominus\sum_{r=1}^{k_1}a_{2,r}g_{1,r}\ominus\sum_{s=1}^{k_2-k_1}a_{2,k_1+s}g_{2/1,s}\end{matrix}\right\},$$

where the summation over $d_1^n,d_2^n,g_{1,i}$ in (71) vanishes because for any choice of $g_{1,r}:r\neq i,g_{2/1}$ the indicator $\mathbb{1}_\alpha$ is non-zero for a unique choice of $d_1^n,d_2^n,g_{1,i}$. Suppose (ii) holds, i.e., there exists $i\in\{k_1+1,\cdots,k_2\}$ such that $\hat{a}_i\neq a_{2i}$, then

$$P\begin{pmatrix}V_j(a_j,m_j)=v_j^n:j\in[2]\\ W^n(\hat{a},\underline{m})=\hat{w}^n\end{pmatrix}=P\begin{pmatrix}a_1G_1\oplus\iota_1(m_1)=v_1^n,a_2G_2\oplus\iota_2(m_2)=v_2^n\\ \hat{a}G_2\oplus\iota_1(m_1)\oplus\iota_2(m_2)=\hat{w}^n\end{pmatrix}$$
$$=P\begin{pmatrix}\sum_{r=1}^{k_1}a_{1,r}G_{1,r}\oplus\iota_1(m_1)=v_1^n,\ \sum_{r=1}^{k_1}a_{2,r}G_{1,r}\oplus\sum_{s=1}^{k_2-k_1}a_{2,k_1+s}G_{2/1,s}\oplus\\ \iota_2(m_2)=v_2^n,\sum_{r=1}^{k_1}\hat{a}_rG_{1,r}\oplus\sum_{s=1}^{k_2-k_1}\hat{a}_{k_1+s}G_{2/1,s}\oplus\iota_1(m_1)\oplus\iota_2(m_2)=\hat{w}^n\end{pmatrix}$$
$$=P\begin{pmatrix}(\hat{a}_i\ominus a_{2i})G_{2/1,i}=\hat{w}^n\ominus v_1^n\ominus v_2^n\ominus\sum_{r=1}^{k_1}(\hat{a}_r\ominus a_{1r}\ominus a_{2r})G_{1,r}\ominus\sum_{s\neq i}(\hat{a}_{k_1+s}\ominus a_{2,k_1+s})G_{2/1,s},\\ \iota_1(m_1)=v_1^n\ominus\sum_{r=1}^{k_1}a_{1,r}G_{1,r},\ \iota_2(m_2)=v_2^n\ominus\sum_{r=1}^{k_1}a_{2,r}G_{1,r}\ominus\sum_{s=1}^{k_2-k_1}a_{2,k_1+s}G_{2/1,s},\end{pmatrix}$$
$$=\sum_{g_1}\sum_{g_{2/1,s}:s\neq i}\frac{1}{q^{k_1n}}\frac{1}{q^{(k_2-k_1-1)n}}\sum_{d_1^n,d_2^n}\sum_{g_{2/1,i}}\frac{1}{q^{3n}}\mathbb{1}_\beta=\sum_{g_1}\sum_{g_{2/1,s}:s\neq i}\frac{1}{q^{k_1n}}\frac{1}{q^{(k_2-k_1-1)n}}\frac{1}{q^{3n}}=\frac{1}{q^{3n}}\ \text{where}\quad(72)$$
$$\beta\triangleq\left\{\begin{matrix}g_{1i}=(\hat{a}_i\ominus a_{2i})^{-1}\big(\hat{w}^n\ominus v_1^n\ominus v_2^n\ominus\sum_{r=1}^{k_1}(\hat{a}_r\ominus a_{1r}\ominus a_{2r})g_{1,r}\ominus\sum_{s\neq i}(\hat{a}_{k_1+s}\ominus a_{2,k_1+s})g_{2/1,s}\big)\\ \iota_1(m_1)=v_1^n\ominus\sum_{r=1}^{k_1}a_{1,r}g_{1,r},\ \iota_2(m_2)=v_2^n\ominus\sum_{r=1}^{k_1}a_{2,r}g_{1,r}\ominus\sum_{s=1}^{k_2-k_1}a_{2,k_1+s}g_{2/1,s}\end{matrix}\right\}.$$

where the summation over $d_1^n,d_2^n,g_{2/1,i}$ in (72) vanishes because for any choice of $g_1:r\neq i,g_{2/1,s}:s\neq i$ the indicator $\mathbb{1}_\beta$ is non-zero for a unique choice of $d_1^n,d_2^n,g_{2/1,i}$. Lastly, we prove the fourth and last statement. Suppose $j\in[2]$ such that $\hat{m}_j\neq m_j$ and $\dot{j}\in\{1,2\}\setminus\{j\}$ is the complement index. Then

$$P\begin{pmatrix}V_j(a_j,m_j)=v_j^n:j\in[2]\\ W^n(\hat{a},\underline{\hat{m}})=\hat{w}^n\end{pmatrix}=P\begin{pmatrix}a_1G_1\oplus\iota_1(m_1)=v_1^n,a_2G_2\oplus\iota_2(m_2)=v_2^n\\ \hat{a}G_2\oplus\iota_j(\hat{m}_j)\oplus\iota_{\dot{j}}(\hat{m}_{\dot{j}})=\hat{w}^n\end{pmatrix}=P\begin{pmatrix}\iota_1(m_1)=v_1^n\ominus a_1G_1,\iota_2(m_2)=v_2^n\ominus a_2G_2,\\ \iota_j(\hat{m}_j)=\hat{w}^n\ominus\iota_{\dot{j}}(\hat{m}_{\dot{j}})\ominus\hat{a}G_2\end{pmatrix}$$
$$=\sum_{g_1}\sum_{g_{2/1}}\sum_{\hat{d}_{\dot{j}}^n}\frac{1}{q^{k_1n}}\frac{1}{q^{n(k_2-k_1)}}\frac{1}{q^n}\sum_{d_1^n,d_2^n,\hat{d}_j^n}\frac{1}{q^{3n}}\mathbb{1}_{\left\{\begin{matrix}d_1^n=v_1^n\ominus a_1g_1,d_2^n=v_2^n\ominus a_2g_2,\\ \hat{d}_j^n=\hat{w}^n\ominus d_{\dot{j}}^n\ominus\hat{a}g_2\end{matrix}\right\}}=\sum_{g_1}\sum_{g_{2/1}}\sum_{\hat{d}_{\dot{j}}^n}\frac{1}{q^{n(k_2+1)}}\frac{1}{q^{3n}}=\frac{1}{q^{3n}}.\quad(73)$$

Lastly we prove the fifth statement which, in light of the uniform distribution stated in the first statement, is a plain computation.

$$\mathbb{E}\{\alpha_j(m_j,s_j^n)\}=\sum_{a_j}P\left((s_j^n,V_j^n(a_j,m_j))\in T_{2\eta}(p_{S_jV_j})\right)=\sum_{a_j}\sum_{v_j^n}P\left(V_j^n(a_j,m_j)=v_j^n\right)\mathbb{1}_{\{v_j^n\in T_{2\eta}(p_{S_jV_j}|s_j^n)\}}$$
$$=\sum_{a_j}\sum_{v_j^n}\frac{1}{q^n}\mathbb{1}_{\{v_j^n\in T_{2\eta}(p_{S_jV_j}|s_j^n)\}}=\frac{q^{k_j}}{q^n}\left|T_{2\eta}(p_{S_jV_j}|s_j^n)\right|\geq\exp\left\{n\left(k_j\log q-n\log q+H(V_j|S_j)-4\eta\right)\right\}\quad(74)$$

where the last equality in (74) follows from the uniform distribution of the codewords proven in the first statement and the inequality follows from bounds on the size of the conditional typical set. □

APPENDIX E
PROOF OF PROPOSITION 1

We are required to derive an upper bound on $\sum_{\underline{s}^n}\mathrm{p}_{\underline{S}}(\underline{s}^n)\sum_{i=1}^2\left(\mathbb{1}_{\mathcal{F}_{ji}}+\frac{\mathbb{1}_{\mathcal{F}_{12i}}}{2}\right)$. For $i,j\in[2]$, let $\zeta_{ji}(\underline{m})\triangleq\sum_{\underline{s}^n}\mathrm{p}_{\underline{S}}(\underline{s}^n)\mathbb{1}_{\mathcal{F}_{ji}}$ and for $j\in[2]$, let $\zeta_{j3}(\underline{m})\triangleq\sum_{\underline{s}^n}\mathrm{p}_{\underline{S}}(\underline{s}^n)\sum_{i=1}^2\frac{\mathbb{1}_{\mathcal{F}_{12i}}}{2}$ where $\mathcal{F}_{ji},\mathcal{F}_{12i}:j,i\in[2]$ is as defined in (11), (12) and let $\overline{\zeta}_{ji}=\mathbb{E}\{\zeta_{ji}(\underline{m})\}$ for $i\in[3],j\in[2]$, where the expectation is with respect to the random code. From Lemma 1, we have $\overline{\zeta}_{j1}\leq 2|\mathcal{S}_j|\exp\left\{-n\eta_3^2\mu_{\mathcal{S}_j}^2(2\log|\underline{\mathcal{S}}\times\underline{\mathcal{V}}\times\mathcal{W}\times\underline{\mathcal{X}}|)^{-2}\right\}$. We

now focus on $\overline{\zeta}_{j2}(\underline{m})$. Recalling $\mathcal{E}_{j1}$ and $\mathcal{E}_{j2}$, note that the encoding rule ensures $\left\{|\mathcal{L}_j(m_j,s_j^n)|\geq L_j\right\}\subseteq\left\{(s_j^n,v_j^n(m_j,s_j^n))\in T_{\eta_3}^n(p_{S_jV_j})\right\}$, hence $\left\{(s_j^n,v_j^n(m_j,s_j^n))\notin T_{\eta_3}^n(p_{S_jV_j})\right\}\subseteq\left\{|\mathcal{L}_j(m_j,s_j^n)|<L_j\right\}$. From this, it is evident that $\mathcal{F}_{j2}\subseteq\left\{s_j^n\in T_{\frac{\eta_3}{2}}^n(\mathtt{p}_{S_j}),|\mathcal{L}_j(m_j,s_j^n)|<L_j\right\}$. From (11), the preceding definition of $\mathcal{E}_{j2}$ therein, the definitions of $\mathcal{L}_j(m_j,s_j^n)$ and $\alpha_j(m_j,s_j^n)$, we have $\alpha_j(m_j,s_j^n)\leq|\mathcal{L}_j(m_j,s_j^n)|$ and hence $\mathcal{F}_{j2}\subseteq\left\{s_j^n\in T_{\frac{\eta_3}{2}}^n(\mathtt{p}_{S_j}),|\mathcal{L}_j(m_j,s_j^n)|<L_j\right\}\subseteq\left\{s_j^n\in T_{\frac{\eta_3}{2}}^n(\mathtt{p}_{S_j}),\alpha_j(m_j,s_j^n)<L_j\right\}$. Thus,

$$\overline{\zeta}_{j2}\leq\mathbb{E}\left\{\sum_{\underline{s}^n}\mathtt{p}_{\underline{S}}(\underline{s}^n)\mathbb{1}_{\left\{s_j^n\in T_{\frac{\eta_3}{2}}^n(\mathtt{p}_{S_j}),\alpha_j(m_j,s_j^n)<L_j\right\}}\right\}\tag{75}$$

and it suffices to derive an upper bound on the RHS of the above inequality. We now compute $\mathbb{E}\{\alpha_j(m_j,s_j^n)\}$ to unravel its relation to L_j. Recall that L_j was defined as $\frac{1}{2}\exp\{k_j\log q-n\log q+nH(V_j|S_j)_\Upsilon-3n\eta_3\}$ prior to (11) and observe that, whenever $s_j^n\in T_{\frac{\eta_3}{2}}(\mathtt{p}_{S_j})$, we have

$$\mathbb{E}\{\alpha_j(m_j,s_j^n)\}=\sum_{a_j}P((V_j^n(a_j,m_j),s_j^n)\in T_{\eta_3}^n(p_{S_jV_j}))=\sum_{a_j}\sum_{v_j^n}\mathbb{1}_{\left\{(s_j^n,v_j^n)\in T_{\eta_3}(p_{S_jV_j})\right\}}P(V_j^n(a_j,m_j)=v_j^n)$$

$$=\sum_{a_j}\sum_{v_j^n}\mathbb{1}_{\left\{(s_j^n,v_j^n)\in T_{\eta_3}(p_{S_jV_j})\right\}}P(a_jG_j\oplus\iota_j(m_j)=v_j^n)=\sum_{a_j}\sum_{v_j^n}\mathbb{1}_{\left\{(s_j^n,v_j^n)\in T_{\eta_3}(p_{S_jV_j})\right\}}\sum_{g_j\in\mathcal{F}_q^{k_j\times n}}\sum_{d_j^n\in\mathcal{F}_q^n}\frac{\mathbb{1}_{\left\{a_jg_j\oplus d_j^n=v_j^n\right\}}}{q^{k_jn}\cdot q^n}$$

$$=\sum_{a_j}\sum_{v_j^n}\mathbb{1}_{\left\{(s_j^n,v_j^n)\in T_{\eta_3}(p_{S_jV_j})\right\}}\sum_{g_j\in\mathcal{F}_q^{k_j\times n}}\sum_{d_j^n}\mathbb{1}_{\left\{d_j^n=v_j^n\ominus a_jg_j\right\}}\frac{1}{q^{k_jn}}\frac{1}{q^n}=\sum_{a_j}\sum_{v_j^n}\mathbb{1}_{\left\{(s_j^n,v_j^n)\in T_{\eta_3}(p_{S_jV_j})\right\}}\sum_{g_j\in\mathcal{F}_q^{k_j\times n}}\frac{1}{q^{k_jn}}\frac{1}{q^n}\tag{76}$$

$$=\sum_{a_j}\sum_{v_j^n}\frac{\mathbb{1}_{\left\{(s_j^n,v_j^n)\in T_{\eta_3}(p_{S_jV_j})\right\}}}{q^n}=q^{k_j}\frac{|T_{\eta_3}(V_j|s_j^n)|}{q^n}\geq\exp\{k_j\log q-n\log q+nH(V_j|S_j)_\Upsilon-3n\eta_3\}=2L_j.\tag{77}$$

where (i) the summation over d_j^n in (76) vanishes, because for any choice of $g_j\in\mathcal{F}_q^{k_j\times n}$ the indicator function $\mathbb{1}_{\left\{d_j^n=v_j^n\ominus a_jg_j\right\}}$ therein is non-zero for a unique choice of d_j^n and (ii) the inequality in (77) holds so long as $n\geq\max\left\{\frac{2}{3\eta_3}\log 2,\frac{\log[4|\underline{\mathcal{S}}\times\underline{\mathcal{V}}\times\mathcal{W}\times\underline{\mathcal{X}}|]}{2(\eta_2-\eta_1)^2\delta(p_{V_jS_j},\eta_1,|\underline{\mathcal{S}}\times\underline{\mathcal{V}}\times\mathcal{W}\times\underline{\mathcal{X}}|)}\right\}$. For sufficiently large n, we therefore have

$$\overline{\zeta}_{j2}(\underline{m})=\mathbb{E}\left\{\sum_{\underline{s}^n}\mathtt{p}_{\underline{S}}(\underline{s}^n)\mathbb{1}_{\left\{s_j^n\in T_{\frac{\eta_3}{2}}^n(\mathtt{p}_{S_j}),\alpha_j(m_j,s_j^n)<L_j\right\}}\right\}\leq\sum_{s_j^n\in T_{\frac{\eta_3}{2}}^n(\mathtt{p}_{S_j})}\mathtt{p}_{S_j}^n(s_j^n)P\left(\alpha_j(m_j,s_j^n)<\frac{\mathbb{E}\{\alpha_j(m_j,s_j^n)\}}{2}\right)$$

$$\leq\sum_{s_j^n\in T_{\frac{\eta_3}{2}}^n(\mathtt{p}_{S_j})}\mathtt{p}_{S_j}(s_j^n)P\left(|\alpha_j(m_j,s_j^n)-\mathbb{E}\left\{\alpha_j(m_j,s_j^n)\right\}|>\frac{\mathbb{E}\left\{\alpha_j(m_j,s_j^n)\right\}}{2}\right)\leq\frac{4\mathrm{Var}(\alpha_j(m_j,s_j^n))}{\left(\mathbb{E}\left\{\alpha_j(m_j,s_j^n)\right\}\right)^2}\tag{78}$$

where $\mathrm{Var}(\alpha_j(m_j,s_j^n))$ denotes the variance of $\alpha_j(m_j,s_j^n)=\sum_{a_j}\mathbb{1}_{\{V_j^n(a_j,m_j)\in T_{\eta_3}^n(V_j|s_j^n)\}}$. In writing the above set of inequalities, we have use the fact that the random codebook is independent of the observed state and the Cheybyshev inequality.

In Lemma 5, we have proved that the codewords in any bin $c_j(m_j)=(V_j^n(a_j,m_j):a_j\in\mathcal{F}_q^{k_j})$ are uniformly distributed and pairwise independent. $\alpha_j(m_j,s_j^n)$ is therefore a sum of q^{k_j} pairwise independent indicator random variables, each of which takes the value 1 with probability $q^{-n}|T_{\eta_3}(V_j|s_j^n)|$. The variance of a sum of pairwise independent indicator random variables is dominated by its mean and we therefore have

$$\frac{4\mathrm{Var}(\alpha_j(m_j,s_j^n))}{\left(\mathbb{E}\left\{\alpha_j(m_j,s_j^n)\right\}\right)^2}\leq\frac{4}{\mathbb{E}\left\{\alpha_j(m_j,s_j^n)\right\}}\leq4\exp\left\{-n\left[\frac{k_j}{n}\log q-(\log q-H(V_j|S_j)_\Upsilon+3\eta_3)\right]\right\},\tag{79}$$

where we have used the lower bound on $\mathbb{E}\left\{\alpha_j(m_j,s_j^n)\right\}$ derived in (77). Substituting (79) in (78) and the earlier stated bound on $\overline{\zeta}_{j1}(\underline{m})$, we have

$$\sum_{i=1}^2\overline{\zeta}_{ji}(\underline{m})\leq2|\mathcal{S}_j|\exp\left\{-n\eta_3^2\mu_{\mathcal{S}_j}^2(2\log|\underline{\mathcal{S}}||\underline{\mathcal{V}}||\mathcal{W}||\underline{\mathcal{X}}|)^{-2}\right\}+4\exp\left\{-n\left[\frac{k_j}{n}\log q-(\log q-H(V_j|S_j)_\Upsilon+3\eta_3)\right]\right\}.\tag{80}$$

We are thus left to bound $\overline{\zeta}_{j3}(\underline{m}) \triangleq \mathbb{E}\{\zeta_{j3}(\underline{m})\}$ on the above, where we recall $\zeta_{j3}(\underline{m}) = \sum_{\underline{s}^n} \mathrm{p}_{\underline{S}}(\underline{s}^n) \sum_{i=1}^{2} \frac{\mathbb{1}_{\mathcal{F}_{12i}}}{2}$ and $\mathcal{F}_{121}, \mathcal{F}_{122}$ are as defined in 12. The analysis of $\mathbb{E}\{\sum_{\underline{s}^n} \mathrm{p}_{\underline{S}}(\underline{s}^n)\mathbb{1}_{\mathcal{F}_{121}}\}$ and $\mathbb{E}\{\sum_{\underline{s}^n} \mathrm{p}_{\underline{S}}(\underline{s}^n)\mathbb{1}_{\mathcal{F}_{122}}\}$ are identical to the analysis of ϵ_3 and ϵ_4 in the proof of [13, Theorem 4]. The analysis of the latter terms are detailed and an upper bound on the same are derived in [13, Appendix D]. In the interest of not repeating the same, we refer the reader to [13, Appendix D] which proves the existence of strictly positive $\kappa > 0$ such that $\overline{\zeta}_{13}(\underline{m}) + \overline{\zeta}_{23}(\underline{m}) \leq \exp\{-n\kappa\eta_3^2\}$. With this, we conclude

$$\sum_{j=1}^{2} \overline{\zeta}_j(\underline{m}) \leq 2|\mathcal{S}_j| \exp\left\{\frac{-n\eta_3^2\mu_{\mathcal{S}_j}^2}{(2\log|\underline{\mathcal{S}}||\underline{\mathcal{Y}}||\mathcal{W}||\underline{\mathcal{X}}|)^2}\right\} + 4\exp\left\{-n\log q\left[\frac{k_j}{n} - \left(1 - \frac{H(V_j|S_j)\Upsilon + 3\eta_3}{\log q}\right)\right]\right\} + \exp\{-n\tilde{\kappa}\eta_3^2\} \quad (81)$$

Appendix F
Proof Of Thm. 2 - Bound on $\zeta_{32}(\underline{m})$

Recall $\zeta_{32}(\underline{m}) = 2\sum_{\underline{s}^n} \mathrm{p}_{\underline{S}}^n(\underline{s}^n)\sqrt{\mathrm{tr}\left\{[\boldsymbol{I} - \pi_{\eta_1}^{\mu}]\,\rho_{\underline{m},\underline{s}^n}\right\}}\mathbb{1}_{\mathcal{E}}$. We are required to prove $\mathrm{tr}\left\{[\boldsymbol{I} - \pi_{\eta_1}^{\mu}]\,\rho_{\underline{m},\underline{s}^n}\right\}$ falls exponentially in n under conditions of event $\mathcal{E}$. For $(\underline{x},\underline{s}) \in \underline{\mathcal{X}} \times \underline{\mathcal{S}}$, let $\rho_{\underline{x},\underline{s}} \in \mathcal{D}(\mathcal{H})$ have a spectral decomposition $\rho_{\underline{x}\underline{s}} = \sum_{\overline{y}\in\mathcal{Y}} q_{\overline{Y}|\underline{X}\underline{S}}^{\rho}(\overline{y}|\underline{x},\underline{s}) \left|e_{\overline{y}|\underline{x}\underline{s}}\right\rangle\!\left\langle e_{\overline{y}|\underline{x}\underline{s}}\right|$ be a spectral decomposition and let $\mu = \sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}}(\underline{x},\underline{s})\rho_{\underline{x},\underline{s}}$ have a spectral decomposition $\mu = \sum_{y\in\mathcal{Y}} q_Y^{\mu}(y)\,|f_y\rangle\langle f_y|$ where $\{|f_y\rangle : y \in \mathcal{Y}\}$ and $\{\left|e_{\overline{y}|\underline{x}\underline{s}}\right\rangle : \overline{y} \in \mathcal{Y}\}$ for each $(\underline{x},\underline{s}) \in \underline{\mathcal{X}} \times \underline{\mathcal{S}}$, are ONBs that span the Hilbert space $\mathcal{H}$. For $(\underline{x},\underline{s},\overline{y},y) \in \underline{\mathcal{X}} \times \underline{\mathcal{S}} \times \mathcal{Y} \times \mathcal{Y}$, let $r_{\underline{X}\underline{S}\overline{Y}Y}(\underline{x},\underline{s},\overline{y},y) \triangleq p_{\underline{X}\underline{S}}(\underline{x},\underline{s})q_{\overline{Y}|\underline{X}\underline{S}}^{\rho}(\overline{y}|\underline{x},\underline{s})|\left\langle f_y\middle|e_{\overline{y}|\underline{x}\underline{s}}\right\rangle|^2$. Since $\sum_{y\in\mathcal{Y}} |\left\langle f_y\middle|e_{\overline{y}|\underline{x}\underline{s}}\right\rangle|^2 = 1$, it can be verified that (i) $r_{\underline{X}\underline{S}\overline{Y}Y}$ is a PMF, (ii) the marginal $r_{\underline{X}\underline{S}} = p_{\underline{X}\underline{S}}$, (iii) conditional PMFs $r_{\overline{Y}|\underline{X}\underline{S}} = q_{\overline{Y}|\underline{X}\underline{S}}$, $r_{Y\overline{Y}|\underline{X}\underline{S}} = q_{\overline{Y}|\underline{X}\underline{S}}|\left\langle f_y\middle|e_{\overline{y}|\underline{x}\underline{s}}\right\rangle|^2$ and $r_{Y|\overline{Y}\underline{X}\underline{S}} = |\left\langle f_y\middle|e_{\overline{y}|\underline{x}\underline{s}}\right\rangle|^2$. Moreover,

$$\begin{aligned}
\langle f_y|\rho_{\underline{x},\underline{s}}|f_y\rangle &= \langle f_y|\sum_{\overline{y}\in\mathcal{Y}} q_{\overline{Y}|\underline{X}\underline{S}}^{\rho}(\overline{y}|\underline{x},\underline{s})\left|e_{\overline{y}|\underline{x},\underline{s}}\right\rangle\!\left\langle e_{\overline{y}|\underline{x},\underline{s}}\right||f_y\rangle = \sum_{\overline{y}\in\mathcal{Y}} q_{\overline{Y}|\underline{X}\underline{S}}^{\rho}(\overline{y}|\underline{x},\underline{s})|\left\langle f_y\middle|e_{\overline{y}|\underline{x}\underline{s}}\right\rangle|^2 = \sum_{\overline{y}\in\mathcal{Y}} r_{Y\overline{Y}|\underline{X}\underline{S}}(y,\overline{y}|\underline{x},\underline{s}) \\
&= r_{Y|\underline{X}\underline{S}}(y|\underline{x},\underline{s}) \text{ and } r_Y(y) = \sum_{\underline{x},\underline{s}}\sum_{\overline{y}\in\mathcal{Y}} r_{\underline{X}\underline{S}\overline{Y}Y}(\underline{x},\underline{s},\overline{y},y) = \sum_{\underline{x},\underline{s}}\sum_{\overline{y}\in\mathcal{Y}} p_{\underline{X}\underline{S}}(\underline{x},\underline{s})q_{\overline{Y}|\underline{X}\underline{S}}^{\rho}(\overline{y}|\underline{x},\underline{s})|\left\langle f_y\middle|e_{\overline{y}|\underline{x}\underline{s}}\right\rangle|^2 \\
&= \langle f_y|\sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}}(\underline{x},\underline{s})\sum_{\overline{y}\in\mathcal{Y}} q_{\overline{Y}|\underline{X}\underline{S}}^{\rho}(\overline{y}|\underline{x},\underline{s})\left|e_{\overline{y}|\underline{x}\underline{s}}\right\rangle\!\left\langle e_{\overline{y}|\underline{x}\underline{s}}\right||f_y\rangle = \langle f_y|\sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}}(\underline{x},\underline{s})\rho_{\underline{x},\underline{s}}\,|f_y\rangle = \langle f_y|\,\mu\,|f_y\rangle \\
&= \langle f_y|\sum_{\tilde{y}} q_Y^{\mu}(\tilde{y})\,|f_{\tilde{y}}\rangle\langle f_{\tilde{y}}|\,|f_y\rangle = q_Y^{\mu}(y). \qquad (82)
\end{aligned}$$

We therefore have $T_\eta(p_{\underline{X}\underline{S}}) = T_\eta(r_{\underline{X}\underline{S}})$ and $T_\eta(q_Y^{\mu}) = T_\eta(r_Y)$ for any $\eta > 0$. From Lemma 1(iii), we note that $T_\eta(Y|\underline{x}^n,\underline{s}^n) \subseteq T_\eta(r_Y)$ for any $\eta > 0$ and $\underline{x}^n,\underline{s}^n$. Since $(x_j^n(m_j,s_j^n), s_j^n : j \in [2]) \in T_{4\eta_3}(p_{\underline{X}\underline{S}})$

$$\begin{aligned}
\mathrm{tr}\left\{\left[\boldsymbol{I} - \pi_{\eta_1}^{\mu}\right]\rho_{\underline{m},\underline{s}^n}\right\} &= \sum_{y^n\notin T_{\eta_1}^n(r_Y)} |\left\langle f_{y^n}|\rho_{\underline{m},\underline{s}^n}|f_{y^n}\right\rangle| = \sum_{y^n\notin T_{\eta_1}^n(r_Y)} r_{Y|\underline{X}\underline{S}}^n(y^n|x_j^n(m_j,s_j^n),s_j^n : j=1,2) \\
&\leq \sum_{y^n\in\mathcal{Y}^n} r_{Y|\underline{X}\underline{S}}^n(y^n|x_j^n(m_j,s_j^n),s_j^n : j=1,2)\mathbb{1}_{\left\{y^n\in\mathcal{Y}^n\setminus T_{\eta_1}(Y|x_j^n(m_j,s_j^n),s_j^n:j=1,2)\right\}} \\
&\leq 2|\mathcal{Y}||\underline{\mathcal{X}}||\underline{\mathcal{S}}|\exp\left\{-n(\eta_1-4\eta_3)^2\delta(r_{Y\underline{X}\underline{S}},4\eta_3,|\mathcal{Y}||\underline{\mathcal{X}}||\underline{\mathcal{S}}|)\right\}
\end{aligned}$$

if $\eta_1 > 4\eta_3$, where $\delta(\cdot)$ is as defined in (67). The last inequality above follows from Lemma 2 (ii). Since the above bound is invariant to $\underline{s}^n \in \underline{\mathcal{S}}^n$, we have

$$\zeta_{32}(\underline{m}) = 2\sum_{\underline{s}^n} \mathrm{p}_{\underline{S}}^n(\underline{s}^n)\sqrt{\mathrm{tr}\left\{[\boldsymbol{I} - \pi_{\eta_1}^{\mu}]\,\rho_{\underline{m},\underline{s}^n}\right\}}\mathbb{1}_{\mathcal{E}} \leq 2|\mathcal{Y}||\underline{\mathcal{X}}||\underline{\mathcal{S}}|\exp\left\{-n(\eta_1-4\eta_3)^2\delta(r_{Y\underline{X}\underline{S}},4\eta_3,|\mathcal{Y}||\underline{\mathcal{X}}||\underline{\mathcal{S}}|)\right\} \quad (83)$$

Appendix G
Proof of Proposition 2

We are required to derive an upper bound on $\overline{\zeta}_{31}(\underline{m}) = \mathbb{E}\{\zeta_{31}(\underline{m})\}$ and we proceed from (20). Defining,

$$\mathcal{G}_{\underline{s}^n} \triangleq \{\underline{S}^n = \underline{s}^n\},\ \mathcal{G}_1^3 \triangleq \left\{\begin{smallmatrix} V_j^n(a_j,m_j)=v_j^n \\ :j\in[2],W^n(a,\underline{m})=w^n\end{smallmatrix}\right\},\ \mathcal{G}_2^3 \triangleq \left\{\begin{smallmatrix} |\mathcal{L}_j(m_j,s_j^n)| \\ \geq L_j : j\in[2]\end{smallmatrix}\right\},\ \mathcal{G}_3^3 \triangleq \left\{\begin{smallmatrix} A_j(m_j,s_j^n) \\ =a_j : j\in[2]\end{smallmatrix}\right\},\ \mathcal{G}_4^3 \triangleq \left\{\begin{smallmatrix} X_j^n(m_j,s_j^n) \\ =x_j^n : j\in[2]\end{smallmatrix}\right\} \quad (84)$$

$$\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{v}^n,\underline{x}^n \\ w^n,a_1,a_2,a\end{smallmatrix}\right) \triangleq \mathbb{1}_{\left\{(s_j^n,v_j^n)\in T_{\eta_3}(p_{S_jV_j}),s_j^n\in T_{\frac{\eta_3}{2}}(\mathrm{p}_{s_j}),a=a_1\ 0^{k_2-k_1}\oplus a_2,(\underline{s}^n,\underline{x}^n\underline{v}^n,w^n)\in T_{4\eta_3}(p_{\underline{S}\underline{X}\underline{V}W})\right\}},\ \text{we have}$$

$$\overline{\zeta}_{31}(\underline{m}) = 2\sum_{\substack{\underline{s}^n,\underline{v}^n\\ w^n,\underline{x}^n}}\sum_{a_1\in\mathcal{F}_q^{k_1}}\sum_{a_2,a\in\mathcal{F}_q^{k_2}} \mathtt{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{w^n}]\,\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)P\left(\bigcap_{i=1}^4\mathcal{G}_i^3\,\middle|\,\mathcal{G}_{\underline{s}^n}\right),\ \text{where} \tag{85}$$

$$\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)P\left(\bigcap_{i=1}^4\mathcal{G}_i^3\,\middle|\,\mathcal{G}_{\underline{s}^n}\right)=\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)P\left(\mathcal{G}_1^3\cap\mathcal{G}_2^3|\mathcal{G}_{\underline{s}^n}\right)P\left(\mathcal{G}_3^3\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^3\cap\mathcal{G}_2^3\right)P\left(\mathcal{G}_4^3\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^3\cap\mathcal{G}_2^3\cap\mathcal{G}_3^3\right)$$

$$\le P\left(\mathcal{G}_1^3|\mathcal{G}_{\underline{s}^n}\right)\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)\prod_{j=1}^2\frac{p^n_{X_j|V_j,S_j}(x_j^n|v_j^n,s_j^n)}{|\mathcal{L}_j(m_j,s_j^n)|}=\frac{\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)}{q^{2n}}\prod_{j=1}^2\frac{p^n_{X_j|V_j,S_j}(x_j^n|v_j^n,s_j^n)}{|\mathcal{L}_j(m_j,s_j^n)|} \tag{86}$$

$$\le\frac{\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)}{q^{2n}}\prod_{j=1}^2\frac{p^n_{X_j|V_j,S_j}(x_j^n|v_j^n,s_j^n)}{L_j}\le\frac{4}{q^{2n}}\prod_{j=1}^2\exp\left\{n\left(\substack{\log q-H(V_j|S_j)\\+3\eta_3-\frac{k_j}{n}\log q}\right)\right\}p^n_{X_j|V_j,S_j}(x_j^n|v_j^n,s_j^n)\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right) \tag{87}$$

where (i) the inequality in (86) follows from the fact that $P\left(\mathcal{G}_3^3\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^3\cap\mathcal{G}_2^3\right)=\frac{1}{|\mathcal{L}_j(m_j,s_j^n)|}$ which is a consequence of the distribution of the random code specified in (16), in particular Remark 2, and $P\left(\mathcal{G}_4^3\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^3\cap\mathcal{G}_2^3\cap\mathcal{G}_3^3\right)=\prod_{j=1}^2 p^n_{X_j|V_j,S_j}(x_j^n|v_j^n,s_j^n)$ from the distribution of the random code specified in (16), (ii) the equality in (86) follows from $P(\mathcal{G}_1^3|\mathcal{G}_{\underline{s}^n})=\frac{1}{q^{2n}}$ that is proven in Lemma 5 in Appendix D, (iii) the inequalities in (87) follows from $|\mathcal{L}_j(m_j,s_j^n)|\ge L_j=\frac{1}{2}\exp\{k_j\log q-n\log q+H(V_j|S_j)-3\eta_3\}$ defined prior to (11). From the law of total probability (LOTP) and substituting the upper bound (87) in (85), we have

$$\overline{\zeta}_{31}(\underline{m})=8\sum_{\substack{\underline{s}^n,\underline{v}^n\\ w^n,\underline{x}^n}}\sum_{\substack{a_1\in\mathcal{F}_q^{k_1}\\ a_2,a\in\mathcal{F}_q^{k_2}}}\frac{\mathtt{p}^n_{\underline{S}}(\underline{s}^n)}{q^{k_1+k_2}}\,\mathrm{tr}\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{w^n}]\,\rho_{\underline{x}^n,\underline{s}^n}\}\,\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)\prod_{j=1}^2\exp\left\{n\left(\substack{-H(V_j|S_j)\\+3\eta_3\log q}\right)\right\}p^n_{X_j|V_j,S_j}(x_j^n|v_j^n,s_j^n), \tag{88}$$

$$\le 8\sum_{\underline{s}^n,\underline{v}^n}\sum_{w^n,\underline{x}^n}\mathtt{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{w^n}]\,\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)\prod_{j=1}^2\exp\left\{\tfrac{9n\eta_3}{2}\right\}p^n_{X_jV_j|S_j}(x_j^n,v_j^n|s_j^n), \tag{89}$$

$$\le 8\sum_{\underline{s}^n,\underline{v}^n}\sum_{w^n,\underline{x}^n}\mathrm{tr}\left\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{w^n}]\,\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)\exp\left\{9n\eta_3\right\}p^n_{\underline{XVS}}(\underline{x}^n,\underline{v}^n,\underline{s}^n),$$

$$\le 8\sum_{\underline{s}^n,\underline{v}^n}\sum_{w^n,\underline{x}^n}\mathrm{tr}\left\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{w^n}]\,\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)\exp\left\{9n\eta_3\right\}p^n_{\underline{XVS}W}(\underline{x}^n,\underline{v}^n,\underline{s}^n,w^n) \tag{90}$$

$$\le\sum_{w^n}8p^n_W(w^n)\sum_{\underline{s}^n,\underline{v}^n,\underline{x}^n}\mathrm{tr}\left\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{w^n}]\,\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathbb{1}_{\{w^n\in T_{4\eta_3}(p_W)\}}\exp\left\{9n\eta_3\right\}p^n_{\underline{XVS}|W}(\underline{x}^n,\underline{v}^n,\underline{s}^n|w^n) \tag{91}$$

$$\le\sum_{w^n}8p^n_W(w^n)\mathbb{1}_{\{w^n\in T_{4\eta_3}(p_W)\}}\sum_{\underline{s}^n,\underline{v}^n,\underline{x}^n}p^n_{\underline{XVS}|W}(\underline{x}^n,\underline{v}^n,\underline{s}^n|w^n)\,\mathrm{tr}\left\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{w^n}]\,\rho_{\underline{x}^n,\underline{s}^n}\right\}\exp\left\{9n\eta_3\right\}$$

$$\le\sum_{w^n}\frac{8p^n_W(w^n)\mathbb{1}_{\{w^n\in T_{4\eta_3}(p_W)\}}}{\exp\{-n9\eta_3\}}\,\mathrm{tr}\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{w^n}]\sigma_{w^n}\}\le\sum_{w^n}\frac{p^n_W(w^n)\log\left[\dim(\mathcal{H})\right]}{16^{-1}\exp\{-n9\eta_3\}}\exp\left\{-2n(\eta_2-4\eta_3)^2\delta_{\mathrm{q}}(\sigma,p_W,\eta_1)\right\}$$

$$\le 16\log\left[\dim(\mathcal{H})\right]\exp\left\{-n\left[2(\eta_2-4\eta_3)^2\delta_{\mathrm{q}}(\sigma,p_W,\eta_1)-9\eta_3\right]\right\}. \tag{92}$$

where (i) (89) follows from (a) $\exp\left\{-n\left(H(V_j|S_j)+\frac{3\eta_3}{2}\right)\right\}\le p^n_{V_j|S_j}(v_j^n|s_j^n)$ whenever $s_j^n\in T_{\frac{\eta_3}{2}}(\mathtt{p}_{S_j})$, $(s_j^n,v_j^n)\in T_{\eta_3}(p_{S_jV_j})$, (b) the summand in (88) being invariant to a_1,a_2 and (c) the sum over $a\in\mathcal{F}_q^{k_2}$ being trivial owing to the fact that $a=a_1\,0^{k_2-k_1}\oplus a_2$, (ii) (90) follows from the chosen PMF $p_{\underline{SVX}W}$ satisfying $p_{\underline{SXV}W}(\underline{s},\underline{x},\underline{v},w)=p_{\underline{SXV}}(\underline{s},\underline{x},\underline{v})\mathbb{1}_{\{w=v_1\oplus v_2\}}$ and the fact that $w^n=v_1^n\oplus v_2^n$ ensured by the factor $\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)$ which guarantees[4] $(\underline{s}^n,\underline{v}^n,\underline{x}^n,w^n)\in T_{4\eta_3}(p_{\underline{SXV}W})$, (iii) (91) follows from the indicator function in question being larger than or equal to the factor $\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)$, and lastly the inequality in (92) and the inequality prior to that is a result of the substantial overlap of the conditional typical projector $\pi^{\sigma,\eta_2}_{w^n}$ with σ_{w^n} whenever $w^n\in T_{4\eta_3}(p_W)$ and $\eta_2>4\eta_3$ as stated in Lemma 4.

[4]In other words, $p^n_{W|\underline{SVX}}(w^n|\underline{s}^n,\underline{v}^n,\underline{x}^n)=\mathbb{1}_{\{w^n=v_1^n\oplus v_2^n\}}$ and the latter indicator function is evident from the factor $\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n,\underline{x}^n\\ w^n,a_1,a_2,a}\right)$.

APPENDIX H
PROOF OF PROPOSITION 3

We are required to derive an upper bound on $\overline{\zeta}_4(\underline{m}) = \mathbb{E}\left\{\zeta_4(\underline{m})\right\}$ and we proceed from (22). Defining, $\mathcal{G}_1^4 \triangleq \left\{ \begin{smallmatrix} V_j^n(a_j,m_j)=v_j^n : j= \\ 1,2,\ W^n(\hat{a},\underline{m})=\hat{w}^n \end{smallmatrix} \right\}, \mathcal{G}_i^4 \triangleq \mathcal{G}_i^3 : i=2,3,4$ and $\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{v}^n \\ a_2,a_1,\hat{a}\end{smallmatrix}\right) \triangleq \mathbb{1}_{\left\{\hat{a}\neq a_1\ 0^{k_2-k_1}\oplus a_2, s_j^n\in T_{\frac{\eta_3}{2}}(\mathtt{p}_{s_j}),(s_j^n,v_j^n)\in T_{\eta_3}(p_{S_jV_j}):j\in[2]\right\}}$ where $\mathcal{G}_i^3 : i\in[4]$ and $\mathcal{G}_{\underline{s}^n}$ are specified in (84), we have

$$\overline{\zeta}_4(\underline{m}) = 2\sum_{\substack{\underline{s}^n,\underline{v}^n \\ \underline{x}^n,\hat{w}^n}} \sum_{a_1\in\mathcal{F}_q^{k_1}} \sum_{\substack{a_2,\hat{a} \\ \in\mathcal{F}_q^{k_2}}} \mathtt{p}_{\underline{S}}^n(\underline{s}^n)\,\mathrm{tr}\left\{\pi_{\eta_1}^{\mu}\pi_{\hat{w}^n}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu}\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{v}^n \\ a_2,a_1,\hat{a}\end{smallmatrix}\right) P\left(\bigcap_{i=1}^4 \mathcal{G}_i^4 \,\middle|\, \mathcal{G}_{\underline{s}^n}\right), \text{ where} \tag{93}$$

$$\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{v}^n \\ a_2,a_1,\hat{a}\end{smallmatrix}\right) P\left(\bigcap_{i=1}^4 \mathcal{G}_i^4 \,\middle|\, \mathcal{G}_{\underline{s}^n}\right) = P\left(\mathcal{G}_1^4\cap\mathcal{G}_2^4|\mathcal{G}_{\underline{s}^n}\right) P\left(\mathcal{G}_3^4\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^4\cap\mathcal{G}_2^4\right) P\left(\mathcal{G}_4^4\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^4\cap\mathcal{G}_2^4\cap\mathcal{G}_3^4\right)\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{v}^n \\ a_2,a_1,\hat{a}\end{smallmatrix}\right)$$

$$\leq P\left(\mathcal{G}_1^4|\mathcal{G}_{\underline{s}^n}\right)\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{v}^n \\ a_2,a_1,\hat{a}\end{smallmatrix}\right)\prod_{j=1}^2 \frac{p_{X_j|V_j,S_j}^n(x_j^n|v_j^n,s_j^n)}{|\mathcal{L}_j(m_j,s_j^n)|} \leq \frac{1}{q^{3n}}\prod_{j=1}^2 \frac{p_{X_j|V_j,S_j}^n(x_j^n|v_j^n,s_j^n)}{L_j} \tag{94}$$

$$= \frac{4}{q^{3n}}\prod_{j=1}^2 \exp\left\{n\log q - k_j\log q - H(V_j|S_j) + 3\eta_3\right\} p_{X_j|V_j,S_j}^n(x_j^n|v_j^n,s_j^n) \tag{95}$$

where (i) the first inequality in (94) follows from the fact that $P\left(\mathcal{G}_3^4\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^4\cap\mathcal{G}_2^4\right) = \frac{1}{|\mathcal{L}_j(m_j,s_j^n)|}$ which is a consequence of the distribution of the random code specified in (16), in particular Remark 2, and $P\left(\mathcal{G}_4^4\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^4\cap\mathcal{G}_2^4\cap\mathcal{G}_3^4\right) = \prod_{j=1}^2 p_{X_j|V_j,S_j}^n(x_j^n|v_j^n,s_j^n)$ from the distribution of the random code specified in (16), (ii) the second inequality in (94) follows from $P(\mathcal{G}_1^4|\mathcal{G}_{\underline{s}^n}) = \frac{1}{q^{3n}}$ that is proven in Lemma 5 and $|\mathcal{L}_j(m_j,s_j^n)| \geq L_j = \frac{1}{2}\exp\{k_j\log q - n\log q + H(V_j|S_j) - 3\eta_3\}$ defined prior to (11). From the LOTP and substituting the upper bound (95) in (93), we have

$$\overline{\zeta}_4(\underline{m}) = 8\sum_{\substack{\underline{s}^n,\underline{v}^n \\ \underline{x}^n,\hat{w}^n}} \sum_{a_1\in\mathcal{F}_q^{k_1}} \sum_{\substack{a_2,\hat{a} \\ \in\mathcal{F}_q^{k_2}}} \frac{\mathtt{p}_{\underline{S}}^n(\underline{s}^n)}{q^{k_1+k_2+n}}\,\mathrm{tr}\left\{\pi_{\eta_1}^{\mu}\pi_{\hat{w}^n}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu}\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{v}^n \\ a_2,a_1,\hat{a}\end{smallmatrix}\right)\prod_{j=1}^2 \exp\left\{n\left(\begin{smallmatrix}-H(V_j|S_j) \\ +3\eta_3\log q\end{smallmatrix}\right)\right\} p_{X_j|V_j,S_j}^n(x_j^n|v_j^n,s_j^n),$$

$$\leq 8\sum_{\substack{\underline{s}^n,\underline{v}^n \\ \underline{x}^n,\hat{w}^n}} \sum_{a_1\in\mathcal{F}_q^{k_1}} \sum_{a_2,\hat{a}\in\mathcal{F}_q^{k_2}} \frac{\mathtt{p}_{\underline{S}}^n(\underline{s}^n)}{q^{k_1+k_2+n}}\,\mathrm{tr}\left\{\pi_{\hat{w}^n}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu}\rho_{\underline{x}^n,\underline{s}^n}\pi_{\eta_1}^{\mu}\right\}\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{v}^n \\ a_2,a_1,\hat{a}\end{smallmatrix}\right)\prod_{j=1}^2 \exp\left\{\frac{9n\eta_3}{2}\right\} p_{X_jV_j|S_j}^n(x_j^n,v_j^n|s_j^n) \tag{96}$$

$$\leq \frac{8}{q^{k_1+k_2+n}}\sum_{\substack{\underline{s}^n,\underline{v}^n \\ \underline{x}^n,\hat{w}^n}} \sum_{a_1\in\mathcal{F}_q^{k_1}} \sum_{a_2,\hat{a}\in\mathcal{F}_q^{k_2}} \mathrm{tr}\left\{\pi_{\hat{w}^n}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu}\rho_{\underline{x}^n,\underline{s}^n}\pi_{\eta_1}^{\mu}\right\}\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{v}^n \\ a_2,a_1,\hat{a}\end{smallmatrix}\right)\exp\{9n\eta_3\}\, p_{\underline{XVS}}^n(\underline{x}^n,\underline{v}^n,\underline{s}^n),$$

$$\leq \frac{8}{q^{k_1+k_2+n}}\sum_{\hat{w}^n} \sum_{a_1\in\mathcal{F}_q^{k_1}} \sum_{a_2,\hat{a}\in\mathcal{F}_q^{k_2}} \mathrm{tr}\left\{\pi_{\hat{w}^n}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu}\sum_{\underline{s}^n,\underline{v}^n\underline{x}^n} p_{\underline{XVS}}^n(\underline{x}^n,\underline{v}^n,\underline{s}^n)\rho_{\underline{x}^n,\underline{s}^n}\pi_{\eta_1}^{\mu}\right\}\exp\{9n\eta_3\}$$

$$\leq \frac{8}{q^{k_1+k_2+n}}\sum_{\hat{w}^n} \sum_{a_1\in\mathcal{F}_q^{k_1}} \sum_{a_2,\hat{a}\in\mathcal{F}_q^{k_2}} \mathrm{tr}\left\{\pi_{\hat{w}^n}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu}\mu^{\otimes n}\pi_{\eta_1}^{\mu}\right\}\exp\{9n\eta_3\} \tag{97}$$

$$\leq \frac{8}{q^{k_1+k_2+n}}\sum_{\hat{w}^n} \sum_{a_1\in\mathcal{F}_q^{k_1}} \sum_{a_2,\hat{a}\in\mathcal{F}_q^{k_2}} \mathrm{tr}\left\{\pi_{\hat{w}^n}^{\sigma,\eta_2}\pi_{\eta_1}^{\mu}\right\}\exp\{-n(H(Y)_\Upsilon - 9\eta_3 - \eta_1)\} \tag{98}$$

$$\leq \frac{8}{q^{k_1+k_2+n}}\sum_{\hat{w}^n} \sum_{a_1\in\mathcal{F}_q^{k_1}} \sum_{a_2,\hat{a}\in\mathcal{F}_q^{k_2}} \mathrm{tr}\left\{\pi_{\hat{w}^n}^{\sigma,\eta_2}\boldsymbol{I}\right\}\exp\{n(H(Y)_\Upsilon - 9\eta_3 - \eta_1)\} \tag{99}$$

$$\leq \frac{8}{q^{k_1+k_2+n}} \sum_{\hat{w}^n \in T_{\eta_2}(p_W)} \sum_{a_1 \in \mathcal{F}_q^{k_1}} \sum_{a_2, \hat{a} \in \mathcal{F}_q^{k_2}} \exp\left\{-n\left(H(Y)_\Upsilon - H(Y|W)_\Upsilon - 9\eta_3 - 2\eta_1 - \eta_2\right)\right\}$$

$$+ \frac{8}{q^{k_1+k_2+n}} \sum_{\hat{w}^n \in \mathcal{W}^n \setminus T_{\eta_2}(p_W)} \sum_{a_1 \in \mathcal{F}_q^{k_1}} \sum_{a_2, \hat{a} \in \mathcal{F}_q^{k_2}} 0 \cdot \exp\left\{n(H(Y)_\Upsilon - 9\eta_3 - \eta_1)\right\} \tag{100}$$

$$\leq 8 \exp\left\{-n\left(H(Y)_\Upsilon - H(Y|W)_\Upsilon + \log q - H(W)_\Upsilon - 9\eta_3 - 2\eta_1 - \eta_2 - \frac{k_2}{n}\log q\right)\right\} \tag{101}$$

where (i) (96) follows from the $\exp\left\{-n\left(H(V_j|S_j) + \frac{3\eta_3}{2}\right)\right\} \leq p^n_{V_j|S_j}(v_j^n|s_j^n)$ whenever $s_j^n \in T_{\frac{\eta_3}{2}}(\mathtt{p}_{S_j})$ and $(s_j^n, v_j^n) \in T_{\eta_3}(p_{S_jV_j})$, (ii) (97) follows from $\mu = \sum_{\underline{x},\underline{s}} p_{\underline{X}\underline{S}}(\underline{x},\underline{s})\rho_{\underline{x},\underline{s}}$ as defined prior to (32), (iii) (98) follows from Lemma 3, since $\pi^\mu_{\eta_1}$ is the typical projector of $\mu^{\otimes n}$ (iv) (100) follows from the fact stated in Lemma 4 which states that the conditional typical projector is the zero operator if the conditioning codeword is not typical with the same parameter, and otherwise, its trace is dominated as specified in Lemma 4, and finally (v) (101) follows from the bound on the size of the typical set $|T_{\eta_2}(p_W)|$ as stated in Lemma 1.

APPENDIX I
PROOF OF PROPOSITION 4

We are required to derive an upper bound on $\overline{\zeta}_5(\underline{m}) = \mathbb{E}\left\{\zeta_5(\underline{m})\right\}$ and we proceed from (23). Defining,

$\mathcal{G}_1^5 \triangleq \left\{\substack{V_j^n(a_j,m_j)=v_j^n:j=\\ 1,2,\ W^n(\hat{a},\underline{\hat{m}})=\hat{w}^n}\right\}, \mathcal{G}_i^5 \triangleq \mathcal{G}_i^3 : i = 2,3,4$ and $\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n\\ \underline{m},\underline{\hat{m}}}\right) \triangleq \mathbb{1}_{\left\{\underline{\hat{m}} \neq \underline{m}, s_j^n \in T_{\frac{\eta_3}{2}}(\mathtt{p}_{s_j}), (s_j^n, v_j^n) \in T_{\eta_3}(p_{S_jV_j}) : j \in [2]\right\}}$

where $\mathcal{G}_i^3 : i \in [4]$ and $\mathcal{G}_{\underline{s}^n}$ are specified in (84), we have

$$\overline{\zeta}_5(\underline{m}) = 2 \sum_{\substack{\underline{s}^n,\underline{v}^n\\ \underline{x}^n,\hat{w}^n}} \sum_{\substack{a_1 \in \mathcal{F}_q^{k_1}\\ \underline{\hat{m}} \neq \underline{m}}} \sum_{\substack{a_2,\hat{a}\\ \in \mathcal{F}_q^{k_2}}} \mathtt{p}^n_{\underline{S}}(\underline{s}^n) \operatorname{tr}\left\{\pi^\mu_{\eta_1} \pi^{\sigma,\eta_2}_{\hat{w}^n} \pi^\mu_{\eta_1} \rho_{\underline{x}^n,\underline{s}^n}\right\} \mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n\\ \underline{m},\underline{\hat{m}}}\right) P\left(\bigcap_{i=1}^4 \mathcal{G}_i^5 \middle| \mathcal{G}_{\underline{s}^n}\right), \text{ where}$$

$$\mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n\\ \underline{m},\underline{\hat{m}}}\right) P\left(\bigcap_{i=1}^4 \mathcal{G}_i^5 \middle| \mathcal{G}_{\underline{s}^n}\right) = \mathcal{G}\left(\substack{\underline{s}^n,\underline{v}^n\\ \underline{m},\underline{\hat{m}}}\right) P\left(\mathcal{G}_1^5 \cap \mathcal{G}_2^5 | \mathcal{G}_{\underline{s}^n}\right) P\left(\mathcal{G}_3^5 \middle| \mathcal{G}_{\underline{s}^n} \cap \mathcal{G}_1^5 \cap \mathcal{G}_2^5\right) P\left(\mathcal{G}_4^5 \middle| \mathcal{G}_{\underline{s}^n} \cap \mathcal{G}_1^5 \cap \mathcal{G}_2^5 \cap \mathcal{G}_3^5\right)$$

$$\leq \mathcal{G}\left(\underline{s}^n,\underline{v}^n,\underline{m},\underline{\hat{m}}\right) P\left(\mathcal{G}_1^5 | \mathcal{G}_{\underline{s}^n}\right) \prod_{j=1}^2 \frac{p^n_{X_j|V_j,S_j}(x_j^n|v_j^n,s_j^n)}{|\mathcal{L}_j(m_j,s_j^n)|} \leq \frac{\mathcal{G}\left(\underline{s}^n,\underline{v}^n,\underline{m},\underline{\hat{m}}\right)}{q^{3n}} \prod_{j=1}^2 \frac{p^n_{X_j|V_j,S_j}(x_j^n|v_j^n,s_j^n)}{L_j} \tag{102}$$

$$= \frac{4\mathcal{G}\left(\underline{s}^n,\underline{v}^n,\underline{m},\underline{\hat{m}}\right)}{q^{3n}} \prod_{j=1}^2 \exp\left\{n\left(\log q - H(V_j|S_j) + 3\eta_3 - \frac{k_j}{n}\log q\right)\right\} p^n_{X_j|V_j,S_j}(x_j^n|v_j^n,s_j^n) \tag{103}$$

where (i) the first inequality in (102) follows from the fact that $P\left(\mathcal{G}_3^5 \middle| \mathcal{G}_{\underline{s}^n} \cap \mathcal{G}_1^5 \cap \mathcal{G}_2^5\right) = \frac{1}{|\mathcal{L}_j(m_j,s_j^n)|}$ which is a consequence of the distribution of the random code specified in (16), in particular Remark 2, and $P\left(\mathcal{G}_4^5 \middle| \mathcal{G}_{\underline{s}^n} \cap \mathcal{G}_1^5 \cap \mathcal{G}_2^5 \cap \mathcal{G}_3^5\right) = \prod_{j=1}^2 p^n_{X_j|V_j,S_j}(x_j^n|v_j^n,s_j^n)$ from the distribution of the random code specified in (16), (ii) the second inequality in (102) follows from $P(\mathcal{G}_1^5|\mathcal{G}_{\underline{s}^n}) = \frac{1}{q^{3n}}$ that is proven in Lemma 5 in Appendix D and $|\mathcal{L}_j(m_j,s_j^n)| \geq L_j = \frac{1}{2}\exp\{k_j \log q - n\log q + H(V_j|S_j) - 3\eta_3\}$ defined prior to (11). From LOTP and substituting the upper bound (103) in (102), we have

$$\overline{\zeta}_5(\underline{m}) = 8\sum_{\substack{\underline{s}^n,\underline{v}^n\\ \underline{x}^n,\hat{w}^n}}\sum_{\substack{a_1\in\mathcal{F}_q^{k_1}\\ \hat{\underline{m}}\neq\underline{m}}}\sum_{\substack{a_2,\hat{a}\\ \in\mathcal{F}_q^{k_2}}}\frac{\mathtt{p}^n_{\underline{S}}(\underline{s}^n)}{q^{k_1+k_2+n}}\operatorname{tr}\left\{\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{\hat{w}^n}\pi^{\mu}_{\eta_1}\rho_{\underline{x}^n,\underline{s}^n}\right\}\mathcal{G}\left(\tfrac{\underline{s}^n,\underline{v}^n}{\underline{m},\hat{\underline{m}}}\right)\prod_{j=1}^{2}\exp\left\{n\begin{pmatrix}-H(V_j|S_j)\\ +3\eta_3\log q\end{pmatrix}\right\}p^n_{X_j|V_j,S_j}(x_j^n|v_j^n,s_j^n),$$

$$\leq 8\sum_{\underline{s}^n,\underline{v}^n}\sum_{\underline{x}^n,\hat{w}^n}\sum_{\substack{a_1\in\mathcal{F}_q^{k_1}\\ \hat{\underline{m}}\neq\underline{m}}}\sum_{\substack{a_2,\hat{a}\\ \in\mathcal{F}_q^{k_2}}}\frac{\mathtt{p}^n_{\underline{S}}(\underline{s}^n)}{q^{k_1+k_2+n}}\operatorname{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{w}^n}\pi^{\mu}_{\eta_1}\rho_{\underline{x}^n,\underline{s}^n}\pi^{\mu}_{\eta_1}\right\}\mathcal{G}\left(\tfrac{\underline{s}^n,\underline{v}^n}{\underline{m},\hat{\underline{m}}}\right)\prod_{j=1}^{2}\exp\left\{\frac{9n\eta_3}{2}\right\}p^n_{X_jV_j|S_j}(x_j^n,v_j^n|s_j^n), \tag{104}$$

$$\leq \frac{8}{q^{k_1+k_2+n}}\sum_{\underline{s}^n,\underline{v}^n}\sum_{\underline{x}^n,\hat{w}^n}\sum_{a_1\in\mathcal{F}_q^{k_1}}\sum_{\hat{\underline{m}}\neq\underline{m}}\sum_{a_2,\hat{a}\in\mathcal{F}_q^{k_2}}\operatorname{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{w}^n}\pi^{\mu}_{\eta_1}\rho_{\underline{x}^n,\underline{s}^n}\pi^{\mu}_{\eta_1}\right\}\mathcal{G}\left(\tfrac{\underline{s}^n,\underline{v}^n}{\underline{m},\hat{\underline{m}}}\right)\exp\{9n\eta_3\}\,p^n_{\underline{X}\underline{V}\underline{S}}(\underline{x}^n,\underline{v}^n,\underline{s}^n), \tag{105}$$

$$\leq \frac{8}{q^{k_1+k_2+n}}\sum_{\hat{w}^n}\sum_{a_1\in\mathcal{F}_q^{k_1}}\sum_{\substack{\hat{\underline{m}}\neq\underline{m}\\ \hat{\underline{m}}\neq\underline{m}}}\sum_{\substack{a_2,\hat{a}\\ \in\mathcal{F}_q^{k_2}}}\operatorname{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{w}^n}\pi^{\mu}_{\eta_1}\sum_{\underline{s}^n,\underline{v}^n,\underline{x}^n}p^n_{\underline{X}\underline{V}\underline{S}}(\underline{x}^n,\underline{v}^n,\underline{s}^n)\rho_{\underline{x}^n,\underline{s}^n}\pi^{\mu}_{\eta_1}\right\}\exp\{9n\eta_3\}$$

$$\leq \frac{8}{q^{k_1+k_2+n}}\sum_{\hat{w}^n}\sum_{a_1\in\mathcal{F}_q^{k_1}}\sum_{\hat{\underline{m}}\neq\underline{m}}\sum_{\substack{a_2,\hat{a}\\ \in\mathcal{F}_q^{k_2}}}\operatorname{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{w}^n}\pi^{\mu}_{\eta_1}\mu^{\otimes n}\pi^{\mu}_{\eta_1}\right\}\exp\{9n\eta_3\} \tag{106}$$

$$\leq \frac{8}{q^{k_1+k_2+n}}\sum_{\hat{w}^n}\sum_{a_1\in\mathcal{F}_q^{k_1}}\sum_{\hat{\underline{m}}\neq\underline{m}}\sum_{\substack{a_2,\hat{a}\\ \in\mathcal{F}_q^{k_2}}}\operatorname{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{w}^n}\pi^{\mu}_{\eta_1}\right\}\exp\left\{-n\left(H(Y)_{\Upsilon}-{}^{9\eta_3-\eta_1}\right)\right\} \tag{107}$$

$$\leq \frac{8}{q^{k_1+k_2+n}}\sum_{\hat{w}^n}\sum_{a_1\in\mathcal{F}_q^{k_1}}\sum_{\hat{\underline{m}}\neq\underline{m}}\sum_{a_2,\hat{a}\in\mathcal{F}_q^{k_2}}\operatorname{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{w}^n}\boldsymbol{I}\right\}\exp\left\{-n\left(H(Y)_{\Upsilon}-{}^{9\eta_3-\eta_1}\right)\right\}$$

$$\leq \frac{8}{q^{k_1+k_2+n}}\sum_{\hat{w}^n\in T_{\eta_2}(p_W)}\sum_{a_1\in\mathcal{F}_q^{k_1}}\sum_{\hat{\underline{m}}\neq\underline{m}}\sum_{a_2,\hat{a}\in\mathcal{F}_q^{k_2}}\exp\left\{-n\left(H(Y)_{\Upsilon}-H(Y|W)_{\Upsilon}-9\eta_3-\eta_1-\eta_2\right)\right\} \tag{108}$$

$$+\frac{8}{q^{k_1+k_2+n}}\sum_{\hat{w}^n\in\mathcal{W}^n\setminus T_{\eta_2}(p_W)}\sum_{a_1\in\mathcal{F}_q^{k_1}}\sum_{\hat{\underline{m}}\neq\underline{m}}\sum_{a_2,\hat{a}\in\mathcal{F}_q^{k_2}}0\cdot\exp\left\{-n\left(H(Y)_{\Upsilon}-{}^{9\eta_3-\eta_1}\right)\right\} \tag{109}$$

$$\leq 8\exp\left\{-n\left(H(Y)_{\Upsilon}-H(Y|W)_{\Upsilon}+\log q-H(W)_{\Upsilon}-9\eta_3-\eta_1-2\eta_2-\frac{k_2+l_1+l_2}{n}\log q\right)\right\} \tag{110}$$

where (i) (104) follows from the $\exp\left\{-n\left(H(V_j|S_j)+\frac{3\eta_3}{2}\right)\right\}\leq p^n_{V_j|S_j}(v_j^n|s_j^n)$ whenever $s_j^n\in T_{\frac{\eta_3}{2}}(\mathtt{p}_{S_j})$ and $(s_j^n,v_j^n)\in T_{\eta_3}(p_{S_jV_j})$, (ii) (106) follows from $\mu=\sum_{\underline{x},\underline{s}}p_{\underline{X}\underline{S}}(\underline{x},\underline{s})\rho_{\underline{x},\underline{s}}$ as defined prior to (32), (iii) (107) follows from Lemma 3, since $\pi^{\mu}_{\eta_1}$ is the typical projector of $\mu^{\otimes n}$ (iv) (109) follows from the fact stated in Lemma 4 which states that the conditional typical projector is the zero operator if the conditioning codeword is not typical with the same parameter, and otherwise, its trace is dominated as specified in Lemma 4, and finally (v) (110) follows from the bound on the size of the typical set $|T_{\eta_2}(p_W)|$ as stated in Lemma 1.

APPENDIX J
PROOF OF PROPOSITION 7

We are required to derive an upper bound on $\overline{\xi_3}(\underline{m})=\mathbb{E}\left\{\xi_3(\underline{m})\right\}$ and we proceed from (40). Let

$$\mathcal{G}_{\underline{s}^n}\triangleq\left\{\underline{S}^n=\underline{s}^n\right\},\mathcal{G}_1^3\triangleq\left\{\begin{smallmatrix}V_j^n(a_j,m_{j2})=v_j^n\\ U_j^n(m_{j1},b_j)=u_j^n:j\in[2]\end{smallmatrix}\right\},\mathcal{G}_2^3\triangleq\left\{\begin{smallmatrix}|\mathcal{L}_{j1}(m_{j1},s_j^n)|\geq L_{j1}\\ |\mathcal{L}_{j2}(\underline{m}_j,s_j^n)|\geq L_{j2}:j\in[2]\end{smallmatrix}\right\},\mathcal{G}_3^3\triangleq\left\{\begin{smallmatrix}B_j(m_{j1},s_j^n)=B_j^*=b_j\\ A_j(\underline{m}_j,s_j^n)=a_j:j\in[2]\end{smallmatrix}\right\} \tag{111}$$

$$\mathcal{G}_4^3\triangleq\left\{\begin{smallmatrix}X_j^n(\underline{m}_j,s_j^n)\\ =x_j^n:j\in[2]\end{smallmatrix}\right\},\mathcal{G}\left(\tfrac{\underline{s}^n,\underline{u}^n}{\underline{v}^n,\underline{x}^n}\right)\triangleq\mathbb{1}_{\left\{\begin{smallmatrix}(s_j^n,u_j^n,v_j^n)\in T_{\eta_5}(p_{S_jU_jV_j}),(s_j^n,u_j^n)\in T_{\frac{\eta_5}{2}}(p_{S_jU_j})\\ s_j^n\in T_{\frac{\eta_5}{4}}(\mathtt{p}_{s_j}),(\underline{s}^n,\underline{u}^n\underline{v}^n,\underline{x}^n)\in T_{4\eta_5}(p_{\underline{S}\underline{U}\underline{V}\underline{X}})\end{smallmatrix}\right\}},\text{ we have} \tag{112}$$

$$\bar{\xi}_3(\underline{m}) = 3 \sum_{\substack{\underline{s}^n,\underline{u}^n \\ \underline{v}^n,\underline{x}^n}} \sum_{b_1,b_2} \sum_{a_1,a_2} \mathrm{p}^n_{\underline{S}}(\underline{s}^n) \left\| \pi^{2,\eta_3}_{u_2^n} \rho_{\underline{x}^n,\underline{s}^n} \pi^{2,\eta_3}_{u_2^n} - \rho_{\underline{x}^n,\underline{s}^n} \right\|_1 \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n \\ \underline{v}^n,\underline{x}^n}\right) P\left(\bigcap_{i=1}^4 \mathcal{G}_i^3 \,\middle|\, \mathcal{G}_{\underline{s}^n} \right), \text{ where} \tag{113}$$

$$\mathcal{G}\left(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n\right) P\left(\bigcap_{i=1}^4 \mathcal{G}_i^3 \,\middle|\, \mathcal{G}_{\underline{s}^n} \right) = \mathcal{G}\left(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n\right) P\left(\mathcal{G}_1^3 \cap \mathcal{G}_2^3 | \mathcal{G}_{\underline{s}^n}\right) P\left(\mathcal{G}_3^3 \,|\mathcal{G}_{\underline{s}^n} \cap \mathcal{G}_1^3 \cap \mathcal{G}_2^3\right) P\left(\mathcal{G}_4^3 \,|\mathcal{G}_{\underline{s}^n} \cap \mathcal{G}_1^3 \cap \mathcal{G}_2^3 \cap \mathcal{G}_3^3\right)$$

$$\leq P\left(\mathcal{G}_1^3 | \mathcal{G}_{\underline{s}^n}\right) \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n \\ \underline{v}^n,\underline{x}^n}\right) \prod_{j=1}^2 \frac{p^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n)}{|\mathcal{L}_{j1}(m_{j1},s_j^n)||\mathcal{L}_{j2}(\mathfrak{m}_j,s_j^n)|} \leq \frac{\mathcal{G}\left(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n\right)}{q^{2n}} \prod_{j=1}^2 \frac{p^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n) p^n_{U_j}(u_j^n)}{L_{j1}L_{j2}} \tag{114}$$

$$\leq \mathcal{G}\left(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n\right) \prod_{j=1}^2 \frac{4q^n p^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n) p^n_{U_j}(u_j^n) \exp\left\{\frac{9n\eta_5}{2}\right\}}{q^{n+k_j} \exp\{n\left(\frac{\log|\mathcal{B}_j|}{n} - I(U_j;S_j)_\Upsilon + H(V_j|S_j,U_j)_\Upsilon\right)\}} \tag{115}$$

$$\leq \mathcal{G}\left(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n\right) \prod_{j=1}^2 \frac{1}{q^{k_j}|\mathcal{B}_j|} 4p^n_{X_j,U_j,V_j|S_j}(x_j^n,u_j^n,v_j^n|s_j^n) \exp\left\{\frac{29n\eta_5}{4}\right\}, \tag{116}$$

where (i) the first inequality in (114) follows from the fact that $P\left(\mathcal{G}_3^3 \,|\mathcal{G}_{\underline{s}^n} \cap \mathcal{G}_1^3 \cap \mathcal{G}_2^3\right) = \frac{1}{|\mathcal{L}_{j1}(m_j,s_j^n)||\mathcal{L}_{j2}(\mathfrak{m}_j,s_j^n)|}$ and $P\left(\mathcal{G}_4^3 \,|\mathcal{G}_{\underline{s}^n} \cap \mathcal{G}_1^3 \cap \mathcal{G}_2^3 \cap \mathcal{G}_3^3\right) = \prod_{j=1}^2 p^n_{X_j|U_jV_jS_j}(x_j^n|u_j^n,v_j^n,s_j^n)$, both of which are consequences of the distribution of the random code (see Remark 4 for the former equality), (ii) the second inequality in (114) follows from (a) $P(\mathcal{G}_1^3|\mathcal{G}_{\underline{s}^n}) = \frac{1}{q^{2n}} \prod_{j=1}^2 p^n_{U_j}(u_j^n)$ - a consequence of Lemma 5 in Appendix D, the random $\mathcal{U}_1-,\mathcal{U}_2-$codebooks being mutually independent with the codewords of the $\mathcal{U}_j-$codebook distributed as $p^n_{U_j}$ - and (b) $|\mathcal{L}_{j1}(m_j,s_j^n)| \geq L_{j1}$, $|\mathcal{L}_{j2}(\mathfrak{m}_j,s_j^n)| \geq L_{j2}$, (iii) the inequalities in (115) follows from $|\mathcal{L}_{j1}(m_j,s_j^n)| \geq L_{j1} = \frac{1}{2}\exp\{n\left(\frac{\log|\mathcal{B}_j|}{n} - I(U_j;S_j)_\Upsilon - \frac{3\eta_5}{2}\right)\}$ and $|\mathcal{L}_{j2}(\mathfrak{m}_j,s_j^n)| \geq L_{j2} = \frac{1}{2}\exp\{n\left(\frac{\log q^{k_j}}{n} - H(V_j|U_jS_j)_\Upsilon - 3\eta_5\right)\}$ defined prior to (36) and finally (iv) (116) follows from the bound $\frac{p^n_{U_jV_jS_j}(u_j^n,v_j^n,s_j^n)}{p^n_{U_j}(u_j^n)p^n_{S_j}(s_j^n)} \geq \exp\{-n\left(H(S_j,U_j,V_j)_\Upsilon + \eta_5 - H(S_j)_\Upsilon - H(U_j)_\Upsilon + \frac{\eta_5}{2} + \frac{\eta_5}{4}\right)\} = \exp\{-n\left(H(V_j|U_j,S_j)_\Upsilon - I(U_j;S_j)_\Upsilon + \frac{7\eta_5}{4}\right)\}$ for $(\underline{s}^n,\underline{u}^n,\underline{v}^n)$ satisfying the conditions of $\mathcal{G}\left(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n\right)$, imply-

ing $1 \leq \frac{p^n_{U_jV_jS_j}(u_j^n,v_j^n,s_j^n)}{p^n_{U_j}(u_j^n)p^n_{S_j}(s_j^n)} \exp\left\{n\left(H(V_j|U_j,S_j)_\Upsilon - I(U_j;S_j)_\Upsilon + \frac{7\eta_5}{4}\right)\right\}$. We have

$$\overline{\xi}_3(\underline{m}) \leq 48 \sum_{\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}} \sum_{\substack{b_1,b_2\\ a_1,a_2}} \frac{\mathtt{p}^n_{\underline{S}}(\underline{s}^n)}{q^{k_1+k_2}} \left\| \pi^{2,\eta_3}_{u_2^n} \rho_{\underline{x}^n,\underline{s}^n} \pi^{2,\eta_3}_{u_2^n} - \rho_{\underline{x}^n,\underline{s}^n} \right\|_1 \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}\right) \prod_{j=1}^2 \frac{p^n_{X_jU_j,V_j|S_j}(x_j^n,u_j^n,v_j^n|s_j^n)\exp\{\frac{n29\eta_5}{4}\}}{|\mathcal{B}_j|}, \tag{117}$$

$$\leq 96 \sum_{\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}} \sum_{\substack{b_1,b_2\\ a_1,a_2}} \frac{\mathtt{p}^n_{\underline{S}}(\underline{s}^n)}{q^{k_1+k_2}} \sqrt{\mathrm{tr}\{(\boldsymbol{I}-\pi^{2,\eta_3}_{u_2^n})\rho_{\underline{x}^n,\underline{s}^n}\}} \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}\right) \prod_{j=1}^2 \frac{p^n_{X_jU_j,V_j|S_j}(x_j^n,u_j^n,v_j^n|s_j^n)\exp\{\frac{n29\eta_5}{4}\}}{|\mathcal{B}_j|}, \tag{118}$$

$$\leq \frac{96\exp\{\frac{n29\eta_5}{4}\}}{q^{k_1+k_2}|\mathcal{B}_1||\mathcal{B}_2|} \sum_{\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}} \sum_{\substack{b_1,b_2\\ a_1,a_2}} p^n_{\underline{S}U\underline{V}X}(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n) \sqrt{\mathrm{tr}\{(\boldsymbol{I}-\pi^{2,\eta_3}_{u_2^n})\rho_{\underline{x}^n,\underline{s}^n}\}} \mathcal{G}\left(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n\right) \tag{119}$$

$$= \frac{96\exp\{\frac{n29\eta_5}{4}\}}{q^{k_1+k_2}|\mathcal{B}_1||\mathcal{B}_2|} \sum_{u_2^n} p^n_{U_2}(u_2^n) \sum_{\substack{\underline{s}^n,u_1^n\\ \underline{v}^n,\underline{x}^n}} \sum_{\substack{b_1,b_2\\ a_1,a_2}} p^n_{\underline{S}U_1\underline{V}\underline{X}|U_2}(\underline{s}^n,u_1^n,\underline{v}^n,\underline{x}^n|u_2^n) \sqrt{\mathrm{tr}\{(\boldsymbol{I}-\pi^{2,\eta_3}_{u_2^n})\rho_{\underline{x}^n,\underline{s}^n}\}} \mathbb{1}_{\left\{u_2^n \in T^n_{\eta_5}(p_U)\right\}} \tag{120}$$

$$\leq \frac{96\exp\{\frac{n29\eta_5}{4}\}}{q^{k_1+k_2}|\mathcal{B}_1||\mathcal{B}_2|} \sum_{u_2^n} p^n_{U_2}(u_2^n) \sum_{\substack{b_1,b_2\\ a_1,a_2}} \sqrt{\sum_{\substack{\underline{s}^n,u_1^n\\ \underline{v}^n,\underline{x}^n}} p^n_{\underline{S}U_1\underline{V}\underline{X}|U_2}(\underline{s}^n,u_1^n,\underline{v}^n,\underline{x}^n|u_2^n)\,\mathrm{tr}\{(\boldsymbol{I}-\pi^{2,\eta_3}_{u_2^n})\rho_{\underline{x}^n,\underline{s}^n}\}} \mathbb{1}_{\left\{u_2^n \in T^n_{\eta_5}(p_U)\right\}} \tag{121}$$

$$\leq \frac{96\exp\{\frac{n29\eta_5}{4}\}}{q^{k_1+k_2}|\mathcal{B}_1||\mathcal{B}_2|} \sum_{u_2^n} p^n_{U_2}(u_2^n) \sum_{\substack{b_1,b_2\\ a_1,a_2}} \sqrt{\mathrm{tr}\{(\boldsymbol{I}-\pi^{2,\eta_3}_{u_2^n})\sigma^2_{u_2^n}\}} \mathbb{1}_{\left\{u_2^n \in T^n_{\eta_5}(p_U)\right\}} \tag{122}$$

$$\leq \frac{96\exp\{\frac{n29\eta_5}{4}\}}{q^{k_1+k_2}|\mathcal{B}_1||\mathcal{B}_2|} \sum_{u_2^n} p^n_{U_2}(u_2^n) \sum_{\substack{b_1,b_2\\ a_1,a_2}} \sqrt{\exp\{-n([\eta_3-\eta_5]^2)\}} \leq 96\exp\left\{-n(\frac{(\eta_3-\eta_5)^2}{2} - \frac{29\eta_5}{4})\right\}. \tag{123}$$

where (i) (117) follows from the law of total probability and substituting the upper bound (116) in (113), (ii) (118) follows from [35, Chain of Inequalities 9.205 through to 9.209], (iii) (119) follows from the Markov chains $X_1U_1V_1 - S_1 - S_2U_2V_2X_2$ and $X_2U_2V_2 - S_1 - S_1U_1V_1X_1$ (v) (121) follows from concavity of the square root function, (vii) (122) follows from definition of $\sigma^2_{u_2^n}$ given as $\sigma^2_{u_2^n} = \sum_{\underline{x}^n,\underline{s}^n} p^n_{\underline{X}\underline{S}|U_2}(\underline{x}^n,\underline{s}^n|u_2^n)\rho_{\underline{x}^n,\underline{s}^n}$ and the other terms under the square root not depending on the variables of the summation, (viii) (123) follows from conditional quantum typicality (Lemma 4) since $u_2^n \in T^n_{\eta_5}(p_U)$.

APPENDIX K
PROOF OF PROPOSITION 8

From the definition of $\xi_4(\underline{m}|\underline{s}^n)$ in (41), we have

$$\overline{\xi}_4(\underline{m}) \leq 2 \sum_{\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}} \sum_{b_1,b_2} \sum_{a_1,a_2} \mathtt{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left(\left[\boldsymbol{I} - \pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{u_1^n}\pi^{\sigma,\eta_1}_{\underline{u}^n}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\right]\rho_{\underline{x}^n,\underline{s}^n}\right) \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}\right) P\left(\bigcap_{i=1}^4 \mathcal{G}_i^3 \middle| \mathcal{G}_{\underline{s}^n}\right), \tag{124}$$

where $\mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}\right), \mathcal{G}_i^3 : i \in [4]$ and $\mathcal{G}_{\underline{s}^n}$ are as defined in (111), (112). Substituting the upper bound in (116), we have

$$\overline{\xi}_4(\underline{m}) \leq 16 \sum_{\substack{\underline{s}^n,\underline{u}^n\\ \underline{x}^n,\underline{v}^n}} \sum_{\substack{b_1,b_2\\ a_1,a_2}} \frac{\mathtt{p}^n_{\underline{S}}(\underline{s}^n)}{q^{k_1+k_2}}\,\mathrm{tr}\left(\left[\boldsymbol{I} - \pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{u_1^n}\pi^{\sigma,\eta_1}_{\underline{u}^n}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\right]\rho_{\underline{x}^n\underline{s}^n}\right) \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}\right) \prod_{j=1}^2 \frac{p^n_{X_jU_j,V_j|S_j}(x_j^n,u_j^n,v_j^n|s_j^n)\exp\{\frac{n29\eta_5}{4}\}}{|\mathcal{B}_j|}$$

$$= 16\exp\{\frac{n29\eta_5}{2}\} \sum_{\underline{s}^n,\underline{u}^n} \sum_{\underline{v}^n,\underline{x}^n} p^n_{\underline{S}U\underline{V}X}(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n)\,\mathrm{tr}\left(\left[\boldsymbol{I} - \pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{u_1^n}\pi^{\sigma,\eta_1}_{\underline{u}^n}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\right]\rho_{\underline{x}^n,\underline{s}^n}\right) \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}\right) \tag{125}$$

$$\leq 16\exp\{\frac{n29\eta_5}{2}\} \sum_{\underline{u}^n} p^n_{\underline{U}}(\underline{u}^n) \sum_{\underline{s}^n} \sum_{\underline{v}^n,\underline{x}^n} p^n_{\underline{S}V\underline{X}|\underline{U}}(\underline{s}^n,\underline{v}^n,\underline{x}^n|\underline{u}^n)\mathrm{tr}\left(\left[\boldsymbol{I} - \pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{u_1^n}\pi^{\sigma,\eta_1}_{\underline{u}^n}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\right]\rho_{\underline{x}^n,\underline{s}^n}\right) \mathbb{1}_{\{\underline{u}^n \in T_{\eta_5}(p_{\underline{U}})\}}$$

$$\leq 16\exp\{\frac{n29\eta_5}{2}\} \sum_{\underline{u}^n} p^n_{\underline{U}}(\underline{u}^n)\mathrm{tr}\left(\left[\boldsymbol{I} - \pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{u_1^n}\pi^{\sigma,\eta_1}_{\underline{u}^n}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\right]\sigma_{\underline{u}^n}\right) \mathbb{1}_{\{\underline{u}^n \in T_{\eta_5}(p_{\underline{U}})\}} \tag{126}$$

where (i) (125) follows from the Markov chains $X_1U_1V_1 - S_1 - S_2U_2V_2X_2$ and $X_2U_2V_2 - S_1 - S_1U_1V_1X_1$ evident from (28), (ii) (126) follows from $\sum_{\underline{s}^n}\sum_{\underline{v}^n,\underline{x}^n} p^n_{\underline{S}\underline{V}\underline{X}|\underline{U}}(\underline{s}^n,\underline{v}^n,\underline{x}^n|\underline{u}^n)\rho_{\underline{x}^n,\underline{s}^n} = \sigma_{\underline{u}^n}$ and the fact that the other terms do not depend on the variable $\underline{s}^n, \underline{v}^n, \underline{x}^n$ of the summation. Repeated application of the 'measurement on close states' [35, Exercise 9.1.8] yields

$$\begin{aligned}\operatorname{tr}\left(\left[\boldsymbol{I}-\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{u_1^n}\pi^{\sigma,\eta_1}_{\underline{u}^n}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\right]\sigma_{\underline{u}^n}\right) &\leq \operatorname{tr}\left(\left[\boldsymbol{I}-\pi^{\sigma,\eta_1}_{\underline{u}^n}\right]\sigma_{\underline{u}^n}\right) + \left\|\sigma_{\underline{u}^n} - \pi^{\mu}_{\eta_4}\sigma_{\underline{u}^n}\pi^{\mu}_{\eta_4}\right\|_1 + \left\|\sigma_{\underline{u}^n} - \pi^{1,\eta_2}_{u_1^n}\sigma_{\underline{u}^n}\pi^{1,\eta_2}_{u_1^n}\right\|_1 \\ &\leq \operatorname{tr}\left(\left[\boldsymbol{I}-\pi^{\sigma,\eta_1}_{\underline{u}^n}\right]\sigma_{\underline{u}^n}\right) + 2\sqrt{\operatorname{tr}\left(\left[\boldsymbol{I}-\pi^{\mu}_{\eta_4}\right]\sigma_{\underline{u}^n}\right)} + 2\sqrt{\operatorname{tr}\left(\left[\boldsymbol{I}-\pi^{1,\eta_2}_{u_1^n}\right]\sigma_{\underline{u}^n}\right)} \quad (127)\end{aligned}$$

where the inequality in (127) follows from [35, Chain of Inequalities 9.205 through to 9.209]. Substituting the upper bound (127) into (126), we have $\overline{\xi}_4(\underline{m}) \leq 16\exp\{\frac{29}{4}n\eta_5\}\left[\tilde{\xi}_{41}(\underline{m}) + \tilde{\xi}_{42}(\underline{m}) + \tilde{\xi}_{43}(\underline{m})\right]$, where

$$\begin{aligned}\tilde{\xi}_{41}(\underline{m}) = \sum_{\underline{u}^n} p^n_{\underline{U}}(\underline{u}^n)\operatorname{tr}\left(\left[\boldsymbol{I}-\pi^{\sigma,\eta_1}_{\underline{u}^n}\right]\sigma_{\underline{u}^n}\right)\mathbb{1}_{\{\underline{u}^n\in T_{\eta_5}(p_{\underline{U}})\}},\ \tilde{\xi}_{42}(\underline{m}) = \sum_{\underline{u}^n} p^n_{\underline{U}}(\underline{u}^n)2\sqrt{\operatorname{tr}\left(\left[\boldsymbol{I}-\pi^{\mu}_{\eta_4}\right]\sigma_{\underline{u}^n}\right)}\mathbb{1}_{\{\underline{u}^n\in T_{\eta_5}(p_{\underline{U}})\}} \\ \text{and } \tilde{\xi}_{43}(\underline{m}) = \sum_{\underline{u}^n} p^n_{\underline{U}}(\underline{u}^n)2\sqrt{\operatorname{tr}\left(\left[\boldsymbol{I}-\pi^{1,\eta_2}_{u_1^n}\right]\sigma_{\underline{u}^n}\right)}\mathbb{1}_{\{\underline{u}^n\in T_{\eta_5}(p_{\underline{U}})\}}. \quad (128)\end{aligned}$$

Since $\underline{u}^n \in T^n_{\eta_5}(p_{\underline{U}})$, from the quantum conditional typicality (Lemma 4) and 'pinching' lemma (Property 15.2.7) - a version of it is proved from basic principles in Appendix F - we have

$$\tilde{\xi}_{41}(\underline{m}) + \tilde{\xi}_{42}(\underline{m}) \leq \sum_{\underline{u}^n} p^n_{\underline{U}}(\underline{u}^n)\exp\left\{-n\left(\eta_1-\eta_5\right)^2\right\} + \sum_{\underline{u}^n} p^n_{\underline{U}}(\underline{u}^n)2\exp\left\{-\frac{n}{2}\left(\eta_4-\eta_5\right)^2\right\}. \quad (129)$$

In regards to $\tilde{\xi}_{43}(\underline{m})$, following from (128) and using the concavity of the square root function, we have

$$\begin{aligned}\tilde{\xi}_{43}(\underline{m}) &\leq 2\sum_{u_1^n} p^n_{U_1}(u_1^n)\sum_{u_2^n} p^n_{U_2|U_1}(u_2^n|u_1^n)2\sqrt{\operatorname{tr}\left(\left[\boldsymbol{I}-\pi^{1,\eta_2}_{u_1^n}\right]\sigma_{\underline{u}^n}\right)}\mathbb{1}_{\{u_1^n\in T_{\eta_5}(p_{U_1})\}} \\ &\leq 2\sum_{u_1^n} p^n_{U_1}(u_1^n)\sqrt{\sum_{u_2^n} p^n_{U_2|U_1}(u_2^n|u_1^n)\operatorname{tr}\left(\left[\boldsymbol{I}-\pi^{1,\eta_2}_{u_1^n}\right]\sigma_{\underline{u}^n}\right)}\mathbb{1}_{\{u_1^n\in T_{\eta_5}(p_{U_1})\}} \leq 2\sum_{u_1^n} p^n_{U_1}(u_1^n)\sqrt{\operatorname{tr}\left(\left[\boldsymbol{I}-\pi^{1,\eta_2}_{u_1^n}\right]\sigma_{u_1^n}\right)}\mathbb{1}_{\{u_1^n\in T_{\eta_5}(p_{U_1})\}} \\ &\leq 2\sum_{u_1^n} p^n_{U_1}(u_1^n)\exp\left\{-\frac{n}{2}\left(\eta_2-\eta_5\right)^2\right\}\mathbb{1}_{\{u_1^n\in T_{\eta_5}(p_{U_1})\}} \leq 2\exp\left\{-\frac{n}{2}\left(\eta_2-\eta_5\right)^2\right\} \quad (130)\end{aligned}$$

where (130) follows from quantum conditional typicality (Lemma 4) since $u_1^n \in T_{\eta_5}(p_{U_1})$.Collating (130) and (129), we have

$$\overline{\xi}_4(\underline{m}) \leq \exp\left\{-n\left(\eta_1-\eta_5\right)^2\right\} + 2\exp\left\{-\frac{n}{2}\left(\eta_4-\eta_5\right)^2\right\} + 2\exp\left\{-\frac{n}{2}\left(\eta_2-\eta_5\right)^2\right\}. \quad (131)$$

APPENDIX L
PROOF OF PROPOSITION 9

We are required to derive an upper bound on $\overline{\xi}_5(\underline{m}) = \mathbb{E}\{\xi_5(\underline{m})\}$ and we proceed from (41). Let

$\mathcal{G}_{\underline{s}^n} \triangleq \{\underline{S}^n = \underline{s}^n\}, \mathcal{G}_1^5 \triangleq \left\{\begin{smallmatrix} V_j^n(a_j,m_{j2})=v_j^n, U_j^n(m_{j1},b_j)=u_j^n \\ \text{for } j=1,2,\ U_1^n(\hat{m}_{11},\hat{b}_1)=\hat{u}_1^n\end{smallmatrix}\right\}, \mathcal{G}_2^5 \triangleq \left\{\begin{smallmatrix} |\mathcal{L}_{j1}(m_{j1},s_j^n)|\geq L_{j1} \\ |\mathcal{L}_{j2}(\underline{m}_j,s_j^n)|\geq L_{j2}: j\in[2]\end{smallmatrix}\right\}, \mathcal{G}_3^5 \triangleq \left\{\begin{smallmatrix} B_j(m_{j1},s_j^n)=B_j^*=b_j \\ A_j(\underline{m}_j,s_j^n)=a_j: j\in[2]\end{smallmatrix}\right\}$

$\mathcal{G}_4^5 \triangleq \left\{\begin{smallmatrix} X_j^n(\underline{m}_j,s_j^n) \\ =x_j^n: j\in[2]\end{smallmatrix}\right\}, \mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{u}^n,\underline{v}^n \\ \hat{m}_{11},\hat{b}_1,b_1\end{smallmatrix}\right) \triangleq \mathbb{1}_{\left\{(s_j^n,u_j^n,v_j^n)\in T_{\eta_5}(p_{S_jU_jV_j}),s_j^n\in T_{\frac{\eta_5}{4}}(\mathtt{p}_{s_j}),(s_j^n,u_j^n)\in T_{\frac{\eta_5}{2}}(p_{S_jU_j}):j\in[2],(\hat{m}_{11},\hat{b}_1)\neq(m_{11},b_1)\right\}}$, we have

$$\overline{\xi}_5(\underline{m}) = \sum_{\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}} \sum_{\substack{\hat{\underline{m}}_1,\hat{\underline{b}}\\ \hat{u}_1^n}} \sum_{\substack{a_1,a_2\\ b_1,b_2}} \mathtt{p}^n_{\underline{S}}(\underline{s}^n)\, \mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{\hat{u}_1^n}\pi^{\sigma,\eta_1}_{\hat{u}_1^n u_2^n}\pi^{1,\eta_2}_{\hat{u}_1^n}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u_2^n}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u_2^n}\right)\mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n\\ \hat{m}_{11},\hat{b}_1,b_1}\right)P\left(\bigcap_{i=1}^4\mathcal{G}_i^5\,\middle|\,\mathcal{G}_{\underline{s}^n}\right), \text{ where (132)}$$

$$\mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n\\ \hat{m}_{11},\hat{b}_1,b_1}\right)P\left(\bigcap_{i=1}^4\mathcal{G}_i^5\,\middle|\,\mathcal{G}_{\underline{s}^n}\right) = \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n\\ \hat{m}_{11},\hat{b}_1,b_1}\right)P\left(\mathcal{G}_1^5\cap\mathcal{G}_2^5|\mathcal{G}_{\underline{s}^n}\right)P\left(\mathcal{G}_3^5\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^5\cap\mathcal{G}_2^5\right)P\left(\mathcal{G}_4^5\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^5\cap\mathcal{G}_2^5\cap\mathcal{G}_3^5\right)$$

$$\leq P\left(\mathcal{G}_1^5|\mathcal{G}_{\underline{s}^n}\right)\mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n\\ \hat{m}_{11},\hat{b}_1,b_1}\right)\prod_{j=1}^2\frac{p^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n)}{|\mathcal{L}_{j1}(m_{j1},s_j^n)||\mathcal{L}_{j2}(\mathfrak{m}_j,s_j^n)|} \tag{133}$$

$$= \frac{p^n_{U_1}(u_1^n)p^n_{U_2}(u_2^n)p^n_{U_1}(\hat{u}_1^n)}{q^{2n}}\mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n\\ \hat{m}_{11},\hat{b}_1,b_1}\right)\prod_{j=1}^2\frac{p^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n)}{L_{j1}L_{j2}} \tag{134}$$

$$\leq \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n\\ \hat{m}_{11},\hat{b}_1,b_1}\right)p^n_{U_1}(\hat{u}_1^n)\prod_{j=1}^2\frac{4q^n p^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n)p^n_{U_j}(u_j^n)\exp\left\{\frac{9n\eta_5}{2}\right\}}{q^{n+k_j}\exp\{n\left(\frac{\log|\mathcal{B}_j|}{n}-I(U_j;S_j)_\Upsilon+H(V_j|S_j,U_j)_\Upsilon\right)\}} \tag{135}$$

$$\leq \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n\\ \hat{m}_{11},\hat{b}_1,b_1}\right)p^n_{U_1}(\hat{u}_1^n)\prod_{j=1}^2\frac{1}{q^{k_j}|\mathcal{B}_j|}4p^n_{X_j,U_j,V_j|S_j}(x_j^n,u_j^n,v_j^n|s_j^n)\exp\left\{\frac{29n\eta_5}{4}\right\}, \tag{136}$$

where (i) (133) follows from the fact that (a) conditioned on the entire codebooks, $B_j(m_{j1},s_j^n)=B_j^*$ and $A_j(\mathfrak{m}_j,s_j^n)$ is uniformly distributed in $\mathcal{L}_{j1}(m_{j1},s_{j^n})$ and $\mathcal{L}_{j2}(\mathfrak{m}_j,s_j^n)$, and in particular conditionally independent of the realization of the codebooks (See Remark 4), which implies that $P\left(\mathcal{G}_3^5\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^5\cap\mathcal{G}_2^5\cap\mathcal{G}_3^5\right)=\frac{1}{|\mathcal{L}_{j1}(m_{j1},s_j^n)||\mathcal{L}_{j2}(\mathfrak{m}_j,s_j^n)|}$ and (b) $P\left(\mathcal{G}_4^5\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^5\cap\mathcal{G}_2^5\cap\mathcal{G}_3^5\right)=\prod_{j=1}^2 p^n_{X_j|U_jV_jS_j}(x_j^n|u_j^n,v_j^n,s_j^n)$, (ii) (134) follows from the facts that (a) the $\mathcal{U}_1-,\mathcal{U}_2-$codebooks are mutually independent and the codewords in the $\mathcal{U}_j-$codebook are mutually independent with distribution $p^n_{U_j}$, (b) two codewords in the two $\mathcal{V}_1-,\mathcal{V}_2-$UCC codes are pairwise independent (Lemma 5) and uniformly distributed in the $\mathcal{F}_q^n$ ambient space and (c) $|\mathcal{L}_{j1}(m_{j1},s_j^n)|\geq L_{j1}=\frac{1}{2}\exp\left\{n\left(\frac{\log|\mathcal{B}_j|}{n}-I(U_j;S_j)_\Upsilon-\frac{3\eta_5}{2}\right)\right\}$, $|\mathcal{L}_{j2}(\mathfrak{m}_j,s_j^n)|\geq L_{j2}=\frac{1}{2}\exp\left\{n\left(\frac{\log|\mathcal{B}_j|}{n}-\log q+H(V_j|S_j,U_j)_\Upsilon-3\eta_5\right)\right\}$, (iii) (135) follows from above definitions of L_{j1},L_{j2}, and (iv) (136) follows from the bound $\frac{p^n_{U_jV_jS_j}(u_j^n,v_j^n,s_j^n)}{p^n_{U_j}(u_j^n)p^n_{S_j}(s_j^n)}\geq\exp\{-n\left(H(S_j,U_j,V_j)_\Upsilon+\eta_5-H(S_j)_\Upsilon-H(U_j)_\Upsilon+\frac{\eta_5}{2}+\frac{\eta_5}{4}\right)\}=\exp\{-n\left(H(V_j|U_j,S_j)_\Upsilon-I(U_j;S_j)_\Upsilon+\frac{7\eta_5}{4}\right)\}$ for $(\underline{s}^n,\underline{u}^n,\underline{v}^n)$ satisfying the conditions of $\mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n\\ \hat{m}_{11},\hat{b}_1,b_1}\right)$, implying $1\leq\frac{p^n_{U_jV_jS_j}(u_j^n,v_j^n,s_j^n)}{p^n_{U_j}(u_j^n)p^n_{S_j}(s_j^n)}\exp\left\{n\left(H(V_j|U_j,S_j)_\Upsilon-I(U_j;S_j)_\Upsilon+\frac{7\eta_5}{4}\right)\right\}$. Substituting the upper bound (136) in (132) and noting that $\mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n\\ \hat{m}_{11},\hat{b}_1,b_1}\right)\leq\mathbb{1}_{\left\{u_2^n\in T_{\frac{\eta_5}{2}}(p_{U_2})\right\}}$, we have

$$\overline{\xi}_5(\underline{m})\leq\sum_{\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}}\sum_{\substack{\hat{\underline{m}}_1,\hat{\underline{b}}\\ \hat{u}_1^n}}\sum_{\substack{a_1,a_2\\ b_1,b_2}}\frac{\mathtt{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{\hat{u}_1^n}\pi^{\sigma,\eta_1}_{\hat{u}_1^n u_2^n}\pi^{1,\eta_2}_{\hat{u}_1^n}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u_2^n}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u_2^n}\right)\mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n\\ \hat{m}_{11},\hat{b}_1,b_1}\right)}{[p^n_{U_1}(\hat{u}_1^n)]^{-1}\exp\{-n8\eta_5\}q^{k_1+k_2}|\mathcal{B}_1||\mathcal{B}_2|}\prod_{j=1}^2 p^n_{X_jU_jV_j|S_j}(x_j^n,u_j^n,v_j^n|s_j^n)$$

$$\leq\sum_{u_2^n}p^n_{U_2}(u_2^n)\sum_{\substack{\underline{s}^n,u_1^n\\ \underline{v}^n,\underline{x}^n}}\sum_{\substack{\hat{\underline{m}}_1,\hat{\underline{b}}\\ \hat{u}_1^n}}\sum_{\substack{a_1,a_2\\ b_1,b_2}}\mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{\hat{u}_1^n}\pi^{\sigma,\eta_1}_{\hat{u}_1^n u_2^n}\pi^{1,\eta_2}_{\hat{u}_1^n}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u_2^n}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u_2^n}\right)\frac{p^n_{\underline{S}U_1\underline{V}\underline{X}|U_2}(\underline{s}^n,u_1^n,\underline{v}^n,\underline{x}^n|u_2^n)p^n_{U_1}(\hat{u}_1^n)\mathbb{1}_{\{u_2^n\in T_{\frac{\eta_5}{2}}(p_{U_2})\}}}{q^{k_1+k_2}|\mathcal{B}_1||\mathcal{B}_2|\exp\{-8n\eta_5\}}$$

$$= \sum_{u_2^n} p_{U_2}^n(u_2^n) \sum_{\substack{\underline{s}^n,u_1^n \\ \underline{v}^n,\underline{x}^n}} \sum_{\substack{\hat{\underline{m}}_1,\hat{\underline{b}} \\ \hat{u}_1^n}} \mathrm{tr}\left(\pi_{\eta_4}^{\mu} \pi_{\hat{u}_1^n}^{1,\eta_2} \pi_{\hat{u}_1^n u_2^n}^{\sigma,\eta_1} \pi_{\hat{u}_1^n}^{1,\eta_2} \pi_{\eta_4}^{\mu} \pi_{u_2^n}^{2,\eta_3} \rho_{\underline{x}^n,\underline{s}^n} \pi_{u_2^n}^{2,\eta_3}\right) \frac{p^n_{\underline{S}U_1\underline{V}\underline{X}|U_2}(\underline{s}^n,u_1^n,\underline{v}^n,\underline{x}^n|u_2^n) p_{U_1}^n(\hat{u}_1^n) \mathbb{1}_{\{u_2^n \in T_{\frac{\eta_5}{2}}(p_{U_2})\}}}{\exp\{-8n\eta_5\}} \tag{137}$$

$$\leq \sum_{u_2^n} p_{U_2}^n(u_2^n) \sum_{\hat{\underline{m}}_1} \sum_{\hat{\underline{b}}} \sum_{\hat{u}_1^n} \mathrm{tr}\left(\pi_{\eta_4}^{\mu} \pi_{\hat{u}_1^n}^{1,\eta_2} \pi_{\hat{u}_1^n u_2^n}^{\sigma,\eta_1} \pi_{\hat{u}_1^n}^{1,\eta_2} \pi_{\eta_4}^{\mu} \left[\pi_{u_2^n}^{2,\eta_3} \sigma_{u_2^n}^2 \pi_{u_2^n}^{2,\eta_3}\right]\right) p_{U_1}^n(\hat{u}_1^n) \mathbb{1}_{\{u_2^n \in T_{\frac{\eta_5}{2}}(p_{U_2})\}} \exp\{8n\eta_5\} \tag{138}$$

$$\leq \exp\{-n\left(H(Y|U_2)_\Upsilon - \frac{17\eta_5}{2} - \eta_3\right)\} \sum_{u_2^n} p_{U_2}^n(u_2^n) \sum_{\hat{\underline{m}}_1} \sum_{\hat{\underline{b}}} \sum_{\hat{u}_1^n} \mathrm{tr}\left(\pi_{\eta_4}^{\mu} \pi_{\hat{u}_1^n}^{1,\eta_2} \pi_{\hat{u}_1^n u_2^n}^{\sigma,\eta_1} \pi_{\hat{u}_1^n}^{1,\eta_2} \pi_{\eta_4}^{\mu} [\pi_{u_2^n}^{2,\eta_3}]\right) p_{U_1}^n(\hat{u}_1^n) \tag{139}$$

$$= \exp\{-n\left(H(Y|U_2)_\Upsilon - \frac{17\eta_5}{2} - \eta_3 - R_{11} - B_1\right)\} \sum_{u_2^n} p_{U_2}^n(u_2^n) \sum_{\hat{u}_1^n} \mathrm{tr}\left(\pi_{\hat{u}_1^n u_2^n}^{\sigma,\eta_1} \left[\pi_{\hat{u}_1^n}^{1,\eta_2} \pi_{\eta_4}^{\mu} \pi_{u_2^n}^{2,\eta_3} \pi_{\eta_4}^{\mu} \pi_{\hat{u}_1^n}^{1,\eta_2}\right]\right) p_{U_1}^n(\hat{u}_1^n) \tag{140}$$

$$\leq \exp\{-n\left(H(Y|U_2)_\Upsilon - \frac{17\eta_5}{2} - \eta_3 - R_{11} - B_1\right)\} \sum_{u_2^n} p_{U_2}^n(u_2^n) \sum_{\hat{u}_1^n} \mathrm{tr}\left(\pi_{\hat{u}_1^n u_2^n}^{\sigma,\eta_1} [\boldsymbol{I}]\right) p_{U_1}^n(\hat{u}_1^n) \tag{141}$$

$$= \exp\{-n\Big(H(Y|U_2)_\Upsilon - \frac{17\eta_5}{2} - \eta_3 - R_{11} - B_1\Big)\} \sum_{u_2^n} p_{U_2}^n(u_2^n) \sum_{\hat{u}_1^n} \mathrm{tr}\left(\pi_{\hat{u}_1^n u_2^n}^{\sigma,\eta_1}\right) p_{U_1}^n(\hat{u}_1^n) \left(\mathbb{1}_{\left\{\substack{(\hat{u}_1^n,u_2^n) \\ \notin T_{\eta_1}^n(p_{\underline{U}})}\right\}} + \mathbb{1}_{\left\{\substack{(\hat{u}_1^n,u_2^n) \\ \in T_{\eta_1}^n(p_{\underline{U}})}\right\}}\right),$$

where (i) (137) follows since terms in the summand are invariant to the choice of $a_j \in \mathcal{F}_q^{k_j}$ and $b_j \in \mathcal{B}_j$ for $j \in [2]$, (ii) (138) follows from $\sum_{\underline{s}^n,u_1^n,\underline{v}^n,\underline{x}^n} p^n_{\underline{S}U_1\underline{V}\underline{X}|U_2}(\underline{s}^n,u_1^n,\underline{v}^n,\underline{x}^n|u_2^n)\rho_{\underline{x}^n,\underline{s}^n} = \sigma_{u_2^n}^2$, (iii) (139) follows from quantum conditional typicality (Lemma 4) which states that $\left[\pi_{u_2^n}^{2,\eta_3} \sigma_{u_2^n}^2 \pi_{u_2^n}^{2,\eta_3}\right] \leq \exp\{-n(H(Y|U_2)_\Upsilon - 2\eta_3 - \eta_5)\}\pi_{u_2^n}^{2,\eta_3}$ since $u_2^n \in T_{\frac{\eta_5}{2}}(p_{U_2})$, (iv) (140) follows from cyclicity of trace and the fact that the terms in the summand are invariant to $\hat{m}_{11} \in [\mathcal{M}_{11}]$ and $b_1 \in [\mathcal{B}_1]$, (v) (141) follows from $\pi_{\hat{u}_1^n}^{1,\eta_2} \pi_{\eta_4}^{\mu} \pi_{u_2^n}^{2,\eta_3} \pi_{\eta_4}^{\mu} \pi_{\hat{u}_1^n}^{1,\eta_2} \leq \boldsymbol{I}$. Since $(\hat{u}_1^n, u_2^n) \notin T_{\eta_1}^n(p_{\underline{U}})$ implies $\pi_{\hat{u}_1^n u_2^n}^{\sigma,\eta_1} = 0$ from Lemma 4, we have

$$\bar{\xi}_5 \leq \exp\{-n\left(H(Y|U_2)_\Upsilon - \frac{17\eta_5}{2} - \eta_3 - R_{11} - B_1\right)\} \sum_{u_2^n} p_{U_2}^n(u_2^n) \sum_{\hat{u}_1^n} \mathrm{tr}\left(\pi_{\hat{u}_1^n u_2^n}^{\sigma,2\eta_1}\right) p_{U_1}^n(\hat{u}_1^n) \mathbb{1}_{\left\{(\hat{u}_1^n,u_2^n) \in T_{\eta_1}^n(p_{\underline{U}})\right\}}$$

$$\leq \exp\{-n\Big(H(Y|U_2)_\Upsilon - H(Y|U_1,U_2)_\Upsilon - 2\eta_1 - \frac{17\eta_5}{2} - \eta_3 - R_{11} - B_1\Big)\} \sum_{u_2^n,\hat{u}_1^n} p_{U_2}^n(u_2^n) p_{U_1}^n(\hat{u}_1^n) \mathbb{1}_{\left\{(\hat{u}_1^n,u_2^n) \in T_{\eta_1}^n(p_{\underline{U}})\right\}} \tag{142}$$

$$\leq \exp\left\{-n\left(I(U_1;U_2)_\Upsilon + I(Y;U_1|U_2)_\Upsilon - 5\eta_1 - \frac{17\eta_5}{2} - \eta_3 - R_{11} - B_1\right)\right\}, \tag{143}$$

where (142) follows from $\mathrm{tr}\left(\pi_{\hat{u}_1^n u_2^n}^{\sigma,2\eta_1}\right) \leq \exp\{n(H(Y|U_1,U_2)_\Upsilon - 2\eta_1)\}$ whenever $(\hat{u}_1^n, u_2^n) \in T_{\eta_1}^n(p_{\underline{U}})$ as stated in Lemma 4 and (143) follows from the fact that $|T_{\eta_1}^n(p_{\underline{U}})| \leq \exp\{n[S(U_1,U_2)_\Upsilon + \eta_1]\}$, and $(\hat{u}_1^n, u_2^n) \in T_{\eta_1}^n(p_{\underline{U}})$ implies that $p_{U_1}(\hat{u}_1^n) \leq \exp\{-n[H(U_1) - \eta_1]\}$ and $p_{U_2}(u_2^n) \leq \exp\{-n[H(U_2) - \eta_1]\}$.

Appendix M
Proof of Proposition 10

We now derive an upper bound on $\bar{\xi}_6(\underline{m}) = \mathbb{E}\{\xi_6(\underline{m})\}$, starting from the definition of $\xi_6(\underline{m})$ in (42). As in [37], our technique for this term is different from that used for analyzing $\bar{\xi}_5(\underline{m})$ motivating us to provide a detailed sequence of steps. Let

$$\mathcal{G}_{\underline{s}^n} \triangleq \mathbb{1}_{\{\underline{S}^n = \underline{s}^n\}}, \mathcal{G}_1^6 \triangleq \mathbb{1}_{\left\{\substack{V_j^n(a_j,m_{j2})=v_j^n, U_j^n(m_{j1},b_j)=u_j^n \\ \text{for } j=1,2,\ U_2^n(\hat{m}_{21},\hat{b}_2)=\hat{u}_2^n}\right\}}, \mathcal{G}_2^6 \triangleq \mathbb{1}_{\left\{\substack{|\mathcal{L}_{j1}(m_{j1},s_j^n)| \geq L_{j1} \\ |\mathcal{L}_{j2}(\underline{m}_j,s_j^n)| \geq L_{j2}: j\in[2]}\right\}}, \mathcal{G}_3^6 \triangleq \mathbb{1}_{\left\{\substack{B_j(m_{j1},s_j^n)=B_j^*=b_j \\ A_j(\underline{m}_j,s_j^n)=a_j: j\in[2]}\right\}}$$
$$\mathcal{G}_4^6 \triangleq \mathbb{1}_{\left\{\substack{X_j^n(\underline{m}_j,s_j^n) \\ =x_j^n: j\in[2]}\right\}}, \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n,b_2 \\ \hat{u}_2^n,\hat{m}_{21},\hat{b}_2}\right) \triangleq \mathbb{1}_{\left\{\substack{(s_j^n,u_j^n,v_j^n)\in T_{\eta_5}(p_{S_jU_jV_j}),(s_j^n,u_j^n)\in T_{\frac{\eta_5}{2}}(p_{S_jU_j}) \\ s_j^n \in T_{\frac{\eta_5}{4}}(\mathtt{p}_{s_j}): j\in[2],(u_1^n,\hat{u}_2^n)\in T_{\eta_1}^n(p_{U_1U_2}),(\hat{m}_{21},\hat{b}_2)\neq(m_{21},b_2)}\right\}}, \text{ we have}$$

$$\bar{\xi}_6(\underline{m}) = \sum_{\substack{\underline{s}^n,\underline{u}^n \\ \underline{v}^n,\underline{x}^n}} \sum_{\substack{\hat{\underline{m}}_1,\hat{\underline{b}} \\ \hat{u}_2^n}} \sum_{\substack{a_1,a_2 \\ b_1,b_2}} \mathtt{p}_{\underline{S}}^n(\underline{s}^n) \mathrm{tr}\left(\pi_{\eta_4}^{\mu} \pi_{u_1^n}^{1,\eta_2} \pi_{u_1^n \hat{u}_2^n}^{\sigma,\eta_1} \pi_{u_1^n}^{1,\eta_2} \pi_{\eta_4}^{\mu} \pi_{u_2^n}^{2,\eta_3} \rho_{\underline{x}^n,\underline{s}^n} \pi_{u_2^n}^{2,\eta_3}\right) \mathcal{G}\left(\substack{\underline{s}^n,\underline{u}^n,\underline{v}^n,b_2 \\ \hat{u}_2^n,\hat{m}_{21},\hat{b}_2}\right) P\left(\bigcap_{i=1}^4 \mathcal{G}_i^6 \,\middle|\, \mathcal{G}_{\underline{s}^n}\right), \text{ where} \tag{144}$$

$$\mathcal{G}\begin{pmatrix}\underline{s}^n,\underline{u}^n,\underline{v}^n,b_2\\ \hat{u}_2^n,\hat{m}_{21},\hat{b}_2\end{pmatrix}P\left(\bigcap_{i=1}^{4}\mathcal{G}_i^6\,\middle|\,\mathcal{G}_{\underline{s}^n}\right)=\mathcal{G}\begin{pmatrix}\underline{s}^n,\underline{u}^n,\underline{v}^n,b_2\\ \hat{u}_2^n,\hat{m}_{21},\hat{b}_2\end{pmatrix}P\left(\mathcal{G}_1^6\cap\mathcal{G}_2^6|\mathcal{G}_{\underline{s}^n}\right)P\left(\mathcal{G}_3^6\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^6\cap\mathcal{G}_2^6\right)P\left(\mathcal{G}_4^6\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^6\cap\mathcal{G}_2^6\cap\mathcal{G}_3^6\right)$$

$$\leq P\left(\mathcal{G}_1^6|\mathcal{G}_{\underline{s}^n}\right)\mathcal{G}\begin{pmatrix}\underline{s}^n,\underline{u}^n,\underline{v}^n,b_2\\ \hat{u}_2^n,\hat{m}_{21},\hat{b}_2\end{pmatrix}\prod_{j=1}^{2}\frac{p^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n)}{|\mathcal{L}_{j1}(m_{j1},s_j^n)||\mathcal{L}_{j2}(\underline{m}_j,s_j^n)|} \tag{145}$$

$$=\frac{p^n_{U_1}(u_1^n)p^n_{U_2}(u_2^n)p^n_{U_2}(\hat{u}_2^n)}{q^{2n}}\mathcal{G}\begin{pmatrix}\underline{s}^n,\underline{u}^n,\underline{v}^n,b_2\\ \hat{u}_2^n,\hat{m}_{21},\hat{b}_2\end{pmatrix}\prod_{j=1}^{2}\frac{p^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n)}{L_{j1}L_{j2}} \tag{146}$$

$$\leq\mathcal{G}\begin{pmatrix}\underline{s}^n,\underline{u}^n,\underline{v}^n,b_2\\ \hat{u}_2^n,\hat{m}_{21},\hat{b}_2\end{pmatrix}p^n_{U_2}(\hat{u}_2^n)\prod_{j=1}^{2}\frac{4q^n p^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n)p^n_{U_j}(u_j^n)\exp\left\{\frac{9n\eta_5}{2}\right\}}{q^{n+k_j}\exp\{n\left(\frac{\log|\mathcal{B}_j|}{n}-I(U_j;S_j)_\Upsilon+H(V_j|S_j,U_j)_\Upsilon\right)\}} \tag{147}$$

$$\leq\mathcal{G}\begin{pmatrix}\underline{s}^n,\underline{u}^n,\underline{v}^n,b_2\\ \hat{u}_2^n,\hat{m}_{21},\hat{b}_2\end{pmatrix}p^n_{U_2|U_1}(\hat{u}_2^n|u_1^n)\prod_{j=1}^{2}\frac{\exp\{-nI(U_1;U_2)_\Upsilon\}}{q^{k_j}|\mathcal{B}_j|\exp\{-3n\eta_1\}}4p^n_{X_j,U_j,V_j|S_j}(x_j^n,u_j^n,v_j^n|s_j^n)\exp\left\{\frac{29n\eta_5}{4}\right\}, \tag{148}$$

where (i) (145) follows from the fact that (a) conditioned on the entire codebooks, $B_j(m_{j1},s_j^n)=B_j^*$ and $A_j(\underline{m}_j,s_j^n)$ is uniformly distributed in $\mathcal{L}_{j1}(m_{j1},s_{j^n})$ and $\mathcal{L}_{j2}(\underline{m}_j,s_j^n)$, and in particular conditionally independent of the realization of the codebooks (See Remark 4), which implies that $P\left(\mathcal{G}_3^6\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^6\cap\mathcal{G}_2^6\cap\mathcal{G}_3^6\right)=\frac{1}{|\mathcal{L}_{j1}(m_{j1},s_j^n)||\mathcal{L}_{j2}(\underline{m}_j,s_j^n)|}$ and (b) $P\left(\mathcal{G}_4^6\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^6\cap\mathcal{G}_2^6\cap\mathcal{G}_3^6\right)=\prod_{j=1}^{2}p^n_{X_j|U_jV_jS_j}(x_j^n|u_j^n,v_j^n,s_j^n)$, (ii) (146) follows from the facts that (a) the $\mathcal{U}_1-,\mathcal{U}_2-$codebooks are mutually independent and the codewords in the $\mathcal{U}_j-$codebook are mutually independent with distribution $p^n_{U_j}$, (b) two codewords in the two $\mathcal{V}_1-,\mathcal{V}_2-$UCC codes are pairwise independent (Lemma 5) and uniformly distributed in the $\mathcal{F}_q^n$ ambient space and (c) $|\mathcal{L}_{j1}(m_{j1},s_j^n)|\geq L_{j1}=\frac{1}{2}\exp\left\{n\left(\frac{\log|\mathcal{B}_j|}{n}-I(U_j;S_j)_\Upsilon-\frac{3\eta_5}{2}\right)\right\}$, $|\mathcal{L}_{j2}(\underline{m}_j,s_j^n)|\geq L_{j2}=\frac{1}{2}\exp\left\{n\left(\frac{\log|\mathcal{B}_j|}{n}-\log q+H(V_j|S_j,U_j)_\Upsilon-3\eta_5\right)\right\}$, (iii) (147) follows from above definitions of L_{j1},L_{j2}, and (iv) (148) follows from the bounds $\frac{p^n_{U_jV_jS_j}(u_j^n,v_j^n,s_j^n)}{p^n_{U_j}(u_j^n)p^n_{S_j}(s_j^n)}\geq\exp\{-n\left(H(S_j,U_j,V_j)_\Upsilon+\eta_5-H(S_j)_\Upsilon-H(U_j)_\Upsilon+\frac{\eta_5}{2}+\frac{\eta_5}{4}\right)\}=\exp\{-n\left(H(V_j|U_j,S_j)_\Upsilon-I(U_j;S_j)_\Upsilon+\frac{7\eta_5}{4}\right)\}$ and $\frac{p^n_{U_1\hat{U}_2}(u_1^n,\hat{u}_2^n)}{p^n_{U_2}(\hat{u}_2^n)p^n_{U_1}(u_1^n)}\geq\frac{\exp\{n(I(U_1;U_2)_\Upsilon)\}}{\exp\{3n\eta_1\}}$ for $(\underline{s}^n,\underline{u}^n,\underline{v}^n)$ satisfying the conditions of $\mathcal{G}\begin{pmatrix}\underline{s}^n,\underline{u}^n,\underline{v}^n,b_2\\ \hat{u}_2^n,\hat{m}_{21},\hat{b}_2\end{pmatrix}$, implying $1\leq\frac{p^n_{U_jV_jS_j}(u_j^n,v_j^n,s_j^n)}{p^n_{U_j}(u_j^n)p^n_{S_j}(s_j^n)}\exp\left\{n\left(H(V_j|U_j,S_j)_\Upsilon-I(U_j;S_j)_\Upsilon+\frac{7\eta_5}{4}\right)\right\}$ and $1\leq\frac{p^n_{U_1\hat{U}_2}(u_1^n,\hat{u}_2^n)\exp\{-n(I(U_1;U_2)_\Upsilon)\}}{p^n_{U_2}(\hat{u}_2^n)p^n_{U_1}(u_1^n)\exp\{-3n\eta_1\}}$. The inclusion of the event $\{(u_1^n,\hat{u}_2^n)\in T^n_{\eta_1}(p_{U_1U_2})\}$ in the definition of $\mathcal{G}\begin{pmatrix}\underline{s}^n,\underline{u}^n,\underline{v}^n,b_2\\ \hat{u}_2^n,\hat{m}_{21},\hat{b}_2\end{pmatrix}$ is justified by the fact that the operator $\pi^{\sigma,\eta_1}_{u_1^n\hat{u}_2^n}=0$ whenever $\{(u_1^n,\hat{u}_2^n)\notin T^n_{\eta_1}(p_{U_1U_2})\}$. The positive terms in the summand of $\overline{\xi}_6(\underline{m})$ therefore remain unaltered. Before we substitute the bound (148) in (144), we make the following observations. Note that if $(u_1^n,\hat{u}_2^n)\notin T_{\eta_1}(p_{\underline{U}})$, Lemma 4 guarantees the operator $\pi^{\sigma,\eta_1}_{u_1^n,\hat{u}_2^n}=0$. For $(u_1^n,\hat{u}_2^n)\in T_{\eta_1}(p_{\underline{U}})$, the commutativity of $\sigma_{u_1^n,\hat{u}_2^n}$ and $\pi^{\sigma,\eta_1}_{u_1^n,\hat{u}_2^n}$ and Lemma 4 ensure

$$\begin{aligned}\pi^{\sigma,\eta_1}_{u_1^n,\hat{u}_2^n}\leq\pi^{\sigma,2\eta_1}_{u_1^n,\hat{u}_2^n}&\leq\exp\{n(H(Y|\underline{U})_\Upsilon+6\eta_1)\}\,\pi^{\sigma,2\eta_1}_{u_1^n,\hat{u}_2^n}\sigma_{u_1^n,\hat{u}_2^n}\pi^{\sigma,2\eta_1}_{u_1^n,\hat{u}_2^n}=\exp\{n(H(Y|\underline{U})_\Upsilon+6\eta_1)\}\sqrt{\sigma_{u_1^n,\hat{u}_2^n}}\pi^{\sigma,\eta_1}_{u_1^n,\hat{u}_2^n}\pi^{\sigma,\eta_1}_{u_1^n,\hat{u}_2^n}\sqrt{\sigma_{u_1^n,\hat{u}_2^n}}\\&=\exp\left\{n\left(H(Y|\underline{U})_\Upsilon+6\eta_1\right)\right\}\sqrt{\sigma_{u_1^n,\hat{u}_2^n}}\pi^{\sigma,\eta_1}_{u_1^n,\hat{u}_2^n}\sqrt{\sigma_{u_1^n,\hat{u}_2^n}}\leq\exp\left\{n\left(H(Y|U_1,U_2)_\Upsilon+6\eta_1\right)\right\}\sqrt{\sigma_{u_1^n,\hat{u}_2^n}}\boldsymbol{I}\sqrt{\sigma_{u_1^n,\hat{u}_2^n}}\\&=\exp\left\{n\left(H(Y|U_1,U_2)_\Upsilon+6\eta_1\right)\right\}\sigma_{u_1^n,\hat{u}_2^n}.\end{aligned} \tag{149}$$

Now substituting the upper bound (148) in (144), we have

$$\overline{\xi}_6(\underline{m})\leq\sum_{\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}}\sum_{\substack{\hat{\underline{m}}_1,\hat{\underline{b}}\\ \hat{u}_2^n}}\sum_{\substack{a_1,a_2\\ b_1,b_2}}\frac{\mathrm{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{u_1^n}\pi^{\sigma,\eta_1}_{u_1^n\hat{u}_2^n}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u_2^n}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u_2^n}\right)\mathcal{G}\begin{pmatrix}\underline{s}^n,\underline{u}^n,\underline{v}^n,b_2\\ \hat{u}_2^n,\hat{m}_{21},\hat{b}_2\end{pmatrix}}{\exp\{nI(U_1;U_2)\}[p^n_{U_2|U_1}(\hat{u}_2^n|u_1^n)]^{-1}\exp\{-n8\eta_5-3n\eta_1\}q^{k_1+k_2}|\mathcal{B}_1||\mathcal{B}_2|}\prod_{j=1}^{2}p^n_{X_jU_jV_j|S_j}(x_j^n,u_j^n,v_j^n|s_j^n)$$

$$\leq\sum_{\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}}\sum_{\substack{\hat{\underline{m}}_1,\hat{\underline{b}}\\ \hat{u}_2^n}}\sum_{\substack{a_1,a_2\\ b_1,b_2}}\frac{p^n_{U_2|U_1}(\hat{u}_2^n|u_1^n)\mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{u_1^n}\sigma_{u_1^n,\hat{u}_2^n}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u_2^n}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u_2^n}\right)\mathbb{1}_{\left\{(u_1^n,\hat{u}_2^n)\in T_{\eta_1}(p_{\underline{U}})\right\}}}{\exp\{-n\left(H(Y|U_1,U_2)_\Upsilon-I(U_1;U_2)_\Upsilon+9\eta_1+8\eta_5\right)\}q^{k_1+k_2}|\mathcal{B}_1||\mathcal{B}_2|}p^n_{\underline{SUVX}}(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n) \tag{150}$$

$$\leq \sum_{\underline{s}^n,\underline{u}^n}\sum_{\underline{v}^n,\underline{x}^n}\sum_{\hat{\underline{m}}_1,\hat{\underline{b}}}\sum_{\substack{a_1,a_2\\b_1,b_2}}\frac{\mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{u_1^n}\sigma^{1}_{u_1^n}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u_2^n}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u_2^n}\right)\mathbb{1}_{\left\{u_1^n\in T_{\eta_1}(p_{U_1})\right\}}}{\exp\{-n\left(H(Y|U_1,U_2)_{\Upsilon}-I(U_1;U_2)+9\eta_1+8\eta_5\right)\}q^{k_1+k_2}|\mathcal{B}_1||\mathcal{B}_2|}p^n_{\underline{SUVX}}(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n) \quad (151)$$

$$=\sum_{\underline{s}^n,\underline{u}^n}\sum_{\underline{v}^n,\underline{x}^n}\frac{\mathrm{tr}\left(\pi^{1,\eta_2}_{u_1^n}\sigma^{1}_{u_1^n}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u_2^n}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u_2^n}\pi^{\mu}_{\eta_4}\right)\mathbb{1}_{\left\{u_1^n\in T_{\eta_1}(p_{U_1})\right\}}}{\exp\{-n\left(H(Y|U_1,U_2)_{\Upsilon}-I(U_1;U_2)+9\eta_1+8\eta_5\right)\}}p^n_{\underline{SUVX}}(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n) \quad (152)$$

$$\leq\sum_{\underline{s}^n,\underline{u}^n}\sum_{\underline{v}^n,\underline{x}^n}\frac{\mathrm{tr}\left(\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u_2^n}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u_2^n}\pi^{\mu}_{\eta_4}\right)\mathbb{1}_{\left\{u_1^n\in T_{\eta_1}(p_{U_1})\right\}}\exp\{-nI(U_1;U_2)\}}{\exp\{-n\left(H(Y|U_1,U_2)_{\Upsilon}-H(Y|U_1)_{\Upsilon}+10\eta_1+\eta_2+8\eta_5+R_{21}+B_2\right)\}}p^n_{\underline{SUVX}}(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n) \quad (153)$$

$$=\sum_{\underline{s}^n,\underline{u}^n}\sum_{\underline{v}^n,\underline{x}^n}\frac{\mathrm{tr}\left(\left[\pi^{2,\eta_3}_{u_2^n}\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u_2^n}\right]\rho_{\underline{x}^n,\underline{s}^n}\right)p^n_{\underline{SUVX}}(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n)}{\exp\{-n\left(-I(Y,U_1;U_2)_{\Upsilon}+7\eta_1+\eta_2+8\eta_5+R_{21}+B_2\right)\}} \quad (154)$$

$$\leq\sum_{\underline{s}^n,\underline{u}^n}\sum_{\underline{v}^n,\underline{x}^n}\frac{\mathrm{tr}\left(\boldsymbol{I}\rho_{\underline{x}^n,\underline{s}^n}\right)p^n_{\underline{SUVX}}(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n)\exp\left\{n\left(\eta_2+R_{21}+B_2\right)\right\}}{\exp\{-n\left(-I(Y,U_1;U_2)_{\Upsilon}+7\eta_1+8\eta_5\right)\}} \quad (155)$$

$$\leq\sum_{\underline{s}^n,\underline{u}^n}\sum_{\underline{v}^n,\underline{x}^n}\frac{p^n_{\underline{SUVX}}(\underline{s}^n,\underline{u}^n,\underline{v}^n,\underline{x}^n)\exp\left\{n\left(R_{21}+B_2\right)\right\}}{\exp\{n(I(Y,U_1;U_2)_{\Upsilon}-7\eta_1-8\eta_5-\eta_2)\}}\leq\exp\{-n(I(Y,U_1;U_2)_{\Upsilon}-7\eta_1-\eta_2-8\eta_5-R_{21}-B_2)\} \quad (156)$$

where (i) (150) follows from substituting the upper bound (149) and the Markov chains $X_1U_1V_1-S_1-S_2U_2V_2X_2$ and $X_2U_2V_2-S_1-S_1U_1V_1X_1$, (ii) (151) follows from $\sum_{\hat{u}_2^n}p^n_{U_2|U_1}(\hat{u}_2^n|u_1^n)\sigma_{u_1^n\hat{u}_2^n}=\sigma^1_{u_1^n}$ and the fact that none of the other terms in (151) depend on $\hat{u}_2^n$, (iii) (152) follows from cyclicity of the trace and the fact that none of the terms in the summand therein depend on $\hat{m}_1,\hat{b},a_1,a_2,b_1,b_2$, (iv) (153) follows from $\pi^{1,\eta_2}_{u_1^n}\sigma^1_{u_1^n}\pi^{1,\eta_2}_{u_1^n}\leq\exp\left\{-n\left(H(Y|U_1)_{\Upsilon}-\eta_1-\eta_2\right)\right\}\pi^{1,\eta_2}_{u_1^n}$ which holds since $u_1^n\in T_{\eta_1}(p_{U_1})$ (Lemma 4) (v) (154) follows from cyclicity of the trace, (vi) (155) follows from the operator dominance $\pi^{2,\eta_3}_{u_2^n}\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{u_1^n}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u_2^n}\leq\boldsymbol{I}$ and finally (vii) (156) from $\mathrm{tr}\big(\rho_{\underline{x}^n\underline{s}^n}\big)=1$.

Appendix N
Proof of Proposition 11

We now derive an upper bound on $\overline{\xi}_7(\underline{m})$ starting from (43). Let Let

$\mathcal{G}_{\underline{s}^n}\triangleq\{\underline{S}^n=\underline{s}^n\},\mathcal{G}_1^7\triangleq\left\{\begin{smallmatrix}V_j^n(a_j,m_{j2})=v_j^n,U_j^n(m_{j1},b_j)=u_j^n\\U_j^n(\hat{m}_{j1},\hat{b}_j)=\hat{u}_j^n\text{ for }j=1,2\end{smallmatrix}\right\},\mathcal{G}_2^7\triangleq\left\{\begin{smallmatrix}|\mathcal{L}_{j1}(m_{j1},s_j^n)|\geq L_{j1}\\|\mathcal{L}_{j2}(\underline{m}_j,s_j^n)|\geq L_{j2}:j\in[2]\end{smallmatrix}\right\},\mathcal{G}_3^7\triangleq\left\{\begin{smallmatrix}B_j(m_{j1},s_j^n)=B_j^*=b_j\\A_j(\underline{m}_j,s_j^n)=a_j:j\in[2]\end{smallmatrix}\right\}$

$\mathcal{G}_4^7\triangleq\left\{\begin{smallmatrix}X_j^n(\underline{m}_j,s_j^n)\\=x_j^n:j\in[2]\end{smallmatrix}\right\},\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{u}^n,\underline{b}\\\underline{v}^n,\hat{\underline{m}}_1,\hat{\underline{b}}\end{smallmatrix}\right)\triangleq\mathbb{1}_{\left\{(s_j^n,u_j^n,v_j^n)\in T_{\eta_5}(p_{S_jU_jV_j}),s_j^n\in T_{\frac{\eta_5}{4}}(\mathrm{p}_{s_j}),(s_j^n,u_j^n)\in T_{\frac{\eta_5}{2}}(p_{S_jU_j}),(\hat{m}_{j1},\hat{b}_j)\neq(m_{j1},b_j):j=1,2\right\}}$, we have

$$\overline{\xi}_7(\underline{m})=\sum_{\substack{\underline{s}^n,\underline{u}^n\\\underline{v}^n,\underline{x}^n}}\sum_{\substack{\hat{\underline{m}}_1,\hat{\underline{b}}\\\hat{u}_1^n,\hat{u}_2^n}}\sum_{\substack{a_1,a_2\\b_1,b_2}}\mathrm{p}^n_{\underline{S}}(\underline{s}^n)\mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{\hat{u}_1^n}\pi^{\sigma,\eta_1}_{\hat{u}_1^n\hat{u}_2^n}\pi^{1,\eta_2}_{\hat{u}_1^n}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u_2^n}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u_2^n}\right)\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{u}^n,\underline{b}\\\underline{v}^n,\hat{\underline{m}}_1,\hat{\underline{b}}\end{smallmatrix}\right)P\left(\bigcap_{i=1}^4\mathcal{G}_i^7\,\middle|\,\mathcal{G}_{\underline{s}^n}\right),\text{ where} \quad (157)$$

$$\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{u}^n,\underline{b}\\\underline{v}^n,\hat{\underline{m}}_1,\hat{\underline{b}}\end{smallmatrix}\right)P\left(\bigcap_{i=1}^4\mathcal{G}_i^7\,\middle|\,\mathcal{G}_{\underline{s}^n}\right)=\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{u}^n,\underline{b}\\\underline{v}^n,\hat{\underline{m}}_1,\hat{\underline{b}}\end{smallmatrix}\right)P\left(\mathcal{G}_1^7\cap\mathcal{G}_2^7|\mathcal{G}_{\underline{s}^n}\right)P\left(\mathcal{G}_3^6\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^7\cap\mathcal{G}_2^7\right)P\left(\mathcal{G}_4^7\,|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}_1^7\cap\mathcal{G}_2^7\cap\mathcal{G}_3^7\right)$$

$$\leq P\left(\mathcal{G}_1^7|\mathcal{G}_{\underline{s}^n}\right)\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{u}^n,\underline{b}\\\underline{v}^n,\hat{\underline{m}}_1,\hat{\underline{b}}\end{smallmatrix}\right)\prod_{j=1}^2\frac{p^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n)}{|\mathcal{L}_{j1}(m_{j1},s_j^n)||\mathcal{L}_{j2}(\underline{m}_j,s_j^n)|} \quad (158)$$

$$=\frac{p^n_{U_1}(u_1^n)p^n_{U_2}(u_2^n)p^n_{U_1}(\hat{u}_1^n)p^n_{U_2}(\hat{u}_2^n)}{q^{2n}}\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{u}^n,\underline{b}\\\underline{v}^n,\hat{\underline{m}}_1,\hat{\underline{b}}\end{smallmatrix}\right)\prod_{j=1}^2\frac{p^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n)}{L_{j1}L_{j2}} \quad (159)$$

$$\leq\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{u}^n,\underline{b}\\\underline{v}^n,\hat{\underline{m}}_1,\hat{\underline{b}}\end{smallmatrix}\right)p^n_{U_1}(\hat{u}_1^n)p^n_{U_2}(\hat{u}_2^n)\prod_{j=1}^2\frac{4q^np^n_{X_j|U_j,V_j,S_j}(x_j^n|u_j^n,v_j^n,s_j^n)p^n_{U_j}(u_j^n)\exp\left\{\frac{9n\eta_5}{2}\right\}}{q^{n+k_j}\exp\{n\left(\frac{\log|\mathcal{B}_j|}{n}-I(U_j;S_j)_{\Upsilon}+H(V_j|S_j,U_j)_{\Upsilon}\right)\}} \quad (160)$$

$$\leq \mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{u}^n,\underline{b}\\ \underline{v}^n,\underline{\hat{m}}_1,\underline{\hat{b}}\end{smallmatrix}\right) p^n_{U_1}(\hat{u}^n_1)p^n_{U_2}(\hat{u}^n_2)\prod_{j=1}^{2}\frac{1}{q^{k_j}|\mathcal{B}_j|}4p^n_{X_j,U_j,V_j|S_j}(x^n_j,u^n_j,v^n_j|s^n_j)\exp\left\{\frac{29n\eta_5}{4}\right\}, \tag{161}$$

where (i) (158) follows from the fact that (a) conditioned on the entire codebooks, $B_j(m_{j1},s^n_j)=B^*_j$ and $A_j(\mathfrak{m}_j,s^n_j)$ is uniformly distributed in $\mathcal{L}_{j1}(m_{j1},s_{j^n})$ and $\mathcal{L}_{j2}(\mathfrak{m}_j,s^n_j)$, and in particular conditionally independent of the realization of the codebooks (See Remark 4), which implies that $P\left(\mathcal{G}^7_3\left|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}^7_1\cap\mathcal{G}^7_2\cap\mathcal{G}^7_3\right.\right)=\frac{1}{|\mathcal{L}_{j1}(m_{j1},s^n_j)||\mathcal{L}_{j2}(\mathfrak{m}_j,s^n_j)|}$ and (b) $P\left(\mathcal{G}^7_4\left|\mathcal{G}_{\underline{s}^n}\cap\mathcal{G}^7_1\cap\mathcal{G}^7_2\cap\mathcal{G}^7_3\right.\right)=\prod_{j=1}^{2}p^n_{X_j|U_jV_jS_j}(x^n_j|u^n_j,v^n_j,s^n_j)$, (ii) (159) follows from the facts that (a) the $\mathcal{U}_1-,\mathcal{U}_2-$codebooks are mutually independent and the codewords in the $\mathcal{U}_j-$codebook are mutually independent with distribution $p^n_{U_j}$, (b) two codewords in the two $\mathcal{V}_1-,\mathcal{V}_2-$UCC codes are pairwise independent (Lemma 5) and uniformly distributed in the $\mathcal{F}^n_q$ ambient space and (c) $|\mathcal{L}_{j1}(m_{j1},s^n_j)|\geq L_{j1}=\frac{1}{2}\exp\left\{n\left(\frac{\log|\mathcal{B}_j|}{n}-I(U_j;S_j)_\Upsilon-\frac{3\eta_5}{2}\right)\right\}$, $|\mathcal{L}_{j2}(\mathfrak{m}_j,s^n_j)|\geq L_{j2}=\frac{1}{2}\exp\left\{n\left(\frac{\log|\mathcal{B}_j|}{n}-\log q+H(V_j|S_j,U_j)_\Upsilon-3\eta_5\right)\right\}$, (iii) (160) follows from above definitions of L_{j1},L_{j2}, and (iv) (161) follows from the bound $\frac{p^n_{U_jV_jS_j}(u^n_j,v^n_j,s^n_j)}{p^n_{U_j}(u^n_j)p^n_{S_j}(s^n_j)}\geq\exp\{-n\left(H(S_j,U_j,V_j)_\Upsilon+\eta_5-H(S_j)_\Upsilon-H(U_j)_\Upsilon+\frac{\eta_5}{2}+\frac{\eta_5}{4}\right)\}=\exp\{-n\left(H(V_j|U_j,S_j)_\Upsilon-I(U_j;S_j)_\Upsilon+\frac{7\eta_5}{4}\right)\}$ for $(\underline{s}^n,\underline{u}^n,\underline{v}^n)$ satisfying the conditions of $\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{u}^n,\underline{b}\\ \underline{v}^n,\underline{\hat{m}}_1,\underline{\hat{b}}\end{smallmatrix}\right)$, implying $1\leq\frac{p^n_{U_jV_jS_j}(u^n_j,v^n_j,s^n_j)}{p^n_{U_j}(u^n_j)p^n_{S_j}(s^n_j)}\exp\left\{n\left(H(S_j|U_j,V_j)_\Upsilon-I(U_j;S_j)_\Upsilon+\frac{7\eta_5}{4}\right)\right\}$. Substituting (161) into (157) and recognizing that $\mathcal{G}\left(\begin{smallmatrix}\underline{s}^n,\underline{u}^n,\underline{b}\\ \underline{v}^n,\underline{\hat{m}}_1,\underline{\hat{b}}\end{smallmatrix}\right)\leq 1$, we have

$$\overline{\xi}_7(\underline{m})\leq\sum_{\substack{\underline{s}^n,\underline{u}^n\\ \underline{v}^n,\underline{x}^n}}\sum_{\substack{\underline{\hat{m}}_1,\underline{\hat{b}}\\ \hat{u}^n_1,\hat{u}^n_2}}\sum_{\substack{a_1,a_2\\ b_1,b_2}}\frac{\mathrm{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\sigma,\eta_1}_{\hat{u}^n_1\hat{u}^n_2}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u^n_2}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u^n_2}\right)}{[p^n_{U_1}(\hat{u}^n_1)p^n_{U_2}(\hat{u}^n_2)]^{-1}\exp\{-n8\eta_5\}q^{k_1+k_2}|\mathcal{B}_1||\mathcal{B}_2|}\prod_{j=1}^{2}p^n_{X_jU_jV_j|S_j}(x^n_j,u^n_j,v^n_j|s^n_j)$$

$$\leq\sum_{\underline{s}^n,\underline{u}^n}\sum_{\underline{v}^n,\underline{x}^n}\sum_{\hat{u}^n_1,\hat{u}^n_2}\frac{\mathrm{p}^n_{\underline{S}}(\underline{s}^n)\,\mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\sigma,\eta_1}_{\hat{u}^n_1\hat{u}^n_2}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u^n_2}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u^n_2}\right)p^n_{U_1}(\hat{u}^n_1)}{\exp\{-n\left(8\eta_5+R_{11}+R_{21}+B_1+B_2\right)\}[p^n_{U_2}(\hat{u}^n_2)]^{-1}}\prod_{j=1}^{2}p^n_{X_jU_jV_j|S_j}(x^n_j,u^n_j,v^n_j|s^n_j) \tag{162}$$

$$\leq\sum_{\hat{u}^n_1,\hat{u}^n_2}\sum_{u^n_2}p^n_{U_1}(\hat{u}^n_1)p^n_{U_2}(\hat{u}^n_2)p_{U_2}(u^n_2)\sum_{\substack{\underline{s}^n,u^n_1\\ \underline{v}^n,\underline{x}^n}}p^n_{\underline{S}U_1\underline{V}\underline{X}|U_2}(\underline{s}^n,u^n_1,\underline{v}^n,\underline{x}^n|u^n_2)\frac{\mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\sigma,\eta_1}_{\hat{u}^n_1\hat{u}^n_2}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u^n_2}\rho_{\underline{x}^n,\underline{s}^n}\pi^{2,\eta_3}_{u^n_2}\right)}{\exp\{-n\left(8\eta_5+R_{11}+R_{21}+B_1+B_2\right)\}} \tag{163}$$

$$\leq\sum_{\hat{u}^n_1,\hat{u}^n_2}\sum_{u^n_2}p^n_{U_1}(\hat{u}^n_1)p^n_{U_2}(\hat{u}^n_2)p_{U_2}(u^n_2)\mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\sigma,\eta_1}_{\hat{u}^n_1\hat{u}^n_2}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\mu}_{\eta_4}\pi^{2,\eta_3}_{u^n_2}\sigma^2_{u^n_2}\pi^{2,\eta_3}_{u^n_2}\right)\exp\{n\left(8\eta_5+R_{11}+R_{21}+B_1+B_2\right)\} \tag{164}$$

$$\leq\sum_{\hat{u}^n_1,\hat{u}^n_2}\sum_{u^n_2}p^n_{U_1}(\hat{u}^n_1)p^n_{U_2}(\hat{u}^n_2)p_{U_2}(u^n_2)\mathrm{tr}\left(\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\sigma,\eta_1}_{\hat{u}^n_1\hat{u}^n_2}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\mu}_{\eta_4}\sigma^2_{u^n_2}\right)\exp\{n\left(8\eta_5+R_{11}+R_{21}+B_1+B_2\right)\} \tag{165}$$

$$=\sum_{\hat{u}^n_1,\hat{u}^n_2}\sum_{u^n_2}p^n_{U_1}(\hat{u}^n_1)p^n_{U_2}(\hat{u}^n_2)p_{U_2}(u^n_2)\mathrm{tr}\left(\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\sigma,\eta_1}_{\hat{u}^n_1\hat{u}^n_2}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\mu}_{\eta_4}\sigma^2_{u^n_2}\pi^{\mu}_{\eta_4}\right)\exp\{n\left(8\eta_5+R_{11}+R_{21}+B_1+B_2\right)\} \tag{166}$$

$$=\sum_{\hat{u}^n_1,\hat{u}^n_2}p^n_{U_1}(\hat{u}^n_1)p^n_{U_2}(\hat{u}^n_2)\mathrm{tr}\left(\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\sigma,\eta_1}_{\hat{u}^n_1\hat{u}^n_2}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\mu}_{\eta_4}\mu^{\otimes n}\pi^{\mu}_{\eta_4}\right)\exp\{n\left(8\eta_5+R_{11}+R_{21}+B_1+B_2\right)\} \tag{167}$$

$$\leq\sum_{\hat{u}^n_1,\hat{u}^n_2}p^n_{U_1}(\hat{u}^n_1)p^n_{U_2}(\hat{u}^n_2)\mathrm{tr}\left(\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\sigma,\eta_1}_{\hat{u}^n_1\hat{u}^n_2}\pi^{1,\eta_2}_{\hat{u}^n_1}\pi^{\mu}_{\eta_4}\right)\exp\{n\left(8\eta_5+\eta_4+R_{11}+R_{21}+B_1+B_2-H(Y)_\Upsilon\right)\} \tag{168}$$

where (i) (162) follows from the fact the terms in the earlier summand are invariant with $a_1,a_2,b_1,b_2,\underline{\hat{m}}_1,\underline{\hat{b}}$, (ii) (163) follows from Markov chains $X_1V_1U_1-S_1-S_2U_2V_2X_2$ and $X_2V_2U_2-S_2-S_1U_1V_1X_1$, (iii) (164) follows from $\sum_{\underline{s}^n,u^n_1,\underline{v}^n,\underline{x}^n}p^n_{\underline{S}U_1\underline{V}\underline{X}|U_2}(\underline{s}^n,u^n_1,\underline{v}^n,\underline{x}^n|u^n_2)\rho_{\underline{x}^n\underline{s}^n}=\sigma^2_{u^n_2}$, (iv) (165) follows from the commutativity of $\pi^{2,\eta_3}_{u^n_2}$ and $\sigma^2_{u^n_2}$ implying $\pi^{2,\eta_3}_{u^n_2}\sigma^2_{u^n_2}\pi^{2,\eta_3}_{u^n_2}=\sqrt{\sigma^2_{u^n_2}}\pi^{2,\eta_3}_{u^n_2}\pi^{2,\eta_3}_{u^n_2}\sqrt{\sigma^2_{u^n_2}}=\sqrt{\sigma^2_{u^n_2}}\pi^{2,\eta_3}_{u^n_2}\sqrt{\sigma^2_{u^n_2}}\leq\sqrt{\sigma^2_{u^n_2}}\boldsymbol{I}\sqrt{\sigma^2_{u^n_2}}=\sigma^2_{u^n_2}$, (v) (166) follows from cyclicity of trace, (vi) (167) follows from $\sum_{u^n_2}p_{U_2}(u^n_2)\sigma^2_{u^n_2}=\mu^{\otimes n}$, (vii) (168) follows from quantum

conditional typicality Lemma 4 which states the operator inequality $\pi^{\mu}_{\eta_4}\mu^{\otimes n}\pi^{\mu}_{\eta_4} \leq \exp\{-n(H(Y)_\Upsilon - \eta_4)\}\pi^{\mu}_{\eta_4}$. Proceeding further, we have

$$\begin{aligned}
\overline{\xi}_7(\underline{m}) &\leq \sum_{\hat{u}_1^n,\hat{u}_2^n} p^n_{U_1}(\hat{u}_1^n)p^n_{U_2}(\hat{u}_2^n)\mathrm{tr}\Big(\pi^{\sigma,\eta_1}_{\hat{u}_1^n\hat{u}_2^n}\left[\pi^{1,\eta_2}_{\hat{u}_1^n}\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{\hat{u}_1^n}\right]\Big)\exp\{n\left(8\eta_5+\eta_4+R_{11}+R_{21}+B_1+B_2-H(Y)_\Upsilon\right)\} \\
&\leq \sum_{\hat{u}_1^n,\hat{u}_2^n} p^n_{U_1}(\hat{u}_1^n)p^n_{U_2}(\hat{u}_2^n)\mathrm{tr}\Big(\pi^{\sigma,\eta_1}_{\hat{u}_1^n\hat{u}_2^n}[\boldsymbol{I}]\Big)\exp\{n\left(8\eta_5+\eta_4+R_{11}+R_{21}+B_1+B_2-H(Y)_\Upsilon\right)\} \quad (169)\\
&\leq \sum_{\hat{u}_1^n,\hat{u}_2^n} p^n_{U_1}(\hat{u}_1^n)p^n_{U_2}(\hat{u}_2^n)\mathrm{tr}\Big(\pi^{\sigma,\eta_1}_{\hat{u}_1^n\hat{u}_2^n}\Big)\mathbb{1}_{\{(\hat{u}_1^n,\hat{u}_2^n)\in T^n_{\eta_1}(p_{U_1U_2})\}}\exp\{n\left(8\eta_5+\eta_4+R_{11}+R_{21}+B_1+B_2-H(Y)_\Upsilon\right)\} \quad (170)\\
&\leq \sum_{\hat{u}_1^n,\hat{u}_2^n} p^n_{U_1}(\hat{u}_1^n)p^n_{U_2}(\hat{u}_2^n)\mathbb{1}_{\{(\underline{\hat{u}}^n)\in T^n_{\eta_1}(p_{U_1U_2})\}}\exp\{n\left(8\eta_5+\eta_4+2\eta_1+R_{11}+R_{21}+B_1+B_2-H(Y)_\Upsilon+H(Y|\underline{U})\right)\} \quad (171)\\
&\leq \exp\{-n\left(I(Y;U_1U_2)_\Upsilon + I(U_1;U_2)_\Upsilon - 8\eta_5-\eta_4-5\eta_1-R_{11}-R_{21}-B_1-B_2\right)\} \quad (172)
\end{aligned}$$

where (i) (169) follows from cyclicity of trace, (ii) (169) follows from $\pi^{1,\eta_2}_{\hat{u}_1^n}\pi^{\mu}_{\eta_4}\pi^{1,\eta_2}_{\hat{u}_1^n} \leq \boldsymbol{I}$, (iii) the inclusion of the indicator $\mathbb{1}_{\{(\hat{u}_1^n,\hat{u}_2^n)\in T^n_{\eta_1}(p_{U_1U_2})\}}$ is justified by the fact that $\pi^{\sigma,\eta_1}_{\hat{u}_1^n\hat{u}_2^n} = 0$ if $(\hat{u}_1^n,\hat{u}_2^n) \notin T^n_{\eta_1}(p_{U_1U_2})$, thereby not altering the positive terms in the summand, (iv) (171) follows from the quantum conditional typicality Lemma 4 which states that $\mathrm{tr}\Big(\pi^{\sigma,\eta_1}_{\hat{u}_1^n\hat{u}_2^n}\Big) \leq \mathrm{tr}\Big(\pi^{\sigma,2\eta_1}_{\hat{u}_1^n\hat{u}_2^n}\Big) \leq \exp\{n(H(Y|U_1,U_2)_\Upsilon + 2\eta_1)\}$ whenever $(\hat{u}_1^n,\hat{u}_2^n) \in T^n_{\eta_1}(p_{U_1U_2})$ and $\pi^{\sigma,\eta_1}_{\hat{u}_1^n\hat{u}_2^n} = 0$ if $(\hat{u}_1^n,\hat{u}_2^n) \notin T^n_{\eta_1}(p_{U_1U_2})$.

APPENDIX O
QSTX : PROOF OF PROPOSITION 14

We begin by defining a common set of objects and relations that we shall leverage in the sequel. Suppose

$$\mathcal{G}_{s^n} \triangleq \{S^n = s^n\}, \mathcal{G}_1 \triangleq \{V^n(a,m) = v^n\}, \mathcal{G}_2 \triangleq \{|\mathcal{L}(m,s^n)| \geq L\}, \mathcal{G}_3 \triangleq \left\{A^*_{m,s^n} = a\right\}, \mathcal{G}_4 \triangleq \{X^n(m,s^n)=x^n\} \quad (173)$$

$\mathcal{G}(s^n,v^n) \triangleq \mathbb{1}_{\{s^n\in T_{\frac{\eta_3}{2}}(\mathtt{p}_S),(s^n,v^n)\in T_{\eta_3}(p_{SV})\}}$, then $p^n_{X|VS}(x^n|v^n,s^n)\mathtt{p}^n_S(s^n)\mathcal{G}(s^n,v^n) \leq p^n_{XVS}(x^n,v^n,s^n)\exp\Big\{n\binom{H(V|S)}{+2\eta_3}\Big\}$ (174)

We begin with $\overline{\zeta}_{41}(m) = \mathbb{E}\{\zeta_{41}(m)\}$ as defined through (64). Let $\mathcal{G}_i^4 \triangleq \mathcal{G}_i$ for $i \in [4]$. We have

$$\begin{aligned}
\overline{\zeta}_{41}(\underline{m}) &= 2\sum_{s^n,v^n,x^n}\sum_{a\in\mathcal{V}^k}\mathtt{p}^n_S(s^n)\,\mathrm{tr}\left\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{v^n}]\,\rho_{x^n,s^n}\right\}\mathcal{G}\left(s^n,v^n\right)P\left(\mathcal{G}_1^4\cap\mathcal{G}_2^4\cap\mathcal{G}_3^4\cap\mathcal{G}_4^4\big|\,\mathcal{G}_{s^n}\right), \text{ where} \quad (175)\\
\mathcal{G}\left(s^n,v^n\right)P\left(\mathcal{G}_1^4\cap\mathcal{G}_2^4\cap\mathcal{G}_3^4\cap\mathcal{G}_4^4|\mathcal{G}_{s^n}\right) &= \mathcal{G}\left(s^n,v^n\right)P\left(\mathcal{G}_1^4\cap\mathcal{G}_2^4|\mathcal{G}_{s^n}\right)P\left(\mathcal{G}_3^4\,\big|\mathcal{G}_{s^n}\cap\mathcal{G}_1^4\cap\mathcal{G}_2^4\right)P\left(\mathcal{G}_4^4\,\big|\mathcal{G}_{s^n}\cap\mathcal{G}_1^4\cap\mathcal{G}_2^4\cap\mathcal{G}_3^4\right) \quad (176)\\
&\leq \mathcal{G}\left(s^n,v^n\right)P\left(\mathcal{G}_1^4|\mathcal{G}_{s^n}\right)|\mathcal{L}(m,s^n)|^{-1}p^n_{X|V,S}(x^n|v^n,s^n) = \mathcal{G}\left(s^n,v^n\right)q^{-n}|\mathcal{L}(m,s^n)|^{-1}p^n_{X|V,S}(x^n|v^n,s^n) \quad (177)\\
&\leq \mathcal{G}\left(s^n,v^n\right)q^{-n}L^{-1}p^n_{X|V,S}(x^n|v^n,s^n) \leq \frac{4}{q^n}\mathcal{G}\left(s^n,v^n\right)\exp\left\{n\left(\log q - H(V|S)+3\eta_3-\tfrac{k}{n}\log q\right)\right\}p^n_{X|V,S}(x^n|v^n,s^n) \quad (178)\\
&\leq 4\mathcal{G}(s^n,v^n)\exp\{5n\eta_3 - k\log q\}p^n_{X,V|S}(x^n,v^n|s^n), \quad (179)
\end{aligned}$$

where (i) the inequality in (177) follows from the fact that $P\left(\mathcal{G}_3^4\,\big|\mathcal{G}_{s^n}\cap\mathcal{G}_1^4\cap\mathcal{G}_2^4\right) = \frac{1}{|\mathcal{L}(m,s^n)|}$ as argued in (61) and $P\left(\mathcal{G}_4^4\,\big|\mathcal{G}_{s^n}\cap\mathcal{G}_1^4\cap\mathcal{G}_2^4\cap\mathcal{G}_3^4\right) = p^n_{X|V,S}(x^n|v^n,s^n)$ from the distribution of the random code specified in (60), (ii) the equality in (177) follows from $P(\mathcal{G}_1^4|\mathcal{G}_{s^n}) = \frac{1}{q^n}$ that is proven in Lemma 5 in Appendix D, (iii) the inequalities in (178) follows from $|\mathcal{L}(m,s^n)| \geq L = \frac{1}{2}\exp\{k\log q - n\log q + H(V|S) - 3\eta_3\}$ defined prior to (56) and finally (iv) (179) follows from $\exp\left\{-n\left(H(V|S)+\frac{3\eta_3}{2}\right)\right\} \leq p^n_{V|S}(v^n|s^n)$ whenever $s^n \in T_{\frac{\eta_3}{2}}(\mathtt{p}_S)$, $(s^n,v^n)\in T_{\eta_3}(p_{SV})$, the latter conditions being ensured by the factor $\mathcal{G}(s^n,v^n)$. Substituting (179) into (175), we have

$$\begin{aligned}
\overline{\zeta}_{41}(\underline{m}) &\leq 8\sum_{s^n,v^n}\sum_{x^n}\mathrm{tr}\left\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{v^n}]\,\rho_{x^n,s^n}\right\}\mathcal{G}(s^n,v^n)\exp\{5n\eta_3\}\,p^n_{XVS}(x^n,v^n,s^n), \quad (180)\\
&\leq \sum_{v^n} 8p^n_V(v^n)\mathbb{1}_{\left\{v^n\in T_{\eta_3}(p_V)\right\}}\exp\{5n\eta_3\}\sum_{s^n,x^n}\mathrm{tr}\left\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{v^n}]\,\rho_{x^n,s^n}\right\}p^n_{XS|V}(x^n,s^n|v^n) \quad (181)\\
&\leq \sum_{v^n}\frac{8p^n_V(v^n)\mathbb{1}_{\{v^n\in T_{\eta_3}(p_V)\}}}{\exp\{-n5\eta_3\}}\,\mathrm{tr}\left\{[\boldsymbol{I}-\pi^{\sigma,\eta_2}_{v^n}]\,\sigma_{v^n}\right\} \leq \sum_{v^n}\frac{p^n_V(v^n)\log\left[\dim(\mathcal{H})\right]}{16^{-1}\exp\{-n5\eta_3\}}\exp\left\{-2n(\eta_2-\eta_3)^2\delta_{\mathrm{q}}(\sigma,p_V,\eta_1)\right\} \quad (182)\\
&\leq 16\log\left[\dim(\mathcal{H})\right]\exp\left\{-n\left[2(\eta_2-\eta_3)^2\delta_{\mathrm{q}}(\sigma,p_V,\eta_1)-5\eta_3\right]\right\}.
\end{aligned}$$

where (i) (180) follows from the fact that terms in the summand in (175), after substituting (179), are invariant to $a \in \mathcal{V}^k$, (ii) (181) follows from the indicator function in question being larger than or equal to the factor $\mathcal{G}\left(s^n,v^n\right)$, and lastly (iii) the second inequality in (182) is a result of the substantial overlap of the conditional typical projector $\pi^{\sigma,\eta_2}_{v^n}$ with σ_{v^n} whenever $v^n \in T_{2\eta_3}(p_V)$ and $\eta_2 > \eta_3$ as stated in Lemma 4.

The next term we analyze is $\overline{\zeta}_5(m) = \mathbb{E}\{\zeta_5(m)\}$ as defined in (59). We refer to (173) and let $\mathcal{G}_i^5 \triangleq \mathcal{G}_i$ for $i = 2,3,4$, $\mathcal{G}_1^5 \triangleq \mathcal{G}_1 \cap \{V^n(\hat{a},m) = \hat{v}^n\}$, $\mathcal{G}(s^n,v^n,a,\hat{a}) \triangleq \mathcal{G}(s^n,v^n)\mathbb{1}_{\{\hat{a}\neq a\}}$. With these definitions, it can be verified that

$$\overline{\zeta}_5(m) = 2\sum_{s^n,v^n,\hat{v}^n}\sum_{a\in\mathcal{V}^k}\sum_{\hat{a}\in\mathcal{V}^k} \mathtt{p}_S^n(s^n)\,\mathrm{tr}\left\{\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{\hat{v}^n}\pi^{\mu}_{\eta_1}\rho_{x^n,s^n}\right\}\mathcal{G}\left(s^n,v^n,a,\hat{a}\right)P(\mathcal{G}_1^5\cap\mathcal{G}_2^5\cap\mathcal{G}_3^5\cap\mathcal{G}_4^5\big|\,\mathcal{G}_{s^n}). \quad (183)$$

From Lemma 5, we have $P(\mathcal{G}_1^5|\mathcal{G}_{s^n}) = \frac{1}{q^{2n}}$. Substituting this, recognizing $\mathcal{G}_i^5 \triangleq \mathcal{G}_i$ for $i = 2,3,4$ and following a sequence of steps analogous to (176) - (179), we have

$$\mathcal{G}\left(s^n,v^n,a,\hat{a}\right)P\left(\mathcal{G}_1^5\cap\mathcal{G}_2^5\cap\mathcal{G}_3^5\cap\mathcal{G}_4^5\big|\,\mathcal{G}_{s^n}\right) \leq 4q^{-n}\exp\{5n\eta_3 - k\log q\}p^n_{X,V|S}(x^n,v^n|s^n). \quad (184)$$

Substituting (184) into (183) and recognizing that the terms in the summand do not depend on $a,\hat{a}\in\mathcal{V}^k$, we have

$$\overline{\zeta}_5(m) \leq \frac{8q^k}{q^n}\sum_{\hat{v}^n}\mathrm{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{v}^n}\pi^{\mu}_{\eta_1}\sum_{s^n,v^nx^n}p^n_{XVS}(x^n,v^n,s^n)\rho_{x^n,s^n}\pi^{\mu}_{\eta_1}\right\}\exp\{5n\eta_3\} \leq \frac{8q^k}{q^n}\sum_{\hat{v}^n}\mathrm{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{v}^n}\pi^{\mu}_{\eta_1}\mu^{\otimes n}\pi^{\mu}_{\eta_1}\right\}\exp\{5n\eta_3\} \quad (185)$$

$$\leq \frac{8q^k}{q^n}\sum_{\hat{v}^n}\mathrm{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{v}^n}\pi^{\mu}_{\eta_1}\right\}\exp\{-n(H(Y)_\Upsilon - 5\eta_3-\eta_1)\} \leq \frac{8q^k}{q^n}\sum_{\hat{v}^n}\mathrm{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{v}^n}\boldsymbol{I}\right\}\exp\left\{n(H(Y)_\Upsilon - 5\eta_3-\eta_1)\right\} \quad (186)$$

$$\leq \frac{8q^k}{q^n_{\hat{v}^n\in T_{\eta_2}(p_V)}}\sum\exp\{-n(H(Y)_\Upsilon - H(Y|V)_\Upsilon - 5\eta_3-\eta_1-2\eta_2)\} + \frac{8q^k}{q^n_{\hat{v}^n\in\mathcal{W}^n\setminus T_{\eta_2}(p_V)}}\sum 0\cdot\exp\{n(H(Y)_\Upsilon - 5\eta_3-\eta_1-2\eta_2)\} \quad (187)$$

$$\leq 8\exp\left\{-n(H(Y)_\Upsilon - H(Y|V)_\Upsilon + \log q - H(V)_\Upsilon - 5\eta_3-\eta_1-2\eta_2) + k\log q\right\} \quad (188)$$

where (i) (185) follows from $\mu = \sum_{x,s}p_{XS}(x,s)\rho_{x,s}$ as defined prior to (32), (ii) (186) follows from Lemma 3(ii), since $\pi^{\mu}_{\eta_1}$ is the typical projector of $\mu^{\otimes n}$ (iii) (187) follows from Lemma 4 (i) and the upper bound in Lemma 4, and finally (v) (188) follows from the bound on the size of the typical set $|T_{\eta_2}(p_V)|$ as stated in Lemma 1.

We now derive an upper bound on our last term $\overline{\zeta}_6(m) = \mathbb{E}\{\zeta_6(m)\}$ as defined in (59). Our analysis will be very similar to the one presented above for $\overline{\zeta}_5(m)$. We refer to (173) and let $\mathcal{G}_i^6 \triangleq \mathcal{G}_i$ for $i = 2,3,4$, $\mathcal{G}_1^6 \triangleq \mathcal{G}_1 \cap \{V^n(\hat{a},\hat{m}) = \hat{v}^n\}$, $\mathcal{G}(s^n,v^n,m,\hat{m}) \triangleq \mathcal{G}(s^n,v^n)\mathbb{1}_{\{\hat{m}\neq m\}}$. With these definitions, it can be verified that

$$\overline{\zeta}_6(m) = 2\sum_{s^n,v^n,\hat{v}^n}\sum_{a\in\mathcal{V}^k}\sum_{\hat{a}\in\mathcal{V}^k}\sum_{\hat{m}\in\mathcal{V}^l} \mathtt{p}_S^n(s^n)\,\mathrm{tr}\left\{\pi^{\mu}_{\eta_1}\pi^{\sigma,\eta_2}_{\hat{v}^n}\pi^{\mu}_{\eta_1}\rho_{x^n,s^n}\right\}\mathcal{G}\left(s^n,v^n,m,\hat{m}\right)P(\mathcal{G}_1^6\cap\mathcal{G}_2^6\cap\mathcal{G}_3^6\cap\mathcal{G}_4^6\big|\,\mathcal{G}_{s^n}). \quad (189)$$

From Lemma 5, we have $P(\mathcal{G}_1^6|\mathcal{G}_{s^n}) = \frac{1}{q^{2n}}$. Substituting this, recognizing $\mathcal{G}_i^6 \triangleq \mathcal{G}_i$ for $i = 2,3,4$ and following a sequence of steps analogous to (176) - (179), we have

$$\mathcal{G}\left(s^n,v^n,m,\hat{m}\right)P\left(\mathcal{G}_1^6\cap\mathcal{G}_2^6\cap\mathcal{G}_3^6\cap\mathcal{G}_4^6\big|\,\mathcal{G}_{s^n}\right) \leq 4q^{-n}\exp\{5n\eta_3 - k\log q\}p^n_{X,V|S}(x^n,v^n|s^n). \quad (190)$$

Substituting (190) into (189) and recognizing that the terms are invariant to $a,\hat{a}\in\mathcal{V}^k$ and $\hat{m}\in\mathcal{M}$, we have

$$\overline{\zeta}_5(m) \leq \frac{8q^{k+l}}{q^n}\sum_{\hat{v}^n}\mathrm{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{v}^n}\pi^{\mu}_{\eta_1}\sum_{s^n,v^nx^n}p^n_{XVS}(x^n,v^n,s^n)\rho_{x^n,s^n}\pi^{\mu}_{\eta_1}\right\}\exp\left\{5n\eta_3\right\} \leq \frac{8q^{k+l}}{q^n}\sum_{\hat{v}^n}\mathrm{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{v}^n}\pi^{\mu}_{\eta_1}\mu^{\otimes n}\pi^{\mu}_{\eta_1}\right\}\exp\{5n\eta_3\} \quad (191)$$

$$\leq \frac{8q^{k+l}}{q^n}\sum_{\hat{v}^n}\mathrm{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{v}^n}\pi^{\mu}_{\eta_1}\right\}\exp\left\{-n(H(Y)_\Upsilon - 5\eta_3-\eta_1)\right\} \leq \frac{8q^{k+l}}{q^n}\sum_{\hat{v}^n}\mathrm{tr}\left\{\pi^{\sigma,\eta_2}_{\hat{v}^n}\boldsymbol{I}\right\}\exp\{-n(H(Y)_\Upsilon - 5\eta_3-\eta_1)\} \quad (192)$$

$$\leq \frac{8q^{k+l}}{q^n_{\hat{v}^n\in T_{\eta_2}(p_V)}}\sum\exp\{-n\left(I(U;Y)_\Upsilon - 5\eta_3-\eta_1-2\eta_2\right)\} + \frac{8q^{k+l}}{q^n_{\hat{v}^n\in\mathcal{W}^n\setminus T_{\eta_2}(p_V)}}\sum 0\cdot\exp\{-n(H(Y)_\Upsilon - 5\eta_3-\eta_1-2\eta_2)\} \quad (193)$$

$$\leq 8\exp\left\{-n(H(Y)_\Upsilon - H(Y|V)_\Upsilon + \log q - H(V)_\Upsilon - 5\eta_3-\eta_1-2\eta_2) + k\log q + l\log q\right\} \quad (194)$$

where (i) (191) follows from $\mu = \sum_{x,s}p_{XS}(x,s)\rho_{x,s}$ as defined prior to (32), (ii) (192) follows from Lemma 3(ii), since $\pi^{\mu}_{\eta_1}$ is the typical projector of $\mu^{\otimes n}$ (iii) (193) follows from Lemma 4 (i) and the upper bound in Lemma 4, and finally (v) (194) follows from the bound on the size of the typical set $|T_{\eta_2}(p_V)|$ as stated in Lemma 1.

REFERENCES

[1] C. E. Shannon, "Channels with side information at the transmitter," IBM Journal of Research and Development, vol. 2, no. 4, pp. 289–293, 1958.

[2] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," Probs. of Ctrl. and Info. Th., vol. 9, no. 1, pp. 19–31, 1980.

[3] M. Costa, "Writing on dirty paper (corresp.)," IEEE Transactions on Information Theory, vol. 29, no. 3, pp. 439–441, 1983.

[4] T. Philosof and R. Zamir, "On the loss of single-letter characterization: The dirty multiple access channel," IEEE Trans. on Info. Th., vol. 55, pp. 2442–2454, June 2009.

[5] H. Boche, N. Cai, and J. Nötzel, "The classical-quantum channel with random state parameters known to the sender," Journal of Physics A: Mathematical and Theoretical, vol. 49, no. 19, p. 195302, apr 2016. [Online]. Available: https://doi.org/10.1088/1751-8113/49/19/195302

[6] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," IEEE Transactions on Information Theory, vol. 25, no. 3, pp. 306–311, May 1979.

[7] I. Savov and M. M. Wilde, "Classical codes for quantum broadcast channels," IEEE Transactions on Information Theory, vol. 61, no. 12, pp. 7017–7028, 2015.

[8] J. Yard, P. Hayden, and I. Devetak, "Quantum broadcast channels," IEEE Transactions on Information Theory, vol. 57, no. 10, pp. 7147–7162, 2011.

[9] Y. Kochman and R. Zamir, "Joint wyner-ziv/dirty-paper coding by modulo-lattice modulation," Information Theory, IEEE Transactions on, vol. 55, no. 11, pp. 4878–4889, Nov 2009.

[10] L. Song, J. Chen, and C. Tian, "Broadcasting correlated vector Gaussians," IEEE Transactions on Information Theory, vol. 61, no. 5, pp. 2465–2477, May 2015.

[11] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," IEEE Jrnl on Selected Areas in Commn., vol. 21, no. 5, pp. 684–702, June 2003.

[12] A. V. Kuznetsov and B. S. Tsybakov, "Coding in a memory with defective cells," Problemy Peredachi Informatsii, vol. 10, no. 2, pp. 52–60, OPT 1974, translation available in the journal Problems Information Transmission, 1974 Volume 10, Issue 2, Pages 132-138.

[13] A. Padakandla and S. S. Pradhan, "An Achievable Rate Region Based on Coset Codes for Multiple Access Channel With States," IEEE Transactions on Information Theory, vol. 63, no. 10, pp. 6393–6415, Oct 2017.

[14] S. S. Pradhan, A. Padakandla, and F. Shirani, "An algebraic and probabilistic framework for network information theory," Foundations and Trends® in Communications and Information Theory, vol. 18, no. 2, pp. 173–379, 2020. [Online]. Available: http://dx.doi.org/10.1561/0100000083

[15] T. A. Atif, A. Padakandla, and S. S. Pradhan, "Achievable rate-region for 3—User Classical-Quantum Interference Channel using Structured Codes," in 2021 IEEE International Symposium on Information Theory (ISIT), 2021, pp. 760–765.

[16] ——, "Computing Sum of Sources over a Classical-Quantum MAC," in 2021 IEEE International Symposium on Information Theory (ISIT), 2021, pp. 414–419.

[17] J. Korner and K. Marton, "How to encode the modulo-two sum of binary sources (corresp.)," IEEE Transactions on Information Theory, vol. 25, no. 2, pp. 219–221, 1979.

[18] B. Nazer and M. Gastpar, "Computation over multiple-access channels," IEEE Transactions on Information Theory, vol. 53, no. 10, pp. 3498–3516, 2007.

[19] A. Pastore, S. H. Lim, C. Feng, B. Nazer, and M. Gastpar, "A unified discretization approach to compute–forward: From discrete to continuous inputs," IEEE Transactions on Information Theory, vol. 69, no. 1, pp. 1–46, 2023.

[20] P. Sen, S. H. Lim, and Y.-H. Kim, "On the optimal achievable rates for linear computation with random homologous codes," IEEE Transactions on Information Theory, vol. 66, no. 10, pp. 6200–6221, 2020.

[21] T. Philosof, R. Zamir, U. Erez, and A. J. Khisti, "Lattice strategies for the dirty multiple access channel," IEEE Transactions on Information Theory, vol. 57, no. 8, pp. 5006–5035, 2011.

[22] A. Padakandla and S. S. Pradhan, "Achievable Rate Region for Three User Discrete Broadcast Channel Based on Coset Codes," IEEE Transactions on Information Theory, vol. 64, no. 4, pp. 2267–2297, April 2018.

[23] A. Padakandla, A. G. Sahebi, and S. S. Pradhan, "An Achievable Rate Region for the Three-User Interference Channel Based on Coset Codes," IEEE Transactions on Information Theory, vol. 62, no. 3, pp. 1250–1279, March 2016.

[24] A. Padakandla, "An achievable rate region for $3-$user classical-quantum broadcast channels," 2022. [Online]. Available: https://arxiv.org/abs/2203.00110

[25] M. Heidari, F. Shirani, and S. S. Pradhan, "A new achievable rate region for multiple-access channel with states," in 2017 IEEE International Symposium on Information Theory (ISIT), 2017, pp. 36–40.

[26] ——, "Quasi structured codes for multi-terminal communications," IEEE Transactions on Information Theory, vol. 65, no. 10, pp. 6263–6289, 2019.

[27] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," IEEE Transactions on Information Theory, vol. 45, no. 6, pp. 2007–2019, 1999.

[28] A. Somekh-Baruch, S. Shamai, and S. Verdu, "Cooperative multiple-access encoding with states available at one transmitter," IEEE Transactions on Information Theory, vol. 54, no. 10, pp. 4448–4469, 2008.

[29] G. Cervia, L. Luzzi, M. Le Treust, and M. R. Bloch, "Strong coordination of signals and actions over noisy channels with two-sided state information," IEEE Transactions on Information Theory, vol. 66, no. 8, pp. 4681–4708, 2020.

[30] A. Anshu, M. Hayashi, and N. A. Warsi, "Secure communication over fully quantum gel'fand-pinsker wiretap channel," IEEE Transactions on Information Theory, vol. 66, no. 9, pp. 5548–5566, 2020.

[31] Y. Chen and A. J. Han Vinck, "Wiretap channel with side information," IEEE Transactions on Information Theory, vol. 54, no. 1, pp. 395–402, 2008.

[32] T. S. Han and M. Sasaki, "Wiretap channels with causal state information: Strong secrecy," IEEE Transactions on Information Theory, vol. 65, no. 10, pp. 6750–6765, 2019.
[33] Y.-K. Chia and A. E. Gamal, "Wiretap channel with causal state information," IEEE Transactions on Information Theory, vol. 58, no. 5, pp. 2838–2849, 2012.
[34] J. Nötzel, "Hypothesis testing on invariant subspaces of the symmetric group: part i. quantum sanov's theorem and arbitrarily varying sources," Journal of Physics A: Mathematical and Theoretical, vol. 47, no. 23, p. 235303, may 2014. [Online]. Available: https://dx.doi.org/10.1088/1751-8113/47/23/235303
[35] M. M. Wilde, Quantum Information Theory, 2nd ed. Cambridge University Press, 2017.
[36] I. Herstein, Topics in Algebra. Xerox College Pub., 1975. [Online]. Available: https://books.google.fr/books?id=lgfvAAAAMAAJ
[37] O. Fawzi, P. Hayden, I. Savov, P. Sen, and M. M. Wilde, "Classical communication over a quantum interference channel," IEEE Transactions on Information Theory, vol. 58, no. 6, pp. 3670–3691, 2012.
[38] A. Wyner, "Recent results in the shannon theory," Information Theory, IEEE Transactions on, vol. 20, no. 1, pp. 2 – 10, Jan 1974.
[39] M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," IEEE Transactions on Information Theory, vol. 49, no. 7, pp. 1753–1768, 2003.
[40] A. Winter, "The capacity of the quantum multiple-access channel," IEEE Transactions on Information Theory, vol. 47, no. 7, pp. 3059–3065, 2001.
[41] A. Padakandla, "An Algebraic Framework for Multi-terminal Communication," Ph.D. dissertation, Univ. of Michigan, Ann Arbor, USA, May. 2014, available at http://deepblue.lib.umich.edu/handle/2027.42/107264.
[42] I. Csiszár and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems, 2nd ed. Cambridge University Press, 2011.