Faithful Simulation of Distributed Quantum Measurements with Applications in Distributed Rate-Distortion Theory

Mohsen Heidari, Touheed Anwar Atif, and S. Sandeep Pradhan
Department of Electrical Engineering and Computer Science,
University of Michigan, Ann Arbor, MI 48109, USA.
Email: mohsenhd@umich.edu, touheed@umich.edu, pradhanv@umich.edu

Abstract

We consider faithful simulation of distributed quantum measurements, and characterize a set of sufficient communication and common randomness rates. To achieve this, we introduce random binning and mutual packing lemma for distributed quantum measurements. These techniques can be viewed as the quantum counterpart of their classical analogues. Finally, using these results, we develop a distributed quantum-to-classical rate distortion theory and characterize the rate-distortion region in terms of single-letter quantum mutual information quantities.

I. Introduction

Measurements are the interface between the intricate quantum world and the perceivable macroscopic classical world. A measurement associates to a quantum state a classical attribute. However, quantum phenomena, such as superposition, entanglement and non-commutativity contribute to uncertainty in the measurement outcomes. A key concern, from an information-theoretic standpoint, is to quantify the amount of "relevant information" conveyed by a measurement about a quantum state.

Winter's measurement compression theorem (as elaborated in [1]) quantifies the "relevant information" as the amount of resources needed to simulate the output of a quantum measurement applied to a given state. Imagine that an agent (Alice) performs a measurement M on a quantum state ρ and sends a set of classical bits to a receiver (Bob). Bob intends to *faithfully* recover the outcomes of Alice's measurements

This work was supported by NSF grant CCF 1717299.

without having access to ρ. The measurement compression theorem states that at least quantum mutual information ($I(X;R)$) amount of classical information and conditional entropy ($S(X|R)$) amount of common shared randomness are needed to obtain a *faithful simulation*.

The measurement compression theorem finds its applications in several paradigms including local purity distillation [1] and private classical communication over quantum channels [2]. This theorem was later used by Datta, et al. [3] to develop a quantum-to-classical rate-distortion theory. The problem involved lossy compression of a quantum information source into classical bits, with the task of compression performed by applying a measurement on the source. In essence, the objective of the problem was to minimize the storage of the classical outputs resulting from the measurement while ensuring sufficient reliability so as to be able to recover the quantum state (from classical bits) within a fixed level of distortion from the original quantum source. To achieve this, the authors in [4] advocated the use of measurement compression protocol and subsequently characterized the so called rate-distortion function in terms of single-letter quantum mutual information quantities. The authors further established that by employing a naive approach of measuring individual output of the quantum source, and then applying Shannon's rate-distortion theory to compress the classical data obtained is insufficient to achieve optimal rates.

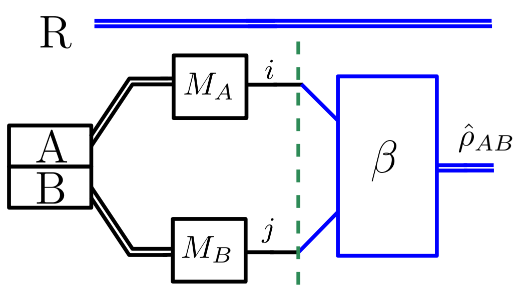

Fig. 1. The diagram of a distributed quantum measurement applied to a bipartite quantum system AB. A tensor product measurement $M_A \otimes M_B$ is performed on many copies of the observed quantum state. The outcomes of the measurements are given by two classical bits. The receiver functions as a classical-to-quantum channel β mapping the classical data to a quantum state.

In this work, we seek to quantify "relevant information" for quantum measurements performed in a distributed fashion. In this setting, as shown in Fig. 1, a composite bipartite quantum system AB is made available at two separate agents, named Alice and Bob. Alice and Bob have access only to sub-systems A and B, respectively. Two separate measurements, one for each sub-system, are performed in a distributed fashion with no communication taking place between Alice and Bob. Imagine that there is a third party (named Eve) who tries to simulate the action of the measurements without any access to the quantum systems. To achieve this objective, Alice and Bob send classical bits to Eve at rate r_1 and r_2, respectively. Eve on receiving these pairs of classical bits from Alice and Bob wishes to reconstruct the joint quantum state ρ_{AB} using a classical-to-quantum channel. The reconstruction has to satisfy a fidelity constraint characterized using a distortion observable or a trace norm.

One strategy is to apply Winter's measurement theorem [5] to compress each individual measurements M_A and M_B separately into $\tilde{M}_A$ and $\tilde{M}_B$. As a result, faithful simulation of M_A by $\tilde{M}_A$ is possible when at least $nI(X;R)$ classical bits of communication and $nS(X|R)$ bits of common randomness are available between Alice and Eve. Similarly, a faithful simulation of M_B by $\tilde{M}_B$ is possible with $nI(X;R)$ classical bits of communication and $nS(Y|R)$ bits of common randomness between Eve and Bob. The challenge here is that the direct use of single-POVM compression theorem for each individual POVMs, M_A and M_B, does not necessarily ensure a "distributed" faithful simulation for the overall measurement, $M_A \otimes M_B$.

One can further reduce the amount of classical communication by exploiting the statistical correlations between Alice's and Bob's measurement outcomes. The challenge here is that the classical outputs of the approximating POVMs (operating on n copies of the source) are not IID sequences — rather they are codewords generated from random coding. Therefore, standard classical source coding techniques are not applicable here. This issue also arises in classical distributed source coding problem which was addressed by Wyner-Ahlswede-Körner [6] by developing Markov Lemma and Mutual Packing Lemma.

Building upon these ideas, we develop a quantum-classical counterpart of these lemmas for the multi-user quantum measurement simulation problem. We characterize a set of sufficient communication and common randomness rates in terms of single-letter quantum information quantities (Theorem 2). To prove this theorem, we develop binning of quantum measurements. This technique can be viewed as the quantum counterpart of its classical analogues. The idea of binning in quantum setting has been used in [7] and [8] for quantum data compression involving side information. However, in this paper we introduce a novel binning technique for measurements which is different from these works. The binning in this work is used to construct measurements for Alice and Bob with fewer outcomes compared to the above individual measurements, i.e., $\tilde{M}_A$ and $\tilde{M}_B$.

Secondly, we use our results on the simulation of distributed measurements to develop a distributed quantum-to-classical rate distortion theory (Theorem 3). For the achievability part, we characterize an achievable rate region analogous to Berger-Tung's [6] in terms of single-letter quantum mutual information quantities. Further, we derive a single-letter outer-bound on the optimal rate-region (Theorem 4).

II. PRELIMINARIES

We here establish all our notations, briefly state few necessary definitions, and also provide Winter's theorem on measurement compression. Let $\mathcal{B}(\mathcal{H})$ denote the algebra of all bounded linear operators acting on a finite dimensional Hilbert space $\mathcal{H}$. Further, let $\mathcal{D}(\mathcal{H})$ denote the set of positive operators of unit trace acting on $\mathcal{H}$. Let I denote the identity operator. The trace distance between two operators A and B is defined as $\|A-B\|_1 \triangleq \operatorname{Tr}|A-B|$, where for any operator Λ we define $|\Lambda| \triangleq \sqrt{\Lambda^\dagger \Lambda}$. The von Neumann entropy of a density operator $\rho \in \mathcal{D}(\mathcal{H})$ is denoted by $S(\rho)$. The quantum mutual information and conditional entropy for a bipartite density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ are defined, respectively, as

$$I(A;B)_\rho \triangleq S(\rho_A) + S(\rho_B) - S(\rho_{AB}),$$

$$S(A|B)_\rho \triangleq S(\rho_{AB}) - S(\rho_B).$$

A positive-operator valued measure (POVM) acting on a Hilbert space $\mathcal{H}$ is a collection of $M \triangleq \{\Lambda_x\}$ of positive operators in $\mathcal{B}(\mathcal{H})$ that form a resolution of the identity:

$$\Lambda_x \geqslant 0, \forall x, \qquad \sum_x \Lambda_x = I.$$

If instead of the equality above, the inequality $\sum_x \Lambda_x \leqslant I$ holds, then the collection is said to be a sub-POVM. A sub-POVM M can be completed to form a POVM, denoted by $[M]$, by adding the operator $\Lambda_0 \triangleq (I - \sum_x \Lambda_x)$ to the collection. Let Ψ^ρ_{RA} denote a purification of a density operator $\rho \in D(\mathcal{H}_A)$. Given a POVM $M \triangleq \{\Lambda^A_x\}$ acting on ρ, the post-measurement state of the reference together with the classical outputs is represented by

$$(\mathrm{id} \otimes M)(\Psi^\rho_{RA}) \triangleq \sum_x |x\rangle\langle x| \otimes \operatorname{Tr}_A\{(I^R \otimes \Lambda^A_x)\Psi^\rho_{RA}\}. \tag{1}$$

Consider two POVMs $M_A = \{\Lambda^A_x\}$ and $M_B = \{\Lambda^B_x\}$ acting on $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Define $M_A \otimes M_B$ as a the collection of all operators of the form $\Lambda^A_x \otimes \Lambda^B_y$, for all x, y. With this definition, $M_A \otimes M_B$ is a POVM acting on $\mathcal{H}_A \otimes \mathcal{H}_B$. By $M^{\otimes n}$ denote the n-fold tensor product of the POVM M with itself.

A. Quantum Information Source

Consider a family of quantum states $\rho_i, i \in [1,m]$ acting on a Hilbert space $\mathcal{H}$. For each state assign a priori probability p_i. We denote such a setup by the ensemble $\{p_i, \rho_i, i \in [1:m]\}$. For such an ensemble, a quantum source is a sequence of states each equal to ρ_i with probability $p_i, i \in [1,m]$. Each realization of the source, after n generations of states, is represented by $\rho_{x^n} \triangleq \bigotimes_{j=1}^n \rho_{x_i}$, where x^n is a vector with elements in $[1,m]$. Let $\rho \triangleq \sum_i p_i \rho_i$, then the average density operator of the source after n generations is $\rho^{\bigotimes n}$.

B. Measurement Compression Theorem

Here, we provide a brief overview of the measurement compression theorem [5].

Definition 1 (Faithful simulation [1])**.** Given a POVM $M \triangleq \{\Lambda_x\}_{x\in\mathcal{X}}$ acting on a Hilbert space $\mathcal{H}_A$ and a density operator $\rho \in \mathcal{D}(\mathcal{H}_A)$, a sub-POVM $\tilde{M}$ acting on $\mathcal{H}_A^{\otimes n}$ is said to be ϵ-faithful to M, for $\epsilon > 0$, if the following holds:

$$\sum_{x^n\in\mathcal{X}^n} \|\sqrt{\rho^{\otimes n}}(\Lambda_{x^n} - \tilde{\Lambda}_{x^n})\sqrt{\rho^{\otimes n}}\|_1 + \mathrm{Tr}\{(I - \sum_{x^n} \tilde{\Lambda}_{x^n})\rho^{\otimes n}\} \leqslant \epsilon, \tag{2}$$

where $\Lambda_{x^n} = \Lambda_{x_1} \otimes \Lambda_{x_2} \otimes \cdots \otimes \Lambda_{x_n}$.

Lemma 1. [1] *For any state $\rho \in \mathcal{D}(\mathcal{H})$ with any purification Ψ^{ρ}_{RA}, and any pair of POVMs M and $\tilde{M}$ acting on $\mathcal{H}$, the following identity holds*

$$\|(id \otimes M)(\Psi^{\rho}_{RA}) - (id \otimes \tilde{M})(\Psi^{\rho}_{RA})\|_1 = \sum_x \|\sqrt{\rho}(\Lambda_x - \tilde{\Lambda}_x)\sqrt{\rho}\|_1, \tag{3}$$

where Λ_x and $\tilde{\Lambda}_x$ are the operators associated with M and $\tilde{M}$, respectively.

Theorem 1. [5] *For any $\epsilon > 0$, any density operator $\rho \in \mathcal{D}(\mathcal{H}_A)$ and any POVM M acting on the Hilbert space $\mathcal{H}_A$, there exist a collection of POVMs $\tilde{M}^{(\mu)}$ for $\mu \in [1, N]$, each acting on $\mathcal{H}_A^{\otimes n}$, and having at most 2^{nR} outcomes, where*

$$R \geqslant I(U;R)_\sigma + \delta(\epsilon), \qquad \frac{1}{n}\log_2 N + R \geqslant S(U)_\sigma + \delta(\epsilon)$$

such that $\tilde{M} \triangleq \frac{1}{N}\sum_\mu \tilde{M}^{(\mu)}$ is ϵ-faithful to M, where $\sigma_{UR} \triangleq (id \otimes M)(\Psi^{\rho}_{RA})$, and $\delta(\epsilon) \searrow 0$ as $\epsilon \searrow 0$.

III. Approximation of Distributed POVMs

We develop a measurement compression protocol in a distributed setting. Consider a bipartite composite quantum system (A, B) represented by Hilbert Space $\mathcal{H}_A \otimes \mathcal{H}_B$. Let ρ_{AB} be a quantum information source on $\mathcal{H}_A \otimes \mathcal{H}_B$. Consider two measurements M_A and M_B on sub-systems A and B, respectively. Imagine that three parties, named Alice, Bob and Eve, are trying to collectively simulate the two measurements, one applied to each sub-system. The three parties share some amount of classical common randomness. Alice and Bob perform a measurement $\tilde{M}_A$ and $\tilde{M}_B$ on n copies of sub-systems A and B, respectively. The measurements are performed in a distributed fashion with no communication taking place between Alice and Bob. Based on their respective measurements and the common randomness, Alice and Bob send some classical bits to Eve. Upon receiving these classical bits, Eve applies a processing operation on them and, then, wishes to produce an n-letter classical sequence. The objective is to construct n-letter measurements $\tilde{M}_A$ and $\tilde{M}_B$ that minimize the classical communication and common randomness bits while ensuring that the overall measurement induced by the action of the three parties is close to $M_A^{\otimes n} \otimes M_B^{\otimes}$. The problem is formally defined in the following.

Definition 2. For a given finite set $\mathcal{Z}$, and a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, a distributed protocol with parameters $(n, \Theta_1, \Theta_2, N)$ is characterized by

1) a collections of Alice's POVMs $\tilde{M}_A^{(\mu)}, \mu \in [1, N]$ each acting on $\mathcal{H}_A^{\otimes n}$ and with outcomes in a subset $\mathcal{L}_1$ satisfying $|\mathcal{L}_1| \leqslant \Theta_1$.

2) a collections of Bob's POVMs $\tilde{M}_B^{(\mu)}, \mu \in [1, N]$ each acting on $\mathcal{H}_B^{\otimes n}$ and with outcomes in a subset $\mathcal{L}_2$, satisfying $|\mathcal{L}_2| \leqslant \Theta_2$.

3) a collection of classical conditional probability distributions $\mathsf{P}^{(\mu)}(z^n|l_1, l_2)$ for all $l_1 \in \mathcal{L}_1, l_2 \in \mathcal{L}_2$, $z^n \in \mathcal{Z}^n$ and $\mu \in [1, N]$.

The overall POVM of this distributed protocol, given by $\tilde{M}_{AB}$, is characterized by the following operators:

$$\tilde{\Lambda}_{z^n} \triangleq \frac{1}{N} \sum_{\mu, l_1, l_2} \mathsf{P}^{(\mu)}(z^n|l_1, l_2)\ \Lambda_{l_1}^{A,(\mu)} \otimes \Lambda_{l_2}^{B,(\mu)}, \quad \forall z^n \in \mathcal{Z}^n, \tag{4}$$

where $\Lambda_{l_1}^{A,(\mu)}$ and $\Lambda_{l_2}^{B,(\mu)}$ are the operators corresponding to the POVMs $\tilde{M}_A^{(\mu)}$ and $\tilde{M}_B^{(\mu)}$, respectively.

In the above definition, (Θ_1, Θ_2) determines the amount of classical bits communicated from Alice and Bob to Eve. The amount of common randomness is determined by N, and μ can be viewed as the common randomness bits distributed among the parties. The classical stochastic mappings induced by $\mathsf{P}^{(\mu)}$ represent the action of Eve on the received classical bits.

Definition 3. Given a POVM M_{AB} acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, and a density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, a triplet (R_1, R_2, C) is said to be achievable, if for all $\epsilon > 0$ and for all sufficiently large n, there exists a distributed protocol with parameters $(n, \Theta_1, \Theta_2, N)$ such that its overall POVM $\tilde{M}_{AB}$ is ϵ-faithful to M_{AB} (see Definition 1), and

$$\begin{aligned} &\frac{1}{n} \log_2 \Theta_i \leqslant R_i + \epsilon, \quad i = 1, 2, \\ &\frac{1}{n} \log_2 N \leqslant C + \epsilon. \end{aligned}$$

Theorem 2. *Given a POVM $M_{AB} = \{\Lambda_z^{AB}\}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ and a density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, a triplet (R_1, R_2, C) is achievable if there exist POVMs $\bar{M}_A = \{\bar{\Lambda}_u^A\}$ and $\bar{M}_B = \{\bar{\Lambda}_v^B\}$ forming the decomposition*

$$\Lambda_z^{AB} = \sum_{u,v} P_{Z|U,V}(z|u, v) \bar{\Lambda}_u^A \otimes \bar{\Lambda}_v^B, \quad \forall z,$$

such that the following inequalities are satisfied:

$$R_1 \geqslant I(U;RB)_{\sigma_1} - I(U;V)_{\sigma_3}, \tag{5a}$$

$$R_2 \geqslant I(V;RA)_{\sigma_2} - I(U;V)_{\sigma_3}, \tag{5b}$$

$$R_1 + R_2 \geqslant I(U;RB)_{\sigma_1} + I(V;RA)_{\sigma_2} - I(U;V)_{\sigma_3}, \tag{5c}$$

$$C + R_1 + R_2 \geqslant S(U,V)_{\sigma_3}, \quad C + R_1 \geqslant S(U|V)_{\sigma_3}, \quad C + R_2 \geqslant S(V|U)_{\sigma_3}, \tag{5d}$$

where the information quantities are computed for the auxiliary states $\sigma_1^{RUB} \triangleq (id^R \otimes \bar{M}_A \otimes id^B)(\Psi^{\rho_{AB}}_{RAB})$, $\sigma_2^{RAV} \triangleq (id^R \otimes id^A \otimes \bar{M}_B)(\Psi^{\rho_{AB}}_{RAB})$, *and* $\sigma_3^{RUV} \triangleq (id^R \otimes \bar{M}_A \otimes \bar{M}_B)(\Psi^{\rho_{AB}}_{RAB})$.[1,2]

Proof. The proof is given in Appendix A. □

Let us consider the regime where the sum rate $(R_1 + R_2)$ is at its minimum achievable, i.e., equation (5c) is active. This requires the largest amount of common randomness, given by the constraint $C \geqslant S(U|RB)_{\sigma_1} + S(V|RA)_{\sigma_2}$. Fig. 2 demonstrates the region in Theorem 2 in terms of the quantum information quantities. It also shows the gains achieved by employing such an approach as opposed to independently compressing the two sources ρ_A and ρ_B.

Next, let us consider the regime where $C = 0$. This implies the following constraints on the rates:

$$R_1 \geqslant S(U|V)_{\sigma_3}, \quad R_2 \geqslant S(V|U)_{\sigma_3}, \quad R_1 + R_2 \geqslant S(U,V)_{\sigma_3}.$$

This regime corresponds to the quantum measurement $M_A \otimes M_B$ followed by classical Slepian-Wolf compression.

A. *Proof Techniques*

***Binning for POVMs**:* We introduce a quantum-counterpart of the classical binning technique used to prove Theorem 2. Here, we describe this technique.

Consider a POVM M with observables $\{\Lambda_{\alpha_1}, \Lambda_{\alpha_2}, .., \Lambda_{\alpha_N}\}$. Given K for which N is divisible, partition $[1, N]$ into K equal bins and for each $i \in [1, K]$, let $B(i)$ denote the i^{th} bin. The binned POVM $\tilde{M}$ is given by the collection of operators $\{\tilde{\Lambda}_{\beta_1}, \tilde{\Lambda}_{\beta_2}, .., \tilde{\Lambda}_{\beta_K}\}$ where $\tilde{\Lambda}_{\beta_i}$ is defined as

$$\tilde{\Lambda}_{\beta_i} = \sum_{j \in \mathcal{B}(i)} \Lambda_{\alpha_j}.$$

[1]Although, in the problem formulation the action Eve is to produce a classical reconstruction, the mutual information quantities are defined for quantum reconstructions. This is done to have a compact representation of the rate-region. An alternative equivalent representation of the rate-region can be obtained in terms of Holevo information.

[2]Note that $S(U)_{\sigma_1} = S(U)_{\sigma_3}$ and $S(V)_{\sigma_2} = S(V)_{\sigma_3}$.

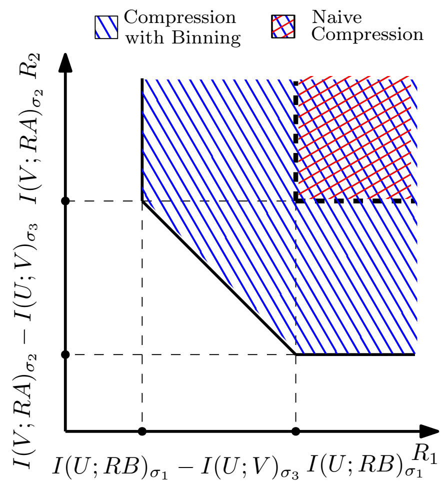

Fig. 2. The achievable rate region is depicted for two different schemes with at least $S(U|RB)_{\sigma_1} + S(V|RA)_{\sigma_2}$ amount of common randomness. The Naive compression scheme is where each quantum source is independently compressed, while the other scheme, in order to exploit the correlation among the measurement outcomes, bins the POVMs before applying the measurements. As a result, the rate achieved by the latter is lower than the naive compression which translates into a larger rate region.

Using the fact that Λ_{α_i} are self-adjoint and positive $\forall i \in [1, N]$ and $\sum_{i=1}^{N} \Lambda_{\alpha_i} = I$, (which is because M is a POVM); it follows that $\tilde{M}$ is a valid POVM.

Mutual Packing Lemma for POVMs: Another technique used to prove Theorem 2 is a quantum version of mutual packing lemma. In what follows, we describe the mutual packing lemma for quantum measurements. For a Hilbert Space $\mathcal{H}_{AB}$ consider a POVM of the form $M_A \otimes M_B$, where (M_A, M_B) are two POVMs each acting on one sub-system. The operators for M_A and M_B are denoted, respectively, by $\Lambda_u^A \in \mathcal{B}(\mathcal{H}_\mathcal{A}), u \in \mathcal{U}$ and $\Lambda_v^B \in \mathcal{B}(\mathcal{H}_\mathcal{B}), v \in \mathcal{V}$, where $\mathcal{U}$ and $\mathcal{V}$ are finite sets. Fix a joint-distribution P_{UV} on the set of all outcomes $\mathcal{U} \times \mathcal{V}$. Fix rates $r_1, r_2 > 0$. For each $l \in [1, 2^{nr_1}]$, let $U^n(l)$ be a random sequence generated according to $\prod_{i=1}^{n} P_U$. Similarly, let $V^n(k)$ be a random sequence distributed according to $\prod_{i=1}^{n} P_V$, where $k \in [1, 2^{nr_2}]$. Suppose $U^n(l)$'s and $V^n(k)$'s are independent. Define the following random operators:

$$A_{u^n} \triangleq |\{l : U^n(l) = u^n\}| \Lambda_{u^n}^A, \quad B_{v^n} \triangleq |\{k : V^n(k) = v^n\}| \Lambda_{v^n}^B$$

where $\Lambda_{u^n}^A = \bigotimes_i \Lambda_{u_i}^A$ and $\Lambda_{v^n}^B = \bigotimes_i \Lambda_{v_i}^B$.

Lemma 2. *For any $\epsilon > 0$ and sufficiently large n, with high probability*

$$\| \sum_{(u^n, v^n) \in \mathcal{T}_\delta^{(n)}(U,V)} A_{u^n} \otimes B_{v^n} \|_\infty \leqslant \epsilon \tag{6}$$

provided that $r_1 + r_2 < I(U;V) - \delta(\epsilon)$.

Proof. From the triangle-inequality and the definition of A_{u^n} and B_{v^n}, the norm in the lemma does not exceed the following

$$\sum_{l,k} \sum_{(u^n,v^n)\in\mathcal{T}_\delta^{(n)}(U,V)} \mathbb{1}\{U^n(l)=u^n, V^n(k)=v^n\}\|\Lambda^A_{u^n}\otimes\Lambda^B_{v^n}\|_\infty$$
$$\leqslant \sum_{l,k}\mathbb{P}\{(U^n(l),V^n(k))\in\mathcal{T}_\delta^{(n)}(U,V)\}$$

where the last inequality holds since $\Lambda^A_{u^n}\otimes\Lambda^B_{v^n}\leqslant I$. The proof completes from the classical mutual packing lemma. $\square$

IV. Quantum-to-Classical (q-c) Distributed Rate Distortion Theory

As an application of the above theorem on faithful simulation of distributed measurements (Theorem 2), we investigate the distributed extension of quantum-to-classical (q-c) rate distortion coding [3]. This problem is a quantum counterpart of the classical distributed source coding. In this setting, many copies of a bipartite quantum information source $\rho_{AB}\in\mathcal{D}(\mathcal{H}_A\otimes\mathcal{H}_B)$ are generated. Alice and Bob have access to the partial trace of the copies denoted by ρ_A and ρ_B, respectively; each performs a measurement on their copies and sends the classical outputs to Eve. The objective of Eve is to produce a reconstruction of the source ρ_{AB} within a targeted distortion threshold which is measured by a given distortion observable. To this end, upon receiving the classical bits sent by Alice and Bob, a reconstruction state is produced by Eve.

A. Problem Formulation

We first formulate this problem as follows. For any quantum information source $\rho_{AB}\in\mathcal{D}(\mathcal{H}_A\otimes\mathcal{H}_B)$, denote its purification by $\Psi^{\rho_{AB}}_{RAB}$.

Definition 4. A q-c source coding setup is characterized by a purified quantum information source $\Psi^{\rho_{AB}}_{RAB}\in\mathcal{D}(\mathcal{H}_R\otimes\mathcal{H}_A\otimes\mathcal{H}_B)$, a reconstruction Hilbert space $\mathcal{H}_{\hat{X}}$, and a distortion observable $\Delta\in\mathcal{B}(\mathcal{H}_R\otimes\mathcal{H}_{\hat{X}})$ which satisfies $\Delta\geqslant 0$.

Next, we formulate the action of Alice, Bob and Eve by the following definition.

Definition 5. An (n,Θ_1,Θ_2) q-c protocol for a given input and reconstruction Hilbert spaces ($\mathcal{H}_A\otimes\mathcal{H}_B$, $\mathcal{H}_{\hat{X}}$) is defined by POVMs $M_A^{(n)}$ and $M_B^{(n)}$ acting on $\mathcal{H}_A^{\otimes n}$ and $\mathcal{H}_B^{\otimes n}$ with Θ_1 and Θ_2 number of outcomes, respectively, and a set of reconstruction states $S_{i,j}\in\mathcal{D}(\mathcal{H}_{\hat{X}}^{\otimes n})$ for all $i\in[1:\Theta_1], j\in[1:\Theta_2]$.

The overall action of Alice, Bob and Eve, as a q-c protocol, on a quantum source ρ_{AB} is denoted by the following operation

$$\mathcal{N}_{A^nB^n\mapsto\hat{X}^n}:\rho_{AB}^{\otimes n}\mapsto\sum_{i,j}\mathrm{Tr}\{(\Lambda_i^A\otimes\Lambda_j^B)\rho_{AB}^{\otimes n}\}\ S_{i,j},\tag{7}$$

where $\{\Lambda_i^A\}$ and $\{\Lambda_j^B\}$ are, respectively, the operators of the POVMs $M_A^{(n)}$ and $M_B^{(n)}$. With this notation and given a q-c source coding setup as in Definition 4, the distortion of a $(n=1,\Theta_1,\Theta_2)$ q-c protocol is measured as

$$d(\rho_{AB},\mathcal{N}_{AB\mapsto\hat{X}})\triangleq\mathrm{Tr}\left\{\Delta\left((\mathrm{id}_R\otimes\mathcal{N}_{AB\mapsto\hat{X}})(\Psi_{RAB}^{\rho_{AB}})\right)\right\}.$$

For an n-letter protocol, we use symbol-wise average distortion observable defined as

$$\Delta^{(n)}=\frac{1}{n}\sum_{i=1}^{n}\Delta_{R_i\hat{X}_i}\otimes I_{R\hat{X}}^{\otimes[n]\setminus i},\tag{8}$$

where $\Delta_{R_i\hat{X}_i}$ is understood as the observable Δ acting on the ith instance space $\mathcal{H}_{R_i}\otimes\mathcal{H}_{\hat{X}_i}$ of the n-letter space $\mathcal{H}_R^{\otimes n}\otimes\mathcal{H}_{\hat{X}}^{\otimes n}$. With this notation, the distortion for an (n,Θ_1,Θ_2) q-c protocol is given by

$$\bar{d}(\rho_{AB}^{\otimes n},\mathcal{N}_{A^nB^n\mapsto\hat{X}^n})\triangleq\mathrm{Tr}\left\{\Delta^{(n)}(\mathrm{id}\otimes\mathcal{N}_{A^nB^n\mapsto\hat{X}^n})(\Psi_{R^nA^nB^n}^{\rho_{AB}})\right\},$$

where $\Psi_{R^nA^nB^n}^{\rho_{AB}}$ is the n-fold tensor product of $\Psi_{RAB}^{\rho_{AB}}$ which is the given purification of the source.

The authors in [3] studied the point-to-point version of the above formulation. They considered a special distortion observable of the form $\Delta=\sum_{\hat{x}\in\hat{\mathcal{X}}}\Delta_{\hat{x}}\otimes|\hat{x}\rangle\langle\hat{x}|$, where $\Delta_{\hat{x}}\geqslant 0$ acts on the reference Hilbert space and $\hat{\mathcal{X}}$ is the reconstruction alphabet. In this paper, we allow Δ to be any non-negative and bounded operator acting on the appropriate Hilbert spaces. Moreover, we allow for the use of any c-q reconstruction mapping as the action of Eve.

Definition 6. For a q-c source coding setup, a rate-distortion triplet (R_1,R_2,D) is said to be achievable, if for all $\epsilon>0$ and all sufficiently large n, there exists an (n,Θ_1,Θ_2) q-c protocol satisfying

$$\frac{1}{n}\log_2\Theta_i\leqslant R_i+\epsilon,\quad i=1,2,$$
$$\bar{d}(\rho_{AB}^{\otimes n},\mathcal{N}_{A^nB^n\mapsto\hat{X}^n})\leqslant D+\epsilon,$$

where $\mathcal{N}_{A^nB^n\mapsto\hat{X}^n}$ is defined as in (7). The set of all achievable rate-distortion triplets (R_1,R_2,D) is called the achievable rate-distortion region.

Our objective is to characterize the achievable rate-distortion region using single-letter information quantities.

B. Inner Bound

Theorem 3. *For a q-c source coding setup with a purified source $\Psi^{\rho_{AB}}_{RAB} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_A \otimes \mathcal{H}_B)$, and distortion observable Δ acting on $\mathcal{H}_R \otimes \mathcal{H}_{\hat{X}}$, any rate-distortion triplet (R_1, R_2, D) satisfying the following inequalities is achievable*

$$
\begin{aligned}
R_1 &\geqslant I(U;RB)_{\sigma_1} - I(U;V)_{\sigma_3},\\
R_2 &\geqslant I(V;RA)_{\sigma_2} - I(U;V)_{\sigma_3},\\
R_1 + R_2 &\geqslant I(U;RB)_{\sigma_1} + I(V;RA)_{\sigma_2} - I(U;V)_{\sigma_3},\\
D &\geqslant d(\rho_{AB}, \mathcal{N}_{AB\mapsto\hat{X}})
\end{aligned}
$$

for some POVMs $M_A \triangleq \{\Lambda^A_u\}_{u\in\mathcal{U}}$, $M_B \triangleq \{\Lambda^B_v\}_{v\in\mathcal{V}}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, and reconstruction states $\{S_{u,v}\}$ with each state in $\mathcal{D}(\mathcal{H}_{\hat{X}})$. The quantum mutual information quantities are computed according to the auxiliary states $\sigma_1^{RUB} \triangleq (id^R \otimes M_A \otimes id^B)(\Psi^{\rho_{AB}}_{RAB})$, $\sigma_2^{RAV} \triangleq (id^R \otimes id^A \otimes M_B)(\Psi^{\rho_{AB}}_{RAB})$, and $\sigma_3^{RUV} \triangleq (id^R \otimes M_A \otimes M_B)(\Psi^{\rho_{AB}}_{RAB})$, where (U,V) represents the output of $M_A \otimes M_B$, and $\mathcal{N}_{AB\mapsto\hat{X}} : \rho_{AB} \mapsto \sum_{u,v} \mathrm{Tr}\{(\Lambda^A_u \otimes \Lambda^B_v)\rho_{AB}\}\ S_{u,v}$.

Proof. The proof follows from Theorem 2. Fix POVMs (M_A, M_B) and reconstruction states $S_{u,v}$ as in the statement of the theorem. Let $\mathcal{N}_{AB\mapsto\hat{X}}$ be the mapping corresponding to these POVMs and the reconstruction states. Then, $d(\rho_{AB}, \mathcal{N}_{AB\mapsto\hat{X}}) \leqslant D$. According to Theorem 2, for any $\epsilon > 0$, there exists an $(n, 2^{nR_1}, 2^{nR_2}, N)$ distributed protocol for ϵ-faithful simulation of $M_A \otimes M_B$ on ρ_{AB} such that (R_1, R_2) satisfies the inequalities in (5) for $\bar{M}_A = M_A$ and $\bar{M}_B = M_B$. Let $\tilde{M}_A^{(\mu)}, \tilde{M}_B^{(\mu)}, \mu \in [1:N]$ and $\mathsf{P}^{(\mu)}$ be the POVM's and the conditional probability distributions of this protocol with $\mathcal{Z} = \mathcal{U} \times \mathcal{V}$. We use these POVM's and mappings to construct a q-c protocol for distributed source coding.

For each $\mu \in [1:N]$, consider the q-c protocol with parameters $\Theta_i = 2^{nR_i}, i = 1,2$, and POVMs $\tilde{M}_A^{(\mu)}, \tilde{M}_B^{(\mu)}$. Moreover, we use n-length reconstruction states $S_{i,j} \triangleq \sum_{u^n,v^n} \mathsf{P}^{(\mu)}(u^n, v^n|i,j) S_{u^n,v^n}$, where $S_{u^n,v^n} = \otimes_i S_{u_i,v_i}$. Further, let the corresponding mappings be denoted as $\tilde{\mathcal{N}}^{(\mu)}_{A^nB^n\mapsto\hat{X}^n}$. With this notation, for the average of these random protocols, the following bounds hold:

$$
\begin{aligned}
\frac{1}{N}\sum_{\mu} \bar{d}(\rho^{\otimes n}_{AB}, \tilde{\mathcal{N}}^{(\mu)}_{A^nB^n\mapsto\hat{X}^n}) &= \frac{1}{N}\sum_{\mu} \mathrm{Tr}\left\{\Delta^{(n)}(\mathrm{id}\otimes\tilde{\mathcal{N}}^{(\mu)}_{A^nB^n\mapsto\hat{X}^n})\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\}\\
&= \mathrm{Tr}\left\{\Delta^{(n)}(\mathrm{id}\otimes\mathcal{N}^{\otimes n}_{AB\mapsto\hat{X}})\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\} + \mathrm{Tr}\left\{\Delta^{(n)}(\mathrm{id}\otimes(\mathcal{N}^{\otimes n}_{AB\mapsto\hat{X}} - \tilde{\mathcal{N}}_{A^nB^n\mapsto\hat{X}^n}))\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\}\\
&\leqslant \mathrm{Tr}\left\{\Delta\left((\mathrm{id}_R\otimes\mathcal{N}_{AB\mapsto\hat{X}})(\Psi^{\rho_{AB}}_{R^nA^nB^n})\right)\right\} + \|\Delta^{(n)}(\mathrm{id}\otimes(\mathcal{N}^{\otimes n}_{AB\mapsto\hat{X}} - \tilde{\mathcal{N}}_{A^nB^n\mapsto\hat{X}^n}))\Psi^{\rho_{AB}}_{R^nA^nB^n}\|_1\\
&\leqslant D + \|\Delta^{(n)}\|_\infty \|(\mathrm{id}\otimes(\mathcal{N}^{\otimes n}_{AB\mapsto\hat{X}} - \tilde{\mathcal{N}}_{A^nB^n\mapsto\hat{X}^n}))\Psi^{\rho_{AB}}_{R^nA^nB^n}\|_1\\
&\leqslant D + \|\Delta^{(n)}\|_\infty \|(\mathrm{id}\otimes(M_A^{\otimes n}\otimes M_B^{\otimes n} - \tilde{M}_{AB}))\Psi^{\rho_{AB}}_{R^nA^nB^n}\|_1
\end{aligned}
$$

$$\leqslant D + \epsilon \|\Delta\|_\infty,$$

where $\tilde{\mathcal{N}}_{AB\mapsto\hat{X}}$ is the average of $\tilde{\mathcal{N}}^{(\mu)}_{AB\mapsto\hat{X}}$, and $\tilde{M}_{AB}$ is the overall POVM of the underlying distributed protocol as given in (4). The first inequaliy holds by the fact that $|\operatorname{Tr}\{A\}| \leqslant \|A\|_1$. The second inequality follows by Lemma 3 given in the sequel. The third inequality is due to the monotonicity of the trace-distance [9] with respect to the quantum channel given by $\mathrm{id} \otimes \mathcal{L}^{\otimes n}_{UV\mapsto\hat{X}}$, where

$$\mathcal{L}_{UV\mapsto\hat{X}}(\omega) \triangleq \sum_{u,v} \langle u, v|\, \omega\, |u, v\rangle\, S_{u,v}.$$

The last inequality follows by Theorem 2, and the fact that $\|\Delta^{(n)}\|_\infty \leqslant \|\Delta\|_\infty$. This completes the proof of the theorem, since Δ is a bounded operator. □

Lemma 3. *For any operator A and B acting on a Hilbert space $\mathcal{H}$ the following inequalities hold.*

$$\|BA\|_1 \leqslant \|B\|_\infty \|A\|_1, \qquad \textit{and} \qquad \|AB\|_1 \leqslant \|B\|_\infty \|A\|_1.$$

Proof. According to Theorem 1.3 in [10], A has a polar decomposition of the form $A = U|A|$, where U is a unitary operator and $|A| = \sqrt{A^\dagger A}$. As $|A|$ is a positive semi-definite operator, it has an eigenvalue decomposition of the form $|A| = \sum_{i=1}^d \lambda_i \,|\phi_i\rangle\!\langle\phi_i|$, where $\lambda_i \geqslant 0$. From triangle-inequality we have

$$\begin{aligned}
\|BA\|_1 = \|BU|A|\|_1 &\leqslant \sum_i \lambda_i \|BU\,|\phi_i\rangle\!\langle\phi_i|\,\|_1 = \sum_i \lambda_i \operatorname{Tr}\sqrt{|\phi_i\rangle\!\langle\phi_i|\, U^\dagger B^\dagger B U\, |\phi_i\rangle\!\langle\phi_i|} \\
&= \sum_i \lambda_i \sqrt{\langle\phi_i|\, U^\dagger B^\dagger B U\, |\phi_i\rangle} = \sum_i \lambda_i \|BU\,|\phi_i\rangle\,\| \\
&\leqslant \sum_i \|B\|_\infty \lambda_i = \|B\|_\infty \|A\|_1,
\end{aligned}$$

where the last inequality holds by the definition of $\|\cdot\|_\infty$ and the fact that U is unitary. For the second statement of the lemma we have

$$\begin{aligned}
\|AB\|_1 &\leqslant \sum_i \lambda_i \|U\,|\phi_i\rangle\!\langle\phi_i|\, B\|_1 = \sum_i \lambda_i \operatorname{Tr}\sqrt{B^\dagger\, |\phi_i\rangle\!\langle\phi_i|\, U^\dagger U\, |\phi_i\rangle\!\langle\phi_i|\, B} \\
&= \sum_i \lambda_i \operatorname{Tr}\sqrt{B^\dagger\, |\phi_i\rangle\!\langle\phi_i|\, B}.
\end{aligned}$$

Let $|\psi_i\rangle \triangleq \frac{B|\phi_i\rangle}{\|B|\phi_i\rangle\|}$. Then

$$\operatorname{Tr}\sqrt{B\,|\phi_i\rangle\!\langle\phi_i|\,B} = \|B\,|\phi_i\rangle\,\| \operatorname{Tr}\Big\{\sqrt{|\psi_i\rangle\!\langle\psi_i|}\Big\} = \|B\,|\phi_i\rangle\,\| \leqslant \|B\|_\infty.$$

Therefore, we obtain $\|AB\|_1 \leqslant \sum_i \|B\|_\infty \lambda_i = \|B\|_\infty \|A\|_1$. □

One can observe that the rate-region in Theorem 3 matches in form with the classical Berger-Tung region when ρ_{AB} is a mixed state of a collection of orthogonal pure states. Note that the rate-region is an inner bound for the set of all achievable rates. The single-letter characterization of the set of achievable rates is still an open problem even in the classical setting. Some progress has been made recently on this problem which provides an improvement over Berger-Tung rate region [11].

C. Outer Bound

In this section, we provide an outer bound for the achievable rate-distortion region.

Theorem 4. *Given a q-c source coding setup with a purified source $\Psi^{\rho_{AB}}_{RAB} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_A \otimes \mathcal{H}_B)$, and distortion observable Δ acting on $\mathcal{H}_R \otimes \mathcal{H}_{\hat{X}}$. If any triplet (R_1, R_2, D) is achievable, then the following inequalities must be satisfied*

$$R_1 \geqslant I(W_1; R|W_2, Q)_\sigma, \tag{9a}$$

$$R_2 \geqslant I(W_2; R|W_1, Q)_\sigma, \tag{9b}$$

$$R_1 + R_2 \geqslant I(W_1, W_2; R|Q)_\sigma, \tag{9c}$$

$$D \geqslant \mathrm{Tr}\{\Delta_{R\hat{X}} \sigma^{R\hat{X}}\}, \tag{9d}$$

where the state $\sigma^{W_1W_2RQ\hat{X}}$ can be written as

$$\sigma^{W_1W_2QR\hat{X}} = (id \otimes \mathcal{N}_{AB\mapsto W_1W_2Q\hat{X}})(\Psi^{\rho_{AB}}_{RAB}),$$

where Q represents an auxiliary quantum state, and $\mathcal{N}_{AB\mapsto W_1W_2Q\hat{X}}$ is a quantum test channel such that $I(R;Q)_\sigma = 0$.

Proof. Suppose the triplet (R_1, R_2, D) is achievable. Then, from Definition 6, for all $\epsilon > 0$, there exists an (n, Θ_1, Θ_2) q-c protocol satisfying the inequalities in the Definition. Let $M_A \triangleq \{\Lambda^A_{l_1}\}$, $M_B \triangleq \{\Lambda^B_{l_2}\}$ and $S_{l_1,l_2} \in \mathcal{D}(\mathcal{H}_{\hat{X}^{\otimes n}})$ be the corresponding POVMs and reconstruction states. Let L_1, L_2 denote the outcomes of the measurements. Then, for Alice's rate, we obtain

$$\begin{aligned} n(R_1 + \epsilon) &\geqslant H(L_1) \geqslant H(L_1|L_2) \\ &\geqslant I(L_1; R^n|L_2)_\tau \\ &= \sum_{j=1}^{n} I(L_1; R_j|L_2, R^{j-1})_\tau \end{aligned}$$

where the state τ is defined as

$$\tau^{L_1L_2R^n\hat{X}^n} \triangleq \sum_{l_1,l_2} |l_1, l_2\rangle\langle l_1, l_2| \otimes \mathrm{Tr}_{A^nB^n}\left\{(\mathrm{id} \otimes \Lambda^A_{l_1} \otimes \Lambda^B_{l_2})\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\} \otimes S_{l_1,l_2}.$$

Note that for each j the corresponding mutual information above is defined for a state in the Hilbert space $\mathcal{H}_{L_1} \otimes \mathcal{H}_{L_2} \otimes \mathcal{H}_R^{\otimes j}$. Next, we convert the above summation into a single-letter quantum mutual information term. For that we proceed with defining a new Hilbert space using direct-sum operation.

Let us recall the direct-sum of Hilbert spaces [12]. Consider a tuple of Hilbert spaces $\mathcal{H}_k$, $k = 1, 2, \ldots, n$ with inner products $\langle \cdot | \cdot \rangle_k$. Define $\bigoplus_{k=1}^n \mathcal{H}_k$ as the collection of tuples of vectors $(|x\rangle_1, |x\rangle_2, \ldots, |x\rangle_n)$. The inner product of two tuples $(|x\rangle_1, |x\rangle_2, \ldots, |x\rangle_n)$ and $(|y\rangle_1, |y\rangle_2, \ldots, |y\rangle_n)$ is given by the sum of inner products of the components, i.e., $\sum_{k=1}^n \langle x_k | y_k \rangle_k$. A linear operator in this space is a tuple of operators given by $(A_1, A_2, \ldots, A_n)$, where A_k operates on $\mathcal{H}_k$, and $\mathrm{Tr}(A) = \sum_{k=1}^n \mathrm{Tr}(A_i)$. A state in $\bigoplus_{k=1}^n \mathcal{H}_k$ is denoted conventionally as $\bigoplus_{k=1}^n |x\rangle_k$. Similarly, a linear operator in this space is written in the form $A = \bigoplus_{k=1}^n A_k$.

With this definition, consider the following single-letterization:

$$\sum_{j=1}^n I(L_1; R_j | L_2, R^{j-1})_\tau = nI(L_1; R | L_2, Q)_\sigma,$$

where the state σ is defined below

$$\sigma^{L_1 L_2 R Q \hat{X}} \triangleq$$

$$\sum_{l_1, l_2} \frac{1}{n} |l_1, l_2\rangle\langle l_1, l_2| \otimes \Big(\bigoplus_{j=1}^n \big(\mathrm{Tr}_{R_{j+1}^n A^n B^n} \Big\{ (\mathrm{id} \otimes \Lambda_{l_1}^A \otimes \Lambda_{l_2}^B) \Psi_{R^n A^n B^n}^{\rho_{AB}} \Big\} \otimes |j\rangle\langle j| \otimes \mathrm{Tr}_{\hat{X}_{\sim j}} \{ S_{l_1, l_2} \} \big) \Big). \quad (10)$$

where $\mathrm{Tr}_{\hat{X}_{\sim j}}$ denotes tracing over $(\hat{X}^{\otimes j-1} \otimes \hat{X}_{j+1}^{\otimes n})$, and $Q \triangleq (R^{J-1}, J)$, and J is an averaging random variable which is uniformly distributed over $[1:n]$. We elaborate on the Hilbert space associated with Q as follows.

Suppose $\{|\phi_i\rangle\}_{i \in \mathcal{I}}$ is an orthonormal basis for $\mathcal{H}_R$. Then, a basis for $\mathcal{H}_R^{\otimes k}$ is given by

$$|\phi_{\mathbf{i}^k}\rangle \triangleq |\phi_{i_1}\rangle \otimes |\phi_{i_2}\rangle \otimes \cdots \otimes |\phi_{i_k}\rangle,$$

for all $\mathbf{i}^k \in \mathcal{I}^k$. Consider the direct-sum of the Hilbert spaces $\bigoplus_{k=1}^n \mathcal{H}_R^{\otimes k}$. Consider the Hilbert space $\mathcal{H}_J \otimes (\bigoplus_{k=1}^n \mathcal{H}_R^{\otimes k})$. With this definition, define $\mathcal{H}_Q$, as the Hilbert space which is spanned by $|j\rangle \otimes |\phi_{\mathbf{i}^{(j-1)}}\rangle$, for all $j \in [1, n]$ and $\mathbf{i}^{(j-1)} \in \mathcal{I}^{(j-1)}$. Therefore, $\mathcal{H}_Q$ is *isometrically isomorphic* to the direct-sum $\bigoplus_k \mathcal{H}_R^{\otimes k}$. Note that $\mathcal{H}_Q$ can be viewed as a multi-particle Hilbert space.

Similarly, for Bob's rate we have

$$R_2 + \epsilon \geqslant I(L_2; R | L_1, Q)_\sigma.$$

For the sum-rate, the following inequalities hold

$$n(R_1 + R_2 + 2\epsilon) \geqslant H(L_1, L_2) \geqslant I(L_1, L_2; R^n)_\tau$$

$$= \sum_{j=1}^{n} I(L_1, L_2; R_j | R^{j-1})_\tau$$

$$= nI(L_1, L_2; R|Q)_\sigma$$

In addition, the distortion of this q-c protocol satisfies $\bar{d}(\rho_{AB}^{\otimes n}, \mathcal{N}_{A^nB^n\mapsto\hat{X}^n}) \leqslant D+\epsilon$, where $\mathcal{N}_{A^nB^n\mapsto\hat{X}^n}$ is the quantum channel associated with the protocol. Therefore, as the distortion observable is symbol-wise additive, we obtain

$$\begin{aligned} D+\epsilon &\geqslant \frac{1}{n}\sum_{j=1}^{n} \mathrm{Tr}\left\{ \left(\Delta_{R_j\hat{X}_j} \otimes I_{R\hat{X}}^{\otimes[n]\backslash j}\right)(\mathrm{id}\otimes\mathcal{N}_{A^nB^n\mapsto\hat{X}^n})(\Psi_{R^nA^nB^n}^{\rho_{AB}})\right\} \\ &= \frac{1}{n}\sum_{j=1}^{n} \mathrm{Tr}\left\{ \left(\Delta_{R_j\hat{X}_j} \otimes I_{R_1^{j-1}} \otimes I_{R_{j+1}^n\hat{X}_{\sim j}}\right)(\mathrm{id}\otimes\mathcal{N}_{A^nB^n\mapsto\hat{X}^n})(\Psi_{R^nA^nB^n}^{\rho_{AB}})\right\} \\ &= \frac{1}{n}\sum_{j=1}^{n} \mathrm{Tr}\left\{ \left(\Delta_{R_j\hat{X}_j} \otimes I_{R_1^{j-1}}\right)\left(\mathrm{Tr}_{R_{j+1}^n\hat{X}_{\sim j}}\{(\mathrm{id}\otimes\mathcal{N}_{A^nB^n\mapsto\hat{X}^n})(\Psi_{R^nA^nB^n}^{\rho_{AB}})\}\right)\right\} \\ &\overset{(a)}{=} \mathrm{Tr}\{(\Delta\otimes I_Q)\sigma^{RQ\hat{X}}\}, \end{aligned}$$

where the third equality holds, because of the following argument. From (10), one can show by partially tracing over (L_1, L_2), that

$$\sigma^{RQ\hat{X}} = \mathrm{Tr}_{L_1,L_2}\{\sigma^{L_1L_2RQ\hat{X}}\} = \bigoplus_j \frac{1}{n}|j\rangle\langle j| \otimes \mathrm{Tr}_{R_{j+1}^n\hat{X}_{\sim j}}\{(\mathrm{id}\otimes\mathcal{N}_{A^nB^n\mapsto\hat{X}^n})(\Psi_{R^nA^nB^n}^{\rho_{AB}})\}, \tag{11}$$

and $I_Q \triangleq \bigoplus_{j=1}^n \left(I_R^{\otimes(j-1)} \otimes |j\rangle\langle j|\right)$. Then, I_Q is the identity operator acting on $\mathcal{H}_Q$. Therefore, the right-hand side of the equality (a) above can be written as

$$\mathrm{Tr}\{(\Delta_{R\hat{X}}\otimes I_Q)\sigma^{RQ\hat{X}}\} = \mathrm{Tr}\left\{\Delta_{R\hat{X}}\sigma^{R\hat{X}}\right\}.$$

Let us identify the single-letter quantum test channel as given in the statement of the theorem. First, due to the distributive property of tensor product over direct sum operation, we can rewrite $\sigma^{L_1L_2RQ\hat{X}}$ as

$$\sigma^{L_1L_2RQ\hat{X}} \triangleq \left(\bigoplus_{j=1}^{n}\frac{1}{n}\sum_{l_1,l_2}|l_1,l_2\rangle\langle l_1,l_2| \otimes \left(\mathrm{Tr}_{R_{j+1}^nA^nB^n}\left\{(\mathrm{id}\otimes\Lambda_{l_1}^A\otimes\Lambda_{l_2}^B)\Psi_{R^nA^nB^n}^{\rho_{AB}}\right\}\otimes|j\rangle\langle j|\otimes\mathrm{Tr}_{\hat{X}_{\sim j}}\{S_{l_1,l_2}\}\right)\right).$$

Next, we identify a quantum channel $\mathcal{N}_{AB\mapsto L_1L_2Q\hat{X}} : \rho_{AB} \mapsto \sigma^{L_1L_2Q\hat{X}}$. For that and for any j define the following intermediate quantum channels:

$$\mathcal{N}_{AB\mapsto L_1L_2R^{(j-1)}\hat{X}}^{(j)}(\omega_{AB}) \triangleq \sum_{l_1,l_2}|l_1,l_2\rangle\langle l_1,l_2| \otimes \left(\mathrm{Tr}_{R_{j+1}^nA^nB^n}\left\{(\mathrm{id}_{R_{\sim j}}\otimes\Lambda_{l_1}^A\otimes\Lambda_{l_2}^B)(\omega_{AB}\otimes E_j)\right\}\otimes\mathrm{Tr}_{\hat{X}_{\sim j}}\{S_{l_1,l_2}\}\right),$$

where $E_j = \Psi^{\rho_{AB}}_{(RAB)_{\sim j}}$. One can verify that $\mathcal{N}^{(j)}_{AB\mapsto L_1L_2R^{(j-1)}\hat{X}}$ is indeed a quantum channel. With these definitions, let

$$\mathcal{N}_{AB\mapsto L_1L_2Q\hat{X}}(\omega_{AB}) \triangleq \bigoplus_j \frac{1}{n}\left(\mathcal{N}^{(j)}_{AB\mapsto L_1L_2R^{(j-1)}\hat{X}}(\omega_{AB}) \otimes |j\rangle\langle j|\right).$$

Using the property of direct-sum operation, one can verify that $\mathcal{N}_{AB\mapsto L_1L_2Q\hat{X}}$ is a valid quantum channel, moreover,

$$\sigma^{L_1L_2RQ\hat{X}} = (\mathrm{id}\otimes\mathcal{N}_{AB\mapsto L_1L_2Q\hat{X}})(\Psi^{\rho_{AB}}_{RAB}).$$

Lastly, we show that the condition $I(R;Q)_\sigma = 0$ is also satisfied. By taking the partial trace of σ over $(L_1, L_2, \hat{X})$ we obtain the following state

$$\begin{aligned}
\sigma^{RQ} = \mathrm{Tr}_{L_1L_2\hat{X}}(\sigma^{L_1L_2RQ\hat{X}}) &= \bigoplus_{j=1}^{n}\frac{1}{n}\sum_{l_1,l_2}\left(\mathrm{Tr}_{R^n_{j+1}A^nB^n}\left\{(\mathrm{id}\otimes\Lambda^A_{l_1}\otimes\Lambda^B_{l_2})\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\}\right)\otimes|j\rangle\langle j|\\
&= \bigoplus_{j=1}^{n}\frac{1}{n}\left(\mathrm{Tr}_{R^n_{j+1}A^nB^n}\left\{\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\}\right)\otimes|j\rangle\langle j|\\
&= \bigoplus_{j=1}^{n}\frac{1}{n}\left(\mathrm{Tr}_{AB}\{\Psi^{\rho_{AB}}_{RAB}\}\right)^{\otimes j}\otimes|j\rangle\langle j|\\
&= \mathrm{Tr}_{AB}\{\Psi^{\rho_{AB}}_{RAB}\}\otimes\left(\bigoplus_{j=1}^{n}\frac{1}{n}\left(\mathrm{Tr}_{AB}\{\Psi^{\rho_{AB}}_{RAB}\}\right)^{\otimes(j-1)}\otimes|j\rangle\langle j|\right),
\end{aligned}$$

where the last equality is due to the distributive property of tensor product over direct sum operation. Hence, σ^{RQ} is in a tensor product of the form $\sigma^R\otimes\sigma^Q$, and therefore, $I(R;Q)_\sigma = 0$. □

V. Conclusion

We established a distributed measurement compression theory. A set of communication rate-pairs and common randomness rate is characterized for faithful simulation of distributed measurements. We further investigated distributed quantum-to-classical rate-distortion theory and provide an inner-bound and an outer bound for that.

Appendix A
Proof of Theorem 2

Assume that the operators of the original POVM M_{AB} are decomposed as

$$\Lambda^{AB}_z = \sum_{u,v} P_{Z|U,V}(z|u,v)\bar{\Lambda}^A_u\otimes\bar{\Lambda}^B_v,\ \forall z, \tag{12}$$

for some POVMs $\bar{M}_A$ and $\bar{M}_B$ with operators denoted by $\{\bar{\Lambda}^A_u : u\in\mathcal{U}\}$ and $\{\bar{\Lambda}^B_v : v\in\mathcal{V}\}$, respectively, where $\mathcal{U},\mathcal{V}$ are two finite sets. The proof follows by constructing a protocol for faithful simulation of

$\bar{M}_A \otimes \bar{M}_B$. We start by generating the canonical ensembles corresponding to $\bar{M}_A$ and $\bar{M}_B$, given by $\{\lambda_u^A, \hat{\rho}_u^A\}_{u\in\mathcal{U}}$, $\{\lambda_v^B, \hat{\rho}_v^B\}_{v\in\mathcal{V}}$, and $\{\lambda_{uv}^{AB}, \hat{\rho}_{uv}^{AB}\}_{(u,v)\in\mathcal{U}\times\mathcal{V}}$, where

$$\lambda_u^A \triangleq \mathrm{Tr}\{\bar{\Lambda}_u^A \rho_A\}, \qquad \lambda_v^B \triangleq \mathrm{Tr}\{\bar{\Lambda}_v^B \rho_B\}, \qquad \lambda_{uv}^{AB} \triangleq \mathrm{Tr}\{(\bar{\Lambda}_u^A \otimes \bar{\Lambda}_v^B)\rho_{AB}\},$$

$$\hat{\rho}_u^A \triangleq \frac{1}{\lambda_u^A}\sqrt{\rho_A}\bar{\Lambda}_u^A\sqrt{\rho_A}, \quad \hat{\rho}_v^B \triangleq \frac{1}{\lambda_v^Y}\sqrt{\rho_B}\bar{\Lambda}_v^B\sqrt{\rho_B}, \quad \hat{\rho}_{uv}^{AB} \triangleq \frac{1}{\lambda_{uv}^{AB}}\sqrt{\rho_{AB}}(\bar{\Lambda}_u^A \otimes \bar{\Lambda}_v^B)\sqrt{\rho_{AB}}$$

where $\rho_A = \mathrm{Tr}_B\{\rho_{AB}\}$ and $\rho_B = \mathrm{Tr}_A\{\rho_{AB}\}$. Note that $\{\lambda_{uv}^{AB}\}$ is a joint probability distribution on $\mathcal{U}\times\mathcal{V}$ with $\{\lambda_u^A\}$ and $\{\lambda_v^B\}$ as the marginals. With this notation, corresponding to each of the probability distributions, we can associate a δ-typical set. Let us denote $\mathcal{T}_\delta^{(n)}(A)$, $\mathcal{T}_\delta^{(n)}(B)$ and $\mathcal{T}_\delta^{(n)}(AB)$ as the δ-typical sets defined for $\{\lambda_u^A\}$, $\{\lambda_v^B\}$ and $\{\lambda_{uv}^{AB}\}$, respectively.

Let Π_{ρ_A} and Π_{ρ_B} denote the δ-typical projectors (as in [10]) for marginal density operators ρ_A and ρ_B, respectively. Also, for any $u^n \in \mathcal{U}^n$ and $v^n \in \mathcal{V}^n$, let $\Pi_{u^n}^A$ and $\Pi_{v^n}^B$ denote the conditional typical projectors (as in [10]) for the canonical ensembles $\{\lambda_u^A, \hat{\rho}_u^A\}$ and $\{\lambda_v^B, \hat{\rho}_v^B\}$, respectively. For each $u^n \in \mathcal{U}^n$ and $v^n \in \mathcal{V}^n$ define

$$\Lambda_{u^n}^{A'} = \Pi_{\rho_A}\Pi_{u^n}^A\hat{\rho}_{u^n}^A\Pi_{u^n}^A\Pi_{\rho_A}, \quad \Lambda_{v^n}^{B'} = \Pi_{\rho_B}\Pi_{v^n}^B\hat{\rho}_{v^n}^B\Pi_{v^n}^B\Pi_{\rho_B}, \tag{13}$$

where $\hat{\rho}_{u^n}^A \triangleq \bigotimes_i \hat{\rho}_{u_i}^A$ and $\hat{\rho}_{v^n}^B \triangleq \bigotimes_i \hat{\rho}_{v_i}^B$.

Let U^n and V^n be random sequences generated independently and according to

$$\mathbb{P}(U^n = u^n) = \frac{\lambda_{u^n}^A}{1-\varepsilon}, \quad \forall u^n \in \mathcal{T}_\delta^{(n)}(A), \tag{14}$$

$$\mathbb{P}(V^n = v^n) = \frac{\lambda_{v^n}^B}{1-\varepsilon'}, \quad \forall v^n \in \mathcal{T}_\delta^{(n)}(B), \tag{15}$$

and $\mathbb{P}(U^n = u^n) = \mathbb{P}(V^n = v^n) = 0$ for any $u^n \notin \mathcal{T}_\delta^{(n)}(A)$ and $v^n \notin \mathcal{T}_\delta^{(n)}(B)$. Note that $|\mathcal{T}_\delta^{(n)}(A)| \leqslant 2^{n(S(\rho_A)+\delta_1)}$ and $|\mathcal{T}_\delta^{(n)}(B)| \leqslant 2^{n(S(\rho_B)+\delta_2)}$ for some $\delta_1, \delta_2 \in (0,1)$ and for all sufficiently large n. Here $\varepsilon > 0$ and $\varepsilon' > 0$ are chosen such that $\mathbb{P}(\mathcal{T}_\delta^{(n)}(A)) = \mathbb{P}(\mathcal{T}_\delta^{(n)}(B)) = 1$. From properties of typical sets, ε and ε' can be made arbitrary small for large enough n. With the notation above, define $\sigma^{A'} \triangleq \mathbb{E}[\Lambda_{U^n}^{A'}]$ and $\sigma^{B'} \triangleq \mathbb{E}[\Lambda_{V^n}^{B'}]$, where the expectation is taken with respect to U^n and V^n, respectively. Let $\hat{\Pi}^A$ and $\hat{\Pi}^B$ be the projectors onto the subspaces spanned by the eigen-states of $\sigma^{A'}$ and $\sigma^{B'}$ corresponding to eigenvalues that are larger than $\epsilon 2^{-n(S(\rho_A)+\delta_1)}$ and $\epsilon 2^{-n(S(\rho_B)+\delta_2)}$, respectively. Lastly, define

$$\Lambda_{u^n}^A \triangleq \hat{\Pi}^A\Lambda_{u^n}^{A'}\hat{\Pi}^A \text{ and } \sigma^A \triangleq \mathbb{E}[\Lambda_{U^n}^A], \tag{16}$$

$$\Lambda_{v^n}^B \triangleq \hat{\Pi}^B\Lambda_{v^n}^{B'}\hat{\Pi}^B \text{ and } \sigma^B \triangleq \mathbb{E}[\Lambda_{V^n}^B]. \tag{17}$$

A. Construction of Random POVMs

In what follows, we construct two random POVMs one for each encoder. Fix a positive integer N and positive real numbers $\tilde{R}_1$ and $\tilde{R}_2$ satisfying $\tilde{R}_1 < S(U)_{\sigma_3}$ and $\tilde{R}_2 < S(V)_{\sigma_3}$, where σ_3 is given in the statement of the theorem. For each $\mu \in [1, N]$, randomly and independently select $2^{n\tilde{R}_1}$ and $2^{n\tilde{R}_2}$ sequences according to P_{U^n} and P_{V^n} (as in (14) and (15)), respectively. Let $(U^{n,(\mu)}(l), V^{n,(\mu)}(k))$ represent the randomly selected sequences for each μ, where $l \in [1, 2^{n\tilde{R}_1}]$ and $k \in [1, 2^{n\tilde{R}_2}]$. Construct operators

$$A^{(\mu)}_{u^n} \triangleq \gamma^{(\mu)}_{u^n}\left(\sqrt{\rho_A}^{-1}\Lambda^A_{u^n}\sqrt{\rho_A}^{-1}\right) \quad \text{and} \quad B^{(\mu)}_{v^n} \triangleq \zeta^{(\mu)}_{v^n}\left(\sqrt{\rho_B}^{-1}\Lambda^B_{v^n}\sqrt{\rho_B}^{-1}\right) \tag{18}$$

where

$$\gamma^{(\mu)}_{u^n} \triangleq \frac{1-\varepsilon}{1+\eta}2^{-n\tilde{R}_1}|\{l : U^{n,(\mu)}(l) = u^n\}| \quad \text{and} \quad \zeta^{(\mu)}_{v^n} \triangleq \frac{1-\varepsilon'}{1+\eta}2^{-n\tilde{R}_2}|\{k : V^{n,(\mu)}(k) = v^n\}|, \tag{19}$$

where $\eta \in (0, 1)$ is a parameter to be determined. Then, for each $\mu \in [1, N]$ construct $M_1^{(n,\mu)}$ and $M_2^{(n,\mu)}$ as in the following

$$M_1^{(n,\mu)} \triangleq \{A^{(\mu)}_{u^n} : u^n \in \mathcal{T}_\delta^{(n)}(A)\}, \quad M_2^{(n,\mu)} \triangleq \{B^{(\mu)}_{v^n} : v^n \in \mathcal{T}_\delta^{(n)}(B)\}. \tag{20}$$

As a first step, one can show that with probability sufficiently close to one, $M_1^{(n,\mu)}$ and $M_2^{(n,\mu)}$ form sub-POVMs for all $\mu \in [1, N]$. More precisely the following Lemma holds.

Lemma 4. *For any positive integer N, and $\varepsilon, \varepsilon', \eta \in (0, 1)$, as in (19), and any $\zeta \in (0, 1)$, there exists $n(\varepsilon, \varepsilon', \eta, \zeta)$ such that for all $n \geqslant n(\varepsilon, \varepsilon', \eta, \zeta)$, the collection of operators $M_1^{(n,\mu)}$ and $M_2^{(n,\mu)}$ form sub-POVMs for all $\mu \in [1, N]$ with probability at least $(1 - \zeta)$, provided that*

$$\tilde{R}_1 > I(U; RB)_{\sigma_1}, \quad \textit{and} \quad \tilde{R}_2 > I(V; RA)_{\sigma_2},$$

where σ_1, σ_2 are defined as in the statement of the theorem. In addition, if

$$\frac{1}{n}\log_2 N + \tilde{R}_1 \geqslant S(U)_{\sigma_1} + \delta_1, \quad \frac{1}{n}\log_2 N + \tilde{R}_2 \geqslant S(V)_{\sigma_2} + \delta_2 \tag{21}$$

then with probability at least $(1 - \zeta)$ the collection of average operators $M_i^{(n)} \triangleq \frac{1}{N}\sum_\mu M_i^{(n,\mu)}, i = 1, 2$ are sub-POVMs and they are ϵ-faithful to M_A and M_B, respectively.

Proof. The proof uses a similar argument as in the proof of Theorem 2 in [5]. Hence it is omitted. □

B. Binning of POVMs

We introduce the quantum counterpart of the so-called *binning* technique which has been widely used in the context of classical distributed source coding. Fix binning rates (R_1, R_2). For each sequence $u^n \in \mathcal{U}^n$ assign an index from $[1, 2^{nR_1}]$ randomly and uniformly. The assignments for different sequences are done independently. A similar random and independent assignment is done for all $v^n \in \mathcal{V}^n$ with indices chosen from $[1, 2^{nR_2}]$. For each $i \in [1, 2^{nR_1}]$ and $j \in [1, 2^{nR_2}]$, let $\mathcal{B}_1(i)$ and $\mathcal{B}_2(j)$ denote the i^{th} and the j^{th} bins, respectively. Define the following operators:

$$\Gamma_i^{A,(\mu)} \triangleq \sum_{u^n \in \mathcal{B}_1(i)} A_{u^n}^{(\mu)}, \qquad \text{and} \qquad \Gamma_j^{B,(\mu)} \triangleq \sum_{v^n \in \mathcal{B}_2(j)} B_{v^n}^{(\mu)},$$

for all $i \in [1, 2^{nR_1}]$ and $j \in [1, 2^{nR_2}]$. The above operators generate the following POVMs

$$M_A^{(n,\mu)} \triangleq \{\Gamma_i^{A,(\mu)}\}_{i\in[1,2^{nR_1}]}, \quad M_B^{(n,\mu)} \triangleq \{\Gamma_j^{B,(\mu)}\}_{j\in[1,2^{nR_2}]} \tag{22}$$

Note that if $M_1^{(n,\mu)}$ and $M_2^{(n,\mu)}$ are sub-POVMs, then so are $M_A^{(n,\mu)}$ and $M_B^{(n,\mu)}$. This is due to the relations

$$\sum_i \Gamma_i^{A,(\mu)} = \sum_{u^n} A_{u^n}^{(\mu)}, \quad \text{and} \quad \sum_j \Gamma_j^{B,(\mu)} = \sum_{v^n} B_{v^n}^{(\mu)}.$$

We use the completion $[M_A^{(n,\mu)}]$ and $[M_B^{(n,\mu)}]$ as the POVMs for each encoder. Note that the index $i = 0$ and $j = 0$ are used, respectively, for $\Gamma_0^{A,(\mu)} = I - \sum_i \Gamma_i^{A,(\mu)}$ and $\Gamma_0^{B,(\mu)} = I - \sum_j \Gamma_j^{B,(\mu)}$. Note that the effect of the binning is in reducing the communication rates from $(\tilde{R}_1, \tilde{R}_2)$ to (R_1, R_2).

C. *Decoder mapping*

Note that the operators $A_{u^n}^{(\mu)} \otimes B_{v^n}^{(\mu)}$ are used to simulate $\bar{M}_A \otimes \bar{M}_B$. The binning can be viewed as partitioning of the set of classical outcomes into bins. Suppose an outcome (U^n, V^n) occurred after the measurement. Then, if the bins are small enough, one might be able to recover the outcomes by knowing the bin numbers. For that we will create a decoder that takes as an input a pair of bin numbers and produces a pair of sequences (U^n, V^n). More precisely, we define a mapping $F^{(\mu)}$ acting on the outputs of $[M_A^{(n,\mu)}]\otimes[M_B^{(n,\mu)}]$ as follows. Let $\mathcal{C}^{(\mu)}$ denote the codebook containing all pairs of codewords $(U^{n,(\mu)}(l), V^{n,(\mu)}(k))$. For each μ and bin numbers (i, j), let us define

$$D_{i,j}^{(\mu)} \triangleq \{(u^n, v^n) \in \mathcal{C}^{(\mu)} : (u^n, v^n) \in \mathcal{T}_\delta^{(n)}(AB) \text{ and } (u^n, v^n) \in \mathcal{B}_1(i) \times \mathcal{B}_2(j)\}. \tag{23}$$

For $i \in [1 : 2^{nR_1}]$ and $j \in [1, 2^{nR_2}]$ define the function $F^{(\mu)}(i, j) = (u^n, v^n)$ if (u^n, v^n) is the only element of $D_{i,j}^{(\mu)}$; otherwise $F^{(\mu)}(i, j) = (u_0^n, v_0^n)$, where (u_0^n, v_0^n) are arbitrary sequences. Further, $F^{(\mu)}(i, j) = (u_0^n, v_0^n)$ for $i = 0$ or $j = 0$. With this mapping, the resulted POVM is denoted by $\tilde{M}_{AB}^{(n)}$ with the following operators

$$\tilde{\Lambda}_{u^n,v^n} \triangleq \frac{1}{N} \sum_{\mu=1}^{N} \sum_{\substack{(i,j): \\ F^{(\mu)}(i,j)=(u^n,v^n)}} \Gamma_i^{A,(\mu)} \otimes \Gamma_j^{B,(\mu)}, \qquad \forall (u^n, v^n) \in \mathcal{U}^n \times \mathcal{V}^n.$$

We show that $\tilde{M}_{AB}^{(n)}$ is a POVM that is ϵ-faithful to the intermediate POVM $\bar{M}_A \otimes \bar{M}_B$. For faithful simulation of the original POVM M_{AB}, we apply the stochastic mapping $P_{Z|U,V}$ to the classical outputs of $\tilde{M}_{AB}^{(n)}$. More precisely, we construct the POVM $\hat{M}_{AB}^{(n)}$ with the following operators:

$$\hat{\Lambda}_{z^n}^{AB} = \sum_{u^n,v^n} P_{Z|U,V}^n(z^n|u^n,v^n)\tilde{\Lambda}_{u^n,v^n}, \ \forall z^n \in \mathcal{Z}^n.$$

D. Trace Distance

In what follows, we show that $\hat{M}_{AB}^{(n)}$ is ϵ-faithful (according to Definition 1) to M_{AB}, where $\epsilon > 0$ can be made arbitrarily small. More precisely, we show that, with probability sufficiently close to 1,

$$\sum_{z^n} \|\sqrt{\rho_{AB}^{\otimes n}}\left(\Lambda_{z^n} - \hat{\Lambda}_{z^n}\right)\sqrt{\rho_{AB}^{\otimes n}}\|_1 \leqslant \epsilon. \tag{24}$$

According to the decomposition of Λ_z, given in (12), the above inequality is equivalent to

$$\sum_{z^n}\left\|\sum_{u^n,v^n} P_{Z|U,V}^n(z^n|u^n,v^n)\left(\sqrt{\rho_{AB}^{\otimes n}}(\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \tilde{\Lambda}_{u^n,v^n})\sqrt{\rho_{AB}^{\otimes n}}\right)\right\|_1 \leqslant \epsilon.$$

From triangle inequality the left-hand side of the above inequality does not exceed the following

$$\begin{aligned}\sum_{z^n}\sum_{u^n,v^n} & P_{Z|U,V}^n(z^n|u^n,v^n)\left\|\sqrt{\rho_{AB}^{\otimes n}}(\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \tilde{\Lambda}_{u^n,v^n})\sqrt{\rho_{AB}^{\otimes n}}\right\|_1 \\ &= \sum_{u^n,v^n}\left\|\sqrt{\rho_{AB}^{\otimes n}}(\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \tilde{\Lambda}_{u^n,v^n})\sqrt{\rho_{AB}^{\otimes n}}\right\|_1\end{aligned}$$

Hence, it is sufficient to show that the above quantity is no greater than ϵ, with probability sufficiently close to 1. This is equivalent to showing that $\tilde{M}_{AB}^{(n)}$ is ϵ-faithful to $\bar{M}_A \otimes \bar{M}_B$. For that we prove the alternative formulation of ϵ-faithful simulation as given in Lemma 1, i.e.,

$$\|(\text{id} \otimes \bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n})(\Psi_{R^nA^nB^n}^{\rho}) - (\text{id} \otimes \tilde{M}_{AB}^{(n)})(\Psi_{R^nA^nB^n}^{\rho})\|_1 \leqslant \epsilon. \tag{25}$$

We characterize the conditions on (n, N, R_1, R_2) under which the inequality given in (25) holds. We start by applying a triangle inequality as in the following, where the binning effect is isolated.

$$\begin{aligned}&\|(\text{id} \otimes \bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n})(\Psi_{R^nA^nB^n}^{\rho}) - (\text{id} \otimes \tilde{M}_{AB}^{(n)})(\Psi_{R^nA^nB^n}^{\rho})\|_1 \\ &\qquad \leqslant \|(\text{id} \otimes \bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n})(\Psi_{R^nA^nB^n}^{\rho}) - \frac{1}{N}\sum_{\mu}(\text{id} \otimes [M_1^{(n,\mu)}] \otimes [M_2^{(n,\mu)}])(\Psi_{R^nA^nB^n}^{\rho})\|_1 \\ &\qquad + \|\frac{1}{N}\sum_{\mu}(\text{id} \otimes [M_1^{(n,\mu)}] \otimes [M_2^{(n,\mu)}])(\Psi_{R^nA^nB^n}^{\rho}) - (\text{id} \otimes \tilde{M}_{AB}^{(n)})(\Psi_{R^nA^nB^n}^{\rho})\|_1\end{aligned} \tag{26}$$

Next, we show that the first term in the right-hand side of (26) is sufficiently small. From triangle inequality and by adding and subtracting $\frac{1}{N}\sum_\mu(\text{id}\otimes[M_1^{(n,\mu)}]\otimes\bar{M}_B^{\otimes n})(\Psi^\rho_{R^nA^nB^n})$ we obtain:

$$\begin{aligned}
&\|(\text{id}\otimes\bar{M}_A^{\otimes n}\otimes\bar{M}_B^{\otimes n})(\Psi^\rho_{R^nA^nB^n})-\frac{1}{N}\sum_\mu(\text{id}\otimes[M_1^{(n,\mu)}]\otimes[M_2^{(n,\mu)}])(\Psi^\rho_{R^nA^nB^n})\|_1\\
&\quad\leqslant\|(\text{id}\otimes\bar{M}_A^{\otimes n}\otimes\bar{M}_B^{\otimes n})(\Psi^\rho_{R^nA^nB^n})-\frac{1}{N}\sum_\mu(\text{id}\otimes[M_1^{(n,\mu)}]\otimes\bar{M}_B^{\otimes n})(\Psi^\rho_{R^nA^nB^n})\|_1\\
&\quad+\|\frac{1}{N}\sum_\mu(\text{id}\otimes[M_1^{(n,\mu)}]\otimes\bar{M}_B^{\otimes n})(\Psi^\rho_{R^nA^nB^n})-\frac{1}{N}\sum_\mu(\text{id}\otimes[M_1^{(n,\mu)}]\otimes[M_2^{(n,\mu)}])(\Psi^\rho_{R^nA^nB^n})\|_1
\end{aligned}\tag{27}$$

We proceed with the following lemma.

Lemma 5. *Let ρ_X be the marginal of a given bipartite state ρ_{XY} corresponding to system X. Suppose, the collection of POVMs $M_X^{(\mu)}$ and $\hat{M}_X^{(\mu)},\mu\in[1:N]$ acting on the system X satisfies the following trace-distance inequality for some $\epsilon\in(0,1)$,*

$$\left\|\frac{1}{N}\sum_\mu\left[(id\otimes M_X^{(\mu)})(\Psi^{\rho_X}_{RX})-(id\otimes\hat{M}_X^{(\mu)})(\Psi^{\rho_X}_{RX})\right]\right\|_1\leqslant\epsilon,$$

where $\Psi^{\rho_X}_{RX}$ is a purification of ρ_X. Then, for any set of POVMs $M_Y^{(\mu)},\mu\in[1:N]$ acting on the system Y, the tensor-product POVMs $\hat{M}_X^{(\mu)}\otimes M_Y^{(\mu)}$ and $M_X^{(\mu)}\otimes M_Y^{(\mu)}$ acting on the system XY satisfy

$$\left\|\frac{1}{N}\sum_\mu\left[(id\otimes M_X^{(\mu)}\otimes M_Y^{(\mu)})(\Psi^{\rho_{XY}}_{R'XY})-(id\otimes\hat{M}_X^{(\mu)}\otimes M_Y^{(\mu)})(\Psi^{\rho_{XY}}_{R'XY})\right]\right\|_1\leqslant\epsilon,\tag{28}$$

for any purification $\Psi^{\rho_{XY}}_{R'XY}$ of ρ_{XY}.

Proof. The proof is given in Appendix B-A. □

Let $M_i^{(n)}\triangleq\frac{1}{N}\sum_\mu M_i^{(n,\mu)},i=1,2$. From Lemma 4, with probability at least $(1-\zeta)$, we have $M_1^{(n)}$ being ϵ-faithful to $\bar{M}_A$. This implies that

$$\left\|(\text{id}\otimes\bar{M}_A^{\otimes n})(\Psi^{\rho_A}_{\tilde{R}^nA^n})-(\text{id}\otimes[M_1^{(n)}])(\Psi^{\rho_A}_{\tilde{R}^nA^n})\right\|_1\leqslant\epsilon,$$

where $\Psi^{\rho_A}_{\tilde{R}^nA^n}$ is the n-fold tensor product of the state $\Psi^{\rho_A}_{\tilde{R}A}$, which is a purification for ρ_A. Also recall that $[M_1^{(n)}]$ is the completion of the sub-POVM $M_1^{(n)}$.

Next, we apply Lemma 5 with $\rho_{XY}=\rho_{AB}^{\otimes n},\rho_X=\rho_A^{\otimes n},M_X^{(\mu)}=\bar{M}_A^{\otimes n},\hat{M}_X^{(\mu)}=[M_1^{(n,\mu)}]$, and $M_Y^{(\mu)}=\bar{M}_B^{\otimes n},\mu\in[1:N]$. As a result, the following inequality holds, with probability at least $(1-\zeta)$:

$$\left\|\frac{1}{N}\sum_\mu\left[(\text{id}\otimes\bar{M}_A^{\otimes n}\otimes\bar{M}_B^{\otimes n})(\Psi^\rho_{R^nA^nB^n})-(\text{id}\otimes[M_1^{(n,\mu)}]\otimes\bar{M}_B^{\otimes n})(\Psi^\rho_{R^nA^nB^n})\right]\right\|_1\leqslant\epsilon.\tag{29}$$

The left-hand side of the above inequality is equals to the first term in the right-hand side of (27).

We use a similar approach for the second term in the right-hand side of (27). From Lemma 4, with probability at least $(1-\zeta)$, $M_2^{(n)}$ is ϵ-faithful to $\bar{M}_B$ and, hence,

$$\left\|(\text{id}\otimes \bar{M}_B^{\otimes n})(\Psi^{\rho_B}_{\hat{R}^nB^n}) - (\text{id}\otimes [M_2^{(n)}])(\Psi^{\rho_B}_{\hat{R}^nB^n})\right\|_1 \leqslant \epsilon,$$

where $\Psi^{\rho_B}_{\hat{R}^nB^n}$ is the n-fold tensor product of $\Psi^{\rho_B}_{\hat{R}B}$, which is a purification for ρ_B. From Lemma 5, with $\rho_X = \rho_B^{\otimes n}, M_X^{(\mu)} = \bar{M}_B^{\otimes n}, \hat{M}_X^{(\mu)} = [M_2^{(n,\mu)}]$, and $M_Y^{(\mu)} = [M_1^{(n,\mu)}]$, the following inequality holds, with probability at least $(1-\zeta)$:

$$\left\|\frac{1}{N}\sum_{\mu}\Big[(\text{id}\otimes [M_1^{(n,\mu)}]\otimes \bar{M}_B^{\otimes n})(\Psi^{\rho}_{R^nA^nB^n}) - (\text{id}\otimes [M_1^{(n,\mu)}]\otimes [M_2^{(n,\mu)}])(\Psi^{\rho}_{R^nA^nB^n})\Big]\right\|_1 \leqslant \epsilon. \quad (30)$$

The trace-norm in the left-hand side of the above inequality is equal to the second term in the right-hand side of (27). Thus, from (29) and (30), the right-hand side of (27) does not exceed 2ϵ, that is, with probability at least $(1-\zeta)$,

$$\left\|(\text{id}\otimes \bar{M}_A^{\otimes n}\otimes \bar{M}_B^{\otimes n})(\Psi^{\rho}_{R^nA^nB^n}) - \frac{1}{N}\sum_{\mu}(\text{id}\otimes [M_1^{(n,\mu)}]\otimes [M_2^{(n,\mu)}])(\Psi^{\rho}_{R^nA^nB^n})\right\|_1 \leqslant 2\epsilon. \quad (31)$$

As a result of the above inequality and the trace distance bound in (26), we obtain, with probability sufficiently close to 1, that

$$\begin{aligned}
&\left\|(\text{id}\otimes \bar{M}_A^{\otimes n}\otimes \bar{M}_B^{\otimes n})(\Psi^{\rho}_{R^nA^nB^n}) - (\text{id}\otimes \tilde{M}_{AB}^{(n)})(\Psi^{\rho}_{R^nA^nB^n})\right\|_1 \\
&\qquad \leqslant 2\epsilon + \left\|\frac{1}{N}\sum_{\mu}(\text{id}\otimes [M_1^{(n,\mu)}]\otimes [M_2^{(n,\mu)}])(\Psi^{\rho}_{R^nA^nB^n}) - (\text{id}\otimes \tilde{M}_{AB}^{(n)})(\Psi^{\rho}_{R^nA^nB^n})\right\|_1 \quad (32)
\end{aligned}$$

In what follows, we show that the second term above is also sufficiently small. For $(u^n,v^n)\in\mathcal{B}_1(i)\times\mathcal{B}_2(j)$ and $(u^n,v^n)\in\mathcal{C}^{(\mu)}$, define $e^{(\mu)}(u^n,v^n)\triangleq F^{(\mu)}(i,j)$. For any $(u^n,v^n)\notin\mathcal{C}^{(\mu)}$ define $e^{(\mu)}(u^n,v^n) = (u_0,v_0)$. Note that $e^{(\mu)}$ captures the overall effect of the binning followed by the decoding function $F^{(\mu)}$. For all $u^n\in\mathcal{U}^n$ and $v^n\in\mathcal{V}^n$, let $\Phi_{u^n,v^n} = |u^n,v^n\rangle\langle u^n,v^n|$. With this notation

$$\begin{aligned}
(\text{id}\otimes\tilde{M}_{AB}^{(n)})(\Psi^{\rho}_{R^nA^nB^n}) &= \frac{1}{N}\sum_{\mu}\sum_{i,j}\Phi_{F^{(\mu)}(i,j)}\otimes \text{Tr}_{AB}\{(\text{id}\otimes\Gamma_i^{A,(\mu)}\otimes\Gamma_j^{B,(\mu)})\Psi^{\rho}_{R^nA^nB^n}\} \\
&= \frac{1}{N}\sum_{\mu}\sum_{i,j\geqslant 1}\sum_{(u^n,v^n)\in\mathcal{B}_1(i)\times\mathcal{B}_2(j)}\Phi_{e^{(\mu)}(u^n,v^n)}\otimes\text{Tr}_{AB}\{(\text{id}\otimes A^{(\mu)}_{u^n}\otimes B^{(\mu)}_{v^n})\Psi^{\rho}_{R^nA^nB^n}\} \\
&+ \frac{1}{N}\sum_{\mu}\sum_{j\geqslant 1}\sum_{v^n\in\mathcal{B}_2(j)}\Phi_{(u_0^n,v_0^n)}\otimes\text{Tr}_{AB}\{(\text{id}\otimes(I-\sum_{u^n}A^{(\mu)}_{u^n})\otimes B^{(\mu)}_{v^n})\Psi^{\rho}_{R^nA^nB^n}\} \\
&+ \frac{1}{N}\sum_{\mu}\sum_{i\geqslant 1}\sum_{u^n\in\mathcal{B}_1(i)}\Phi_{(u_0^n,v_0^n)}\otimes\text{Tr}_{AB}\{(\text{id}\otimes A^{(\mu)}_{u^n}\otimes(I-\sum_{v^n}B^{(\mu)}_{v^n}))\Psi^{\rho}_{R^nA^nB^n}\} \\
&+ \frac{1}{N}\sum_{\mu}\Phi_{(u_0^n,v_0^n)}\otimes\text{Tr}_{AB}\{(\text{id}\otimes(I-\sum_{u^n}A^{(\mu)}_{u^n})\otimes(I-\sum_{v^n}B^{(\mu)}_{v^n}))\Psi^{\rho}_{R^nA^nB^n}\},
\end{aligned}$$

Note that

$$\frac{1}{N}\sum_{\mu}(\mathrm{id}\otimes[M_1^{(n,\mu)}]\otimes[M_2^{(n,\mu)}])(\Psi^{\rho}_{R^nA^nB^n})$$
$$= \frac{1}{N}\sum_{\mu}\sum_{u^n,v^n}\Phi_{(u^n,v^n)}\otimes \mathrm{Tr}_{AB}\{(\mathrm{id}\otimes A^{(\mu)}_{u^n}\otimes B^{(\mu)}_{v^n})\Psi^{\rho}_{R^nA^nB^n}\}$$
$$+ \frac{1}{N}\sum_{\mu}\sum_{v^n}\Phi_{(u_0^n,v_0^n)}\otimes \mathrm{Tr}_{AB}\{(\mathrm{id}\otimes (I-\sum_{u^n}A^{(\mu)}_{u^n})\otimes B^{(\mu)}_{v^n})\Psi^{\rho}_{R^nA^nB^n}\}$$
$$+ \frac{1}{N}\sum_{\mu}\sum_{u^n}\Phi_{(u_0^n,v_0^n)}\otimes \mathrm{Tr}_{AB}\{(\mathrm{id}\otimes A^{(\mu)}_{u^n}\otimes (I-\sum_{v^n}B^{(\mu)}_{v^n}))\Psi^{\rho}_{R^nA^nB^n}\}$$
$$+ \frac{1}{N}\sum_{\mu}\Phi_{(u_0^n,v_0^n)}\otimes \mathrm{Tr}_{AB}\{(\mathrm{id}\otimes (I-\sum_{u^n}A^{(\mu)}_{u^n})\otimes (I-\sum_{v^n}B^{(\mu)}_{v^n}))\Psi^{\rho}_{R^nA^nB^n}\},$$

The second term on the right-hand side of (32) is bounded from above using the triangle inequality as

$$\frac{1}{N}\sum_{\mu}\sum_{u^n,v^n}\left\|(\Phi_{(u^n,v^n)}-\Phi_{e^{(\mu)}(u^n,v^n)})\otimes \mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes A^{(\mu)}_{u^n}\otimes B^{(\mu)}_{v^n})\Psi^{\rho}_{R^nA^nB^n}\right\}\right\|_1$$
$$= \frac{1}{N}\sum_{\mu}\sum_{u^n,v^n}\|\Phi_{(u^n,v^n)}-\Phi_{e^{(\mu)}(u^n,v^n)}\|_1\times\left\|\mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes A^{(\mu)}_{u^n}\otimes B^{(\mu)}_{v^n})\Psi^{\rho}_{R^nA^nB^n}\right\}\right\|_1$$
$$= \frac{1}{N}\sum_{\mu}\sum_{u^n,v^n}\|\Phi_{(u^n,v^n)}-\Phi_{e^{(\mu)}(u^n,v^n)}\|_1\times \mathrm{Tr}\left\{(\mathrm{id}\otimes A^{(\mu)}_{u^n}\otimes B^{(\mu)}_{v^n})\Psi^{\rho}_{R^nA^nB^n}\right\}$$
$$= \frac{1}{N}\sum_{\mu}\sum_{u^n,v^n}\|\Phi_{u^n,v^n}-\Phi_{e^{(\mu)}(u^n,v^n)}\|_1\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n}\Omega_{u^n,v^n}$$

where $\Omega_{u^n,v^n}\triangleq \mathrm{Tr}\left\{\sqrt{\rho_A^{\otimes n}\otimes\rho_B^{\otimes n}}^{-1}(\Lambda^A_{u^n}\otimes\Lambda^B_{v^n})\sqrt{\rho_A^{\otimes n}\otimes\rho_B^{\otimes n}}^{-1}\rho^{\otimes n}\right\}$. As a result of the above argument, the right-hand side of (32) is bounded by

$$2\epsilon+\frac{1}{N}\sum_{u^n,v^n}\sum_{\mu}\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n}\Omega_{u^n,v^n}\|\Phi_{u^n,v^n}-\Phi_{e^{(\mu)}(u^n,v^n)}\|_1. \tag{33}$$

The above summation can be divided into two summations depending whether $(u^n,v^n)\in\mathcal{T}_\delta^{(n)}(AB)$. Noting that $\|\Phi_{u^n,v^n}-\Phi_{e^{(\mu)}(u^n,v^n)}\|_1\leqslant 2$, the above term does not exceed the following

$$2\epsilon+\frac{1}{N}\sum_{\substack{\mu\in[1,n]\\(u^n,v^n)\in\mathcal{T}_\delta^n(AB)}}\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n}\Omega_{u^n,v^n}\left\|\Phi_{u^n,v^n}-\Phi_{e^{(\mu)}(u^n,v^n)}\right\|_1+\frac{1}{N}\sum_{\substack{\mu\in[1,n]\\(u^n,v^n)\notin\mathcal{T}_\delta^n(AB)}}2\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n}\Omega_{u^n,v^n} \tag{34}$$

From (31) and Lemma 1, we have, with probability sufficiently close to 1, that

$$\sum_{(u^n,v^n)}\left\|\sqrt{\rho^{\otimes n}}\Big(\bar\Lambda^A_{u^n}\otimes\bar\Lambda^B_{v^n}-\big(\frac{1}{N}\sum_{\mu}A^{(\mu)}_{u^n}\otimes B^{(\mu)}_{v^n}\big)\Big)\sqrt{\rho^{\otimes n}}\right\|_1\leqslant 2\epsilon. \tag{35}$$

Then from triangle inequality, with probability sufficiently close to 1, for any subset $\mathcal{E}$, we have

$$\left\|\sum_{(u^n,v^n)\in\mathcal{E}}\sqrt{\rho^{\otimes n}}\Big(\bar\Lambda^A_{u^n}\otimes\bar\Lambda^B_{v^n}-\big(\frac{1}{N}\sum_{\mu}A^{(\mu)}_{u^n}\otimes B^{(\mu)}_{v^n}\big)\Big)\sqrt{\rho^{\otimes n}}\right\|_1\leqslant 2\epsilon, \tag{36}$$

which from the fact that $|\operatorname{Tr}\{A\}| \leqslant \|A\|_1$, implies that

$$\left|\operatorname{Tr}\left\{\sum_{(u^n,v^n)\in\mathcal{E}} \sqrt{\rho^{\otimes n}}\Big(\bar{\Lambda}^A_{u^n}\otimes\bar{\Lambda}^B_{v^n} - \big(\frac{1}{N}\sum_{\mu} A^{(\mu)}_{u^n}\otimes B^{(\mu)}_{v^n}\big)\Big)\sqrt{\rho^{\otimes n}}\right\}\right| \leqslant 2\epsilon. \tag{37}$$

As a result of this inequality and the fact that

$$\lambda^{AB}_{u^n,v^n} = \operatorname{Tr}\left\{\left(\bar{\Lambda}^A_{u^n}\otimes\bar{\Lambda}^B_{v^n}\right)\rho^{\otimes n}\right\}, \qquad \gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n}\Omega_{u^n,v^n} = \operatorname{Tr}\left\{(A^{(\mu)}_{u^n}\otimes B^{(\mu)}_{v^n})\rho^{\otimes n}\right\},$$

we have with probability sufficiently close to one, for any subset $\mathcal{E}$ of $\mathcal{U}^n\times\mathcal{V}^n$

$$\sum_{(u^n,v^n)\in\mathcal{E}} \lambda^{AB}_{u^n,v^n} - 2\epsilon \leqslant \frac{1}{N}\sum_{\mu}\sum_{(u^n,v^n)\in\mathcal{E}} \gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n}\Omega_{u^n,v^n} \leqslant \sum_{(u^n,v^n)\in\mathcal{E}} \lambda^{AB}_{u^n,v^n} + 2\epsilon \tag{38}$$

Therefore, the summation on non-typical sequences in (34) can be bounded, with probability sufficiently close to one, as

$$\frac{1}{N}\sum_{\substack{\mu\in[1,n]\\(u^n,v^n)\notin\mathcal{T}^n_\delta(AB)}} 2\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n}\Omega_{u^n,v^n} \leqslant 2\sum_{(u^n,v^n)\notin\mathcal{T}^n_\delta(AB)} \lambda^{AB}_{u^n,v^n} + 4\epsilon = 4\epsilon + 2\varepsilon'', \tag{39}$$

where $\varepsilon'' \triangleq 1-\mathbb{P}(\mathcal{T}^n_\delta(AB))$ which, from the properties of typical sets, can be made arbitrary small for large enough n. Next, we bound the summation on typical sequences in (34). From the definition of $\gamma^{(\mu)}_{u^n}$ and $\zeta^{(\mu)}_{v^n}$, such a summation equals to

$$\frac{(1-\varepsilon)(1-\varepsilon')}{(1+\eta)^2 N}2^{-n(\tilde{R}_1+\tilde{R}_2)}\sum_{(u^n,v^n)\in\mathcal{T}_\delta}\sum_{\mu}\sum_{l,k}\mathbb{1}\{U^{n,(\mu)}(l)=u^n, V^{n,(\mu)}(k)=v^n\}\Omega_{u^n,v^n}\|\Phi_{u^n,v^n}-\Phi_{e^{(\mu)}(u^n,v^n)}\|_1. \tag{40}$$

For any (u^n,v^n), the norm-1 above can be bounded from above by the following

$$2\times\mathbb{1}\Big\{\exists(\tilde{u}^n,\tilde{v}^n,i,j):(u^n,v^n)\in\mathcal{B}_1(i)\times\mathcal{B}_2(j),$$
$$(\tilde{u}^n,\tilde{v}^n)\in\mathcal{C}^{(\mu)}\bigcap\mathcal{T}^{(n)}_\delta(AB), (\tilde{u}^n,\tilde{v}^n)\in\mathcal{B}_1(i)\times\mathcal{B}_2(j), (\tilde{u}^n,\tilde{v}^n)\neq(u^n,v^n)\Big\}$$

Denote such an indicator function by $\mathbb{1}^{(\mu)}(u^n,v^n)$. Therefore, (40) is bounded by

$$S^{(n)} \triangleq \frac{(1-\varepsilon)(1-\varepsilon')}{(1+\eta)^2}2^{-n(\tilde{R}_1+\tilde{R}_2)}\sum_{l,k}\sum_{(u^n,v^n)\in\mathcal{T}_\delta}\Omega_{u^n,v^n}\frac{2}{N}\sum_{\mu}\mathbb{1}^{(\mu)}(u^n,v^n)\mathbb{1}\{U^{n,(\mu)}(l)=u^n, V^{n,(\mu)}(k)=v^n\}. \tag{41}$$

Next, we use the Markov inequality to show that $S^{(n)} \leqslant \epsilon$ with probability sufficiently close to 1. We first show that the expectation of $S^{(n)}$ can be made arbitrary small by taking n large enough. For that we take the expectation of the indicator functions with respect to random variables U^n and V^n which are independent of each other and distributed according to (14) and (15), respectively.

$$\mathbb{E}\Big[\mathbb{1}^{(\mu)}(u^n,v^n)\mathbb{1}\{U^{n,(\mu)}(l)=u^n, V^{n,(\mu)}(k)=v^n\}\Big]$$

$$\leqslant \sum_{\substack{(\tilde{u}^n,\tilde{v}^n)\in\mathcal{T}_\delta^{(n)}(AB)\\(\tilde{u}^n,\tilde{v}^n)\neq(u^n,v^n)}} \sum_{i,j} \sum_{(\tilde{l},\tilde{k})\neq(l,k)} \mathbb{E}\big[\mathbb{1}\{(u^n,v^n)\in\mathcal{B}_1(i)\times\mathcal{B}_2(j)\}\mathbb{1}\{(\tilde{u}^n,\tilde{v}^n)\in\mathcal{B}_1(i)\times\mathcal{B}_2(j)\}$$

$$\times\mathbb{1}\{U^{n,(\mu)}(l)=u^n,V^{n,(\mu)}(k)=v^n\}\mathbb{1}\{U^{n,(\mu)}(\tilde{l})=\tilde{u}^n,V^{n,(\mu)}(\tilde{k})=\tilde{v}^n\}\Big]$$

$$\leqslant \frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)^2(1-\varepsilon')^2}2^{-n(I(U;V)-\delta_1)}\Big[2^{n(\tilde{R}_1-R_1)}2^{n(\tilde{R}_2-R_2)}$$

$$+2^{n(\tilde{R}_1-R_1)}+2^{n(\tilde{R}_2-R_2)}+2^{-n(S(U)-\delta_1)}2^{n\tilde{R}_1}2^{n(\tilde{R}_2-R_2)}+2^{-n(S(V)-\delta_1)}2^{n\tilde{R}_2}2^{n(\tilde{R}_1-R_1)}\Big]$$

$$\leqslant 5\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)^2(1-\varepsilon')^2}2^{-n(I(U;V)-2\delta_1)}2^{n(\tilde{R}_1-R_1)}2^{n(\tilde{R}_2-R_2)},$$

where δ_1 is a function of δ (as in (14)) such that $\delta_1 \searrow 0$ as $\delta \searrow 0$, and we have used the inequalities $\tilde{R}_1 < S(U)$ and $\tilde{R}_2 < S(V)$. Hence, given any $\epsilon \in (0,1)$, the above expectation can be made less than $5\epsilon\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)(1-\varepsilon')}$ for large enough n provided that $(\tilde{R}_1-R_1)+(\tilde{R}_2-R_2)\leqslant I(U;V)$. As a result, for large enough n,

$$\mathbb{E}[S^{(n)}]\leqslant\frac{2^{-n(\tilde{R}_1+\tilde{R}_2)}}{(1+\eta)^2}\sum_{l,k}\sum_{(u^n,v^n)\in\mathcal{T}_\delta}10\epsilon\Omega_{u^n,v^n}\lambda^A_{u^n}\lambda^B_{v^n}=\frac{10\epsilon}{(1+\eta)^2}\sum_{(u^n,v^n)\in\mathcal{T}_\delta}\Omega_{u^n,v^n}\lambda^A_{u^n}\lambda^B_{v^n}$$

We proceed by the following lemma.

Lemma 6. *For any $\epsilon\in(0,\frac{1}{3}),\epsilon'\in(0,1)$, and sufficiently small $\delta>0$, there exists $n(\epsilon,\epsilon',\delta)$ such that for all $n\geqslant n(\epsilon,\epsilon',\delta)$ the inequality*

$$\mathbb{P}\left\{\bigcup_{(u^n,v^n)\in\mathcal{T}_\delta}\left\{\frac{1}{N}\sum_\mu\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n}\leqslant\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1+\eta)^2}(1-\epsilon)\right\}\right\}\leqslant\epsilon', \tag{42}$$

holds, provided that

$$\frac{1}{n}\log_2 N+\tilde{R}_1+\tilde{R}_2\geqslant S(U)_{\sigma_3}+S(V)_{\sigma_3}+\delta_1+\delta_2.$$

Proof. The proof is given in Appendix B-B. ☐

From the lemma, with probability at least $(1-\epsilon')$, we have $\frac{1}{N}\sum_\mu\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n}>(1-\epsilon)\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1+\eta)^2}$ for all $(u^n,v^n)\in\mathcal{T}_\delta$. Furthermore, from (38) for $\mathcal{E}=\mathcal{T}^{(n)}_\delta(AB)$, the following holds with probability sufficiently close to 1:

$$\frac{1}{N}\sum_\mu\sum_{(u^n,v^n)\in\mathcal{T}_\delta}\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n}\Omega_{u^n,v^n}\leqslant 2\epsilon+\sum_{(u^n,v^n)\in\mathcal{T}_\delta}\lambda^{AB}_{u^n,v^n}.$$

These two statements imply that there is a realization of $\{U^{n,(\mu)}(l),U^{n,(\mu)}(k)\}_{l,k,\mu}$ for which

$$\frac{10\epsilon}{(1+\eta)^2}\sum_{(u^n,v^n)\in\mathcal{T}_\delta}\Omega_{u^n,v^n}\lambda^A_{u^n}\lambda^B_{v^n}\leqslant\frac{10\epsilon}{(1-\epsilon)}\sum_{(u^n,v^n)\in\mathcal{T}_\delta}\frac{1}{N}\sum_\mu\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n}\Omega_{u^n,v^n}$$

$$\leqslant\frac{10\epsilon}{(1-\epsilon)}(2\epsilon+\sum_{(u^n,v^n)\in\mathcal{T}_\delta}\lambda^{AB}_{u^n,v^n})$$

$$\leqslant \frac{10\epsilon(1+2\epsilon)}{(1-\epsilon)},$$

where the last inequality holds because $\sum_{(u^n,v^n)\in\mathcal{T}_\delta}\lambda^{AB}_{u^n,v^n}\leqslant 1$. As a result, we showed that $\mathbb{E}[S^{(n)}]\leqslant \frac{10\epsilon(1+2\epsilon)}{(1-\epsilon)}$. From, Markov-inequality, this implies that

$$S^{(n)}\leqslant\sqrt{\epsilon} \tag{43}$$

with probability at least $1-\frac{6\sqrt{\epsilon}(1+2\epsilon)}{(1-\epsilon)}$. Using the equations (43) and (39), the quantity in (33) which is an upper bound on the right-hand side of (32) can be bounded from above by $\sqrt{\epsilon}+6\epsilon+2\varepsilon''$, with probability sufficiently close to one.

To sum-up, we showed that the trace distance inequality in (25) holds for sufficiently large n and with probability sufficiently close to 1, if the following bounds hold:

$$\tilde{R}_1\geqslant I(U;RB)_{\sigma_1} \tag{44a}$$

$$\tilde{R}_2\geqslant I(V;RA)_{\sigma_2} \tag{44b}$$

$$C+\tilde{R}_1\geqslant S(U)_{\sigma_3},\quad C+\tilde{R}_2\geqslant S(V)_{\sigma_3},\quad C+\tilde{R}_1+\tilde{R}_2\geqslant S(U)_{\sigma_3}+S(V)_{\sigma_3} \tag{44c}$$

$$(\tilde{R}_1-R_1)+(\tilde{R}_2-R_2)<I(U;V)_{\sigma_3} \tag{44d}$$

$$\tilde{R}_1\geqslant R_1\geqslant 0,\quad \tilde{R}_2\geqslant R_2\geqslant 0,\quad C\geqslant 0, \tag{44e}$$

where $C\triangleq\frac{1}{n}\log_2 N$. This implies that $\tilde{M}_{AB}$ is ϵ-faithful to $\bar{M}_A\otimes\bar{M}_B$ with probability sufficiently close to one, and hence, $\hat{M}_{AB}$ is also ϵ-faithful to M_{AB}, i.e, (24) is satisfied. Therefore, there exists a distributed protocol with parameters $(n,2^{nR_1},2^{nR_2},2^{nC})$ such that its overall POVM $\hat{M}_{AB}$ is ϵ-faithful to M_{AB}. Lastly, we complete the proof of the theorem using the following lemma.

Lemma 7. *Let $\mathcal{R}_1$ denote the set of all (R_1,R_2,C) for which there exists $(\tilde{R}_1,\tilde{R}_2)$ such that the quintuple $(R_1,R_2,C,\tilde{R}_1,\tilde{R}_1)$ satisfies the inequalities in (44). Let, $\mathcal{R}_2$ denote the set of all triples (R_1,R_2,C) that satisfies the inequalities in (5) given in the statement of the theorem. Then, $\mathcal{R}_1=\mathcal{R}_2$.*

Proof. This follows by Fourier-Motzkin elimination [13]. For that, we eliminate $(\tilde{R}_1,\tilde{R}_2)$ from the system of inequalities given by (44). This gives us an equivalent rate-region described by all (C,R_1,R_2) that satisfies the following set of inequalities:

$$R_1\geqslant I(U;RB)_{\sigma_1}-I(U;V)_{\sigma_3}, \tag{45a}$$

$$R_2\geqslant I(V;RA)_{\sigma_2}-I(U;V)_{\sigma_3}, \tag{45b}$$

$$R_1+R_2\geqslant I(U;RB)_{\sigma_1}+I(V;RA)_{\sigma_2}-I(U;V)_{\sigma_3}, \tag{45c}$$

$$C+R_2\geqslant S(V|U)_{\sigma_3}, \tag{45d}$$

$$C + R_1 \geqslant S(U|V)_{\sigma_3}, \tag{45e}$$

$$2C + R_1 + R_2 \geqslant S(U,V)_{\sigma_3}, \tag{45f}$$

$$C + R_1 + R_2 \geqslant S(U,V)_{\sigma_3}, \tag{45g}$$

$$C + R_1 + R_2 \geqslant I(V;RA)_{\sigma_1} + S(U|V)_{\sigma_3}, \tag{45h}$$

$$C + R_1 + R_2 \geqslant I(U;RB)_{\sigma_2} + S(V|U)_{\sigma_3} \tag{45i}$$

Having (45g), the bounds given by (45f),(45h), and (45i) are redundant. Hence we get the rate-region $\mathcal{R}_2$, and the proof is complete. $\square$

Appendix B
Proof of Lemma 5 and 6

A. Proof of Lemma 5

Let us denote the operators of the given POVMs $M_X^{(\mu)}$ and $M_Y^{(\mu)}$ by $\{\Lambda_u^{X(\mu)} : u \in \mathcal{U}\}$ and $\{\Gamma_v^{Y(\mu)} : v \in \mathcal{V}\}$, respectively, where $\mathcal{U}, \mathcal{V}$ are two finite sets. Further, denote the operators of the approximating POVM $\hat{M}_X^{(\mu)}$ by $\{\hat{\Lambda}_u^{X(\mu)} : u \in \mathcal{U}\}$. Let $\Psi_{R'XY}^{\rho_{XY}}$ be any purification of ρ_{XY}. The left-hand side of equation (28) can be simplified as

$$\begin{aligned}
&\Big\| \frac{1}{N}\sum_{\mu}\Big[(\mathrm{id}\otimes M_X^{(\mu)}\otimes M_Y^{(\mu)})(\Psi_{R'XY}^{\rho_{XY}}) - (\mathrm{id}\otimes \hat{M}_X^{(\mu)}\otimes M_Y^{(\mu)})(\Psi_{R'XY}^{\rho_{XY}})\Big]\Big\|_1 \\
&\quad = \Big\| \frac{1}{N}\sum_{\mu}\Big[(\mathrm{id}\otimes (M_X^{(\mu)} - \hat{M}_X^{(\mu)})\otimes M_Y^{(\mu)})(\Psi_{R'XY}^{\rho_{XY}})\Big]\Big\|_1 \\
&\quad = \Big\| \frac{1}{N}\sum_{\mu}\Big[\sum_{u,v}|u\rangle\langle u|\otimes|v\rangle\langle v|\otimes \mathrm{Tr}_{XY}\{(\mathrm{id}\otimes(\Lambda_u^{X(\mu)} - \hat{\Lambda}_u^{X(\mu)})\otimes\Gamma_v^{Y(\mu)})(\Psi_{R'XY}^{\rho_{XY}})\}\Big]\Big\|_1 \\
&\quad = \Big\| \frac{1}{N}\sum_{\mu}\Big[\sum_{u,v}|u\rangle\langle u|\otimes|v\rangle\langle v|\otimes \mathrm{Tr}_{XY}\{(\mathrm{id}\otimes(\Lambda_u^{X(\mu)} - \hat{\Lambda}_u^{X(\mu)})\otimes\mathrm{id}_Y)(\mathrm{id}\otimes\mathrm{id}_X\otimes\Gamma_v^{Y(\mu)})(\Psi_{R'XY}^{\rho_{XY}})\}\Big]\Big\|_1 \\
&\quad = \Big\| \frac{1}{N}\sum_{\mu}\Big[\sum_{u}|u\rangle\langle u|\otimes \mathrm{Tr}_{X}\{(\mathrm{id}_V\otimes\mathrm{id}\otimes(\Lambda_u^{X(\mu)} - \hat{\Lambda}_u^{X(\mu)})(\sum_{v}|v\rangle\langle v|\otimes\mathrm{Tr}_Y\{(\mathrm{id}\otimes\mathrm{id}_X\otimes\Gamma_v^{Y(\mu)})(\Psi_{R'XY}^{\rho_{XY}})\}\}\Big]\Big\|_1
\end{aligned}$$

The last equality first uses the property that $\mathrm{Tr}_{XY}\{\} = \mathrm{Tr}_X\{\mathrm{Tr}_Y\{\}\}$, followed by the definition of partial trace and its linearity. Here, defining $\sigma_{R'XV}$ as

$$\sigma_{R'XV} = (\mathrm{id}\otimes\mathrm{id}_X\otimes M_Y^{(\mu)})\Psi_{R'XY}^{\rho_{XY}} = \sum_{v}|v\rangle\langle v|\otimes\mathrm{Tr}_Y\{(\mathrm{id}\otimes\mathrm{id}_X\otimes\Gamma_v^{Y(\mu)})(\Psi_{R'XY}^{\rho_{XY}})$$

gives us,

$$\| \frac{1}{N}\sum_{\mu}\Big[(\mathrm{id}\otimes M_X^{(\mu)}\otimes M_Y^{(\mu)})(\Psi_{R'XY}^{\rho_{XY}}) - (\mathrm{id}\otimes \hat{M}_X^{(\mu)}\otimes M_Y^{(\mu)})(\Psi_{R'XY}^{\rho_{XY}})\Big]\|_1$$

$$= \| \frac{1}{N}\sum_{\mu}\left[\sum_{u}|u\rangle\langle u|\otimes \mathrm{Tr}_X\{(\mathrm{id}_V\otimes \mathrm{id}\otimes (\Lambda_u^{X(\mu)}-\hat{\Lambda}_u^{X(\mu)}))(\sigma_{R'XV})\}\right] \|_1 \tag{46}$$

Further, by defining $\Phi^{\sigma_{R'XV}}_{R'XVY'}$ to be any purification of $\sigma_{R'XV}$, we get

$$\begin{aligned}
&\left\|\frac{1}{N}\sum_{\mu}\Big[(id\otimes M_X^{(\mu)}\otimes M_Y^{(\mu)})(\Psi^{\rho_{XY}}_{R'XY})-(\mathrm{id}\otimes \hat{M}_X^{(\mu)}\otimes M_Y^{(\mu)})(\Psi^{\rho_{XY}}_{R'XY})\Big]\right\|_1\\
&\quad=\left\|\frac{1}{N}\sum_{\mu}\left[\sum_{u}|u\rangle\langle u|\otimes \mathrm{Tr}_X\{(\mathrm{id}\otimes (\Lambda_u^{X(\mu)}-\hat{\Lambda}_u^{X(\mu)})\otimes \mathrm{id}_V)\,\mathrm{Tr}_{Y'}\{\Phi^{\sigma_{R'XV}}_{R'XVY'}\}\}\right]\right\|_1\\
&\quad=\left\|\frac{1}{N}\sum_{\mu}\left[\sum_{u}|u\rangle\langle u|\otimes \mathrm{Tr}_{XY'}\{(\mathrm{id}\otimes (\Lambda_u^{X(\mu)}-\hat{\Lambda}_u^{X(\mu)})\otimes \mathrm{id}_{VY'})\Phi^{\sigma_{R'XV}}_{R'XVY'}\}\right]\right\|_1\\
&\quad=\left\|\mathrm{Tr}_{Y'}\left\{\frac{1}{N}\sum_{\mu}\left[\sum_{u}|u\rangle\langle u|\otimes \mathrm{Tr}_X\{(\mathrm{id}\otimes (\Lambda_u^{X(\mu)}-\hat{\Lambda}_u^{X(\mu)})\otimes \mathrm{id}_{VY'})\Phi^{\sigma_{R'XV}}_{R'XVY'}\}\right]\right\}\right\|_1\\
&\quad\leqslant\left\|\frac{1}{N}\sum_{\mu}\left[\sum_{u}|u\rangle\langle u|\otimes \mathrm{Tr}_X\{(\mathrm{id}\otimes (\Lambda_u^{X(\mu)}-\hat{\Lambda}_u^{X(\mu)})\otimes \mathrm{id}_{VY'})\Phi^{\sigma_{R'XV}}_{R'XVY'}\}\right]\right\|_1\\
&\quad=\left\|\frac{1}{N}\sum_{\mu}\Big[(\mathrm{id}_{R'VY'}\otimes M_X^{(\mu)})\Phi^{\sigma_{R'XV}}_{R'XVY'}-(\mathrm{id}_{R'VY'}\otimes \hat{M}_X^{(\mu)})\Phi^{\sigma_{R'XV}}_{R'XVY'}\Big]\right\|_1 &(47)\\
&\quad=\left\|\frac{1}{N}\sum_{\mu}\big[(\mathrm{id}\otimes M_X^{(\mu)})(\Psi^{\rho_X}_{RX})-(\mathrm{id}\otimes \hat{M}_X^{(\mu)})(\Psi^{\rho_X}_{RX})\big]\right\|_1 &(48)\\
&\quad\leqslant \epsilon &(49)
\end{aligned}$$

The first three equalities in equation (47) exploit the linearity of trace, and the first inequality follows from the monotonicity of the trace-distance as given in Corollary 9.1.2 [9]. The last equality is established by first observing that $\Phi^{\sigma_{R'XV}}_{R'XVY'}$ is also a purification of ρ_X, i.e.,

$$\mathrm{Tr}_{R'VY'}\{\Phi^{\sigma_{R'XV}}_{R'XVY'}\}=\mathrm{Tr}_{RV'}\{\sigma_{VXR'}\}=\mathrm{Tr}_{R'}\{\rho_{R'X}\}=\rho_X,$$

followed by application of Lemma 1 to $\Phi^{\sigma_{R'XV}}_{R'XVY'}$ which is true for any purification of ρ_X. This completes the proof.

B. *Proof of Lemma 6*

We use Suen's Inequality (Theorem 2.23, [14]) to prove this lemma. For that we need the following definition:

Dependency Graph: Let $\{X_i\}_{i\in\mathcal{I}}$ be a family of random variables defined on a common probability space. A dependency graph for $\{X_i\}$ is a graph L with vertex set $V(L)=\mathcal{I}$ such that, if $\mathcal{A}$ and $\mathcal{B}$ are two disjoint subsets of $\mathcal{I}$ with no edge between them, then the sub-families $\{X_i\}_{i\in\mathcal{A}}$ and $\{X_i\}_{i\in\mathcal{B}}$ are mutually independent.

Theorem 5 ([14]). *Let $I_i \in Be(p_i), i \in \mathcal{I}$ be a family of Bernoulli random variables having a dependency graph L with vertex set $\mathcal{I}$ and edge set $E(L)$. Let $X = \sum_i I_i$ and $\lambda = \mathbb{E}[X]$. Moreover, write $i \sim j$ if $(i,j) \in E(L)$ and let $\Delta \triangleq \frac{1}{2}\sum_i \sum_{j\sim i} \mathbb{E}[I_i I_j]$, $\bar{\Delta} \triangleq \lambda + 2\Delta$ and $\delta \triangleq \max_i \sum_{k\sim i} p_k$. Then for $\epsilon \in (0,1)$*

$$\mathbb{P}(X \leqslant (1-\epsilon)\lambda) \leqslant \exp\left\{-\min\left(\frac{\epsilon^2\lambda^2}{4\bar{\Delta}}, \frac{\epsilon\lambda}{6\delta}\right)\right\}$$

We start by defining the Bernoulli random variables $\kappa_{l,k}^{(\mu)}(u^n, v^n)$ for $(u^n, v^n) \in \mathcal{T}_\delta(AB)$ as

$$\kappa_{l,k}^{(\mu)}(u^n, v^n) \triangleq \mathbb{1}\{U^{n,(\mu)}(l) = u^n, V^{n,(\mu)}(k) = v^n\}.$$

Note that $(\kappa_{l,k}^{(\mu)}(u^n, v^n), \kappa_{l,k}^{(\mu')}(u^n, v^n))$ are independent for $\mu \neq \mu'$ while $(\kappa_{l,k}^{(\mu)}(u^n, v^n), \kappa_{l',k}^{(\mu)}(u^n, v^n))$ and $(\kappa_{l,k}^{(\mu)}(u^n, v^n), \kappa_{l,k'}^{(\mu)}(u^n, v^n))$ are not independent. This gives us

$$X_{(u^n,v^n)} = \sum_{\mu,l,k} \kappa_{l,k}^{(\mu)}(u^n, v^n) = \frac{(1+\eta)^2}{(1-\varepsilon)(1-\varepsilon')} 2^{n(\tilde{R}_1+\tilde{R}_2)} \sum_{\mu} \gamma_{u^n}^{(\mu)} \zeta_{v^n}^{(\mu)}, \tag{50}$$

and

$$\begin{aligned}
\lambda_{(u^n,v^n)} = \mathbb{E}[X_{(u^n,v^n)}] &= \sum_{\mu,l,k} \mathbb{E}[\kappa_{l,k}^{(\mu)}(u^n, v^n)] \\
&= \sum_{\mu,l,k} \mathbb{P}\{U^{n,(\mu)}(l) = u^n, V^{n,(\mu)}(k) = v^n\} \\
&= \sum_{\mu,l,k} \frac{\lambda_{u^n}^A}{1-\varepsilon} \frac{\lambda_{v^n}^B}{1-\varepsilon'} \\
&= N \cdot 2^{n(\tilde{R}_1+\tilde{R}_2)} \frac{\lambda_{u^n}^A \lambda_{v^n}^B}{(1-\varepsilon)(1-\varepsilon')}.
\end{aligned} \tag{51}$$

The above calculation of $\mathbb{E}[X_{(u^n,v^n)}]$ uses the distribution defined in (14) and (15). We can also compute $\delta_{(u^n,v^n)}$ and $\Delta_{(u^n,v^n)}$ as

$$\begin{aligned}
\delta_{(u^n,v^n)} &= \max_{\mu,l,k} \sum_{(\mu',l',k')\sim(\mu,l,k)} \mathbb{E}[\kappa_{l,k}^{(\mu)}(u^n, v^n)] \\
&= \max_{\mu,l,k} \left[\sum_{(l':l'\neq l)} \frac{\lambda_{u^n}^A \lambda_{v^n}^B}{(1-\varepsilon)(1-\varepsilon')} + \sum_{(k':k'\neq k)} \frac{\lambda_{u^n}^A \lambda_{v^n}^B}{(1-\varepsilon)(1-\varepsilon')} \right] \\
&= \max_{\mu,l,k} \frac{\lambda_{u^n}^A \lambda_{v^n}^B}{(1-\varepsilon)(1-\varepsilon')} (2^{n\tilde{R}_1} + 2^{n\tilde{R}_2} - 2) = \frac{\lambda_{u^n}^A \lambda_{v^n}^B}{(1-\varepsilon)(1-\varepsilon')} (2^{n\tilde{R}_1} + 2^{n\tilde{R}_2} - 2).
\end{aligned}$$

The first equality uses the simplification from (51), and the second equality follows from the dependency relations between $\kappa_{l,k}^{(\mu)}(u^n, v^n)$. Similarly,

$$\Delta_{(u^n,v^n)} = \frac{1}{2} \sum_{\mu,l,k} \sum_{(\mu',l',k')\sim(\mu,l,k)} \mathbb{E}[\kappa_{l,k}^{(\mu)}(u^n, v^n) \cdot \kappa_{l',k'}^{(\mu')}(u^n, v^n)]$$

$$= \frac{1}{2}\sum_{\mu,l,k}\left[\sum_{(l':l'\neq l)} \mathbb{E}[\kappa_{l,k}^{(\mu)}(u^n,v^n)\cdot \kappa_{l',k}^{(\mu)}(u^n,v^n))] + \sum_{(k':k'\neq k)} \mathbb{E}[\kappa_{l,k}^{(\mu)}(u^n,v^n)\cdot \kappa_{l,k'}^{(\mu)}(u^n,v^n))]\right]$$

$$= \frac{1}{2}\sum_{\mu,l,k}\left[\sum_{(l':l'\neq l)} \frac{(\lambda^A_{u^n})^2\lambda^B_{v^n}}{(1-\varepsilon)^2(1-\varepsilon')} + \sum_{(k':k'\neq k)} \frac{\lambda^A_{u^n}(\lambda^B_{v^n})^2}{(1-\varepsilon)(1-\varepsilon')^2}\right]$$

$$= \frac{N}{2}2^{n(\tilde{R}_1+\tilde{R}_2)}\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)(1-\varepsilon')}\left[(2^{n\tilde{R}_1}-1)\frac{\lambda^A_{u^n}}{(1-\varepsilon)} + (2^{n\tilde{R}_2}-1)\frac{\lambda^B_{v^n}}{(1-\varepsilon')}\right].$$

Now we evaluate the exponent in the statement of the Suen's Inequality given by $-\min\left(\frac{\epsilon^2\lambda^2}{4\bar{\Delta}}, \frac{\epsilon\lambda}{6\delta}\right)$, where we omit the subscript (u^n,v^n) in the terms δ, Δ and λ. Consider,

$$\frac{\epsilon^2\lambda^2}{4\bar{\Delta}} = \frac{\left(\epsilon N\cdot 2^{n(\tilde{R}_1+\tilde{R}_2)}\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)(1-\varepsilon')}\right)^2}{4\left(N2^{n(\tilde{R}_1+\tilde{R}_2)}\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)(1-\varepsilon')}\right)\left((2^{n\tilde{R}_1}-1)\frac{\lambda^A_{u^n}}{(1-\varepsilon)} + (2^{n\tilde{R}_2}-1)\frac{\lambda^B_{v^n}}{(1-\varepsilon')}+1\right)}$$

$$= \epsilon^2\frac{\left(N\cdot 2^{n(\tilde{R}_1+\tilde{R}_2)}\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)(1-\varepsilon')}\right)}{4\left((2^{n\tilde{R}_1}-1)\frac{\lambda^A_{u^n}}{(1-\varepsilon)} + (2^{n\tilde{R}_2}-1)\frac{\lambda^B_{v^n}}{(1-\varepsilon')}+1\right)}$$

$$\leqslant \epsilon^2\frac{\left(N\cdot 2^{n(\tilde{R}_1+\tilde{R}_2)}\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)(1-\varepsilon')}\right)}{4\left(2^{n\tilde{R}_1}\frac{\lambda^A_{u^n}}{(1-\varepsilon)} + 2^{n\tilde{R}_2}\frac{\lambda^B_{v^n}}{(1-\varepsilon')}\right)} = \epsilon^2\frac{N\cdot 2^{n(\tilde{R}_1+\tilde{R}_2)}}{4\left(2^{n\tilde{R}_1}\frac{(1-\varepsilon')}{\lambda^B_{v^n}} + 2^{n\tilde{R}_2}\frac{(1-\varepsilon)}{\lambda^A_{u^n}}\right)}. \tag{52}$$

where the inequality follows, since for large enough n, we have $\left(1-\frac{\lambda^A_{u^n}}{(1-\varepsilon)} - \frac{\lambda^B_{v^n}}{(1-\varepsilon')}\right) \geqslant 0$. As for the other term in the exponent, we have

$$\frac{\epsilon\lambda}{6\delta} = \frac{\epsilon N\cdot 2^{n(\tilde{R}_1+\tilde{R}_2)}\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)(1-\varepsilon')}}{6\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)(1-\varepsilon')}(2^{n\tilde{R}_1}+2^{n\tilde{R}_2}-2)} \geqslant \frac{\epsilon N\cdot 2^{n(\tilde{R}_1+\tilde{R}_2)}}{6(2^{n\tilde{R}_1}+2^{n\tilde{R}_2})}. \tag{53}$$

Assuming $\varepsilon,\varepsilon' \in (0,\frac{1}{2})$ and $\epsilon \in (0,\frac{1}{3})$, we have $3\lambda^A_{u^n}\epsilon \leqslant 3\epsilon < 2(1-\varepsilon)$, hence $\frac{3}{2}\epsilon \leqslant \frac{(1-\varepsilon)}{\lambda^A_{u^n}}$ and similarly, $\frac{3}{2}\epsilon \leqslant \frac{(1-\varepsilon')}{\lambda^B_{v^n}}$. These inequalities provide an upper bound for (52) as in the following:

$$\epsilon^2\frac{N\cdot 2^{n(\tilde{R}_1+\tilde{R}_2)}}{4\left(2^{n\tilde{R}_1}\frac{(1-\varepsilon')}{\lambda^B_{v^n}} + 2^{n\tilde{R}_2}\frac{(1-\varepsilon)}{\lambda^A_{u^n}}\right)} \leqslant \epsilon^2\frac{N\cdot 2^{n(\tilde{R}_1+\tilde{R}_2)}}{4\left(2^{n\tilde{R}_1}\frac{3}{2}\epsilon + 2^{n\tilde{R}_2}\frac{3}{2}\epsilon\right)} = \frac{\epsilon N\cdot 2^{n(\tilde{R}_1+\tilde{R}_2)}}{6(2^{n\tilde{R}_1}+2^{n\tilde{R}_2})} \leqslant \frac{\epsilon\lambda}{6\delta}.$$

where the last inequality follows from (53). This implies

$$\exp\left\{-\min\left(\frac{\epsilon^2\lambda^2}{4\bar{\Delta}}, \frac{\epsilon\lambda}{6\delta}\right)\right\} = \exp\left\{-\frac{\epsilon^2\lambda^2}{4\bar{\Delta}}\right\}$$

Further, $\exp\left\{-\frac{\epsilon^2\lambda^2}{4\bar{\Delta}}\right\}$ can be simplified as

$$\exp\left\{-\frac{\epsilon^2\lambda^2}{4\bar{\Delta}}\right\} = \exp\left\{-\epsilon^2\frac{\left(N\cdot 2^{n(\tilde{R}_1+\tilde{R}_2)}\frac{\lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)(1-\varepsilon')}\right)}{4\left((2^{n\tilde{R}_1}-1)\frac{\lambda^A_{u^n}}{(1-\varepsilon)} + (2^{n\tilde{R}_2}-1)\frac{\lambda^B_{v^n}}{(1-\varepsilon')}+1\right)}\right\}$$

Next, to provide an upper bound to the above expression, we bound the denominator from above. Using the fact that $(u^n,v^n) \in \mathcal{T}_\delta^{(n)}(AB)$, note that $\lambda^A_{u^n} \leqslant 2^{-n(S(U)-\delta_1)}$ and $\lambda^B_{v^n} \leqslant 2^{-n(S(V)-\delta_2)}$ for all sufficiently large n, where $\delta_i \searrow 0, i = 1, 2$, as $\delta \searrow 0$. From the assumption that $\tilde{R}_1 < S(U)_{\sigma_3}$ and $\tilde{R}_2 < S(V)_{\sigma_3}$, we can choose the parameter δ in $\mathcal{T}_\delta^{(n)}(AB)$ sufficiently small so that $R_1 < S(U)_{\sigma_3} - \delta_1$ and $R_2 < S(V)_{\sigma_3} - \delta_2$. Hence, we get

$$4\left((2^{n\tilde{R}_1} - 1)\frac{\lambda^A_{u^n}}{(1-\varepsilon)} + (2^{n\tilde{R}_2} - 1)\frac{\lambda^B_{v^n}}{(1-\varepsilon')} + 1\right) \leqslant 4\Big(\frac{2^{n(\tilde{R}_1 - S(U)+\delta_1)}}{(1-\varepsilon)} + \frac{2^{n(\tilde{R}_2 - S(V)+\delta_2)}}{(1-\varepsilon')} + 1\Big)$$
$$\leqslant 12$$

As a result of this inequality and using the fact that for all sufficiently large n, $\lambda^A_{u^n} \geqslant 2^{-n(S(U)+\delta_1)}$ and $\lambda^B_{v^n} \geqslant 2^{-n(S(V)+\delta_2)}$, we obtain

$$\exp\left\{-\frac{\epsilon^2\lambda^2}{4\bar{\Delta}}\right\} \leqslant \exp\left\{-\frac{N\epsilon^2}{12(1-\varepsilon)(1-\varepsilon')}\ 2^{n(\tilde{R}_1+\tilde{R}_2)}2^{-n(S(U)_\sigma+S(V)_\sigma+\delta_1+\delta_2)}\right\}.$$

This implies that for

$$\frac{1}{n}\log N + \tilde{R}_1 + \tilde{R}_2 \geqslant S(U)_\sigma + S(V)_\sigma + \delta_1 + \delta_2,$$

we get $\exp\left\{-\frac{\epsilon^2\lambda^2}{4\bar{\Delta}}\right\} \searrow 0$ as $\epsilon \searrow 0$.

Using the above definitions of $X_{(u^n,v^n)}, \delta_{(u^n,v^n)}, \Delta_{(u^n,v^n)}$, we can now write Suen's inequality as

$$\mathbb{P}\left\{\sum_{\mu,l,k}\kappa^{(\mu)}_{l,k}(u^n,v^n) \leqslant N\cdot 2^{n(\tilde{R}_1+\tilde{R}_2)}\frac{\lambda^A_{u^n}\lambda^B_{v^n}(1-\epsilon)}{(1-\varepsilon)(1-\varepsilon')}\right\} \leqslant \exp\left\{-\frac{\epsilon^2\lambda^2}{4\bar{\Delta}}\right\}.$$

Therefore, from (50), the following inequality holds

$$\mathbb{P}\left\{\frac{1}{N}\sum_{\mu}\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n} \leqslant \frac{\lambda^A_{u^n}\lambda^B_{v^n}(1-\epsilon)}{(1+\eta)^2}\right\} \leqslant \exp\left\{-\frac{\epsilon^2\lambda^2}{4\bar{\Delta}}\right\}. \tag{54}$$

By taking the union over all $(u^n, v^n) \in \mathcal{T}_\delta^{(n)}(AB)$, the following upper-bound is obtained

$$\mathbb{P}\left\{\bigcup_{(u^n,v^n)\in\mathcal{T}_\delta}\left\{\frac{1}{N}\sum_{\mu}\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n} \leqslant \frac{\lambda^A_{u^n}\lambda^B_{v^n}(1-\epsilon)}{(1+\eta)^2}\right\}\right\} \leqslant \sum_{(u^n,v^n)\in\mathcal{T}_\delta}\mathbb{P}\left\{\frac{1}{N}\sum_{\mu}\gamma^{(\mu)}_{u^n}\zeta^{(\mu)}_{v^n} \leqslant \frac{\lambda^A_{u^n}\lambda^B_{v^n}(1-\epsilon)}{(1+\eta)^2}\right\}$$
$$\leqslant |\mathcal{T}_\delta^{(n)}(AB)|\exp\left\{-\frac{\epsilon^2\lambda^2}{4\bar{\Delta}}\right\} \leqslant \epsilon',$$

for sufficiently large n. This proves the lemma.

References

[1] M. M. Wilde, P. Hayden, F. Buscemi, and M.-H. Hsieh, “The information-theoretic costs of simulating quantum measurements,” *Journal of Physics A: Mathematical and Theoretical*, vol. 45, no. 45, p. 453001, 2012.

[2] I. Devetak, “The private classical capacity and quantum capacity of a quantum channel,” *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 44–55, 2005b.

[3] N. Datta, M.-H. Hsieh, M. M. Wilde, and A. Winter, “Quantum-to-classical rate distortion coding,” *Journal of Mathematical Physics*, vol. 54, no. 4, p. 042201, 2013.

[4] N. Datta, M. H. Hsieh, and M. M. Wilde, “Quantum rate distortion, reverse Shannon theorems, and source-channel separation,” *IEEE Transactions on Information Theory*, vol. 59, pp. 615–630, 2013.

[5] A. Winter, “”Extrinsic” and ”intrinsic” data in quantum measurements: asymptotic convex decomposition of positive operator valued measures,” *Communication in Mathematical Physics*, vol. 244, no. 1, pp. 157–185, 2004.

[6] T. Berger, “Multiterminal source coding,” *The Inform. Theory Approach to Communications, G. Longo, Ed., New York: Springer-Verlag*, 1977.

[7] I. Devetak and A. Winter, “Classical data compression with quantum side information,” *Physical Review A*, vol. 68, no. 4, p. 042301, 2003.

[8] A. Anshu, R. Jain, and N. A. Warsi, “A generalized quantum Slepian–Wolf,” *IEEE Transactions on Information Theory*, vol. 64, no. 3, pp. 1436–1453, March 2018.

[9] M. M. Wilde, “From classical to quantum shannon theory,” *arXiv preprint arXiv:1106.1445*, 2011.

[10] A. S. Holevo, *Quantum systems, channels, information: a mathematical introduction*. Walter de Gruyter, 2012, vol. 16.

[11] F. Shirani and S. S. Pradhan, “Finite block-length gains in distributed source coding,” in *2014 IEEE International Symposium on Information Theory*, June 2014, pp. 1702–1706.

[12] J. B. Conway, *A Course in Functional Analysis*. Springer New York, 1985.

[13] G. M. Ziegler, *Lectures on polytopes*. Springer Science & Business Media, 2012, vol. 152.

[14] S. Janson, T. Luczak, and A. Rucinski, *Random graphs*. John Wiley & Sons, 2011, vol. 45.