

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

Touheed Anwar Atif <sup>\*</sup>, Mohsen Heidari <sup>†</sup> and S. Sandeep Pradhan <sup>\*</sup>,

<sup>\*</sup> EECS Dept., University of Michigan, Ann Arbor, MI.

<sup>†</sup> CS Dept., Purdue University, West Lafayette, IN.

<sup>\*</sup>{`touheed`, `pradhanv`}@umich.edu, <sup>†</sup>`mheidari`@purdue.edu

## Abstract

We consider the task of faithfully simulating a distributed quantum measurement, wherein we provide a protocol for the three parties, Alice, Bob and Eve, to simulate a repeated action of a distributed quantum measurement using a pair of non-product approximating measurements by Alice and Bob, followed by a stochastic mapping at Eve. The objective of the protocol is to utilize minimum resources, in terms of classical bits needed by Alice and Bob to communicate their measurement outcomes to Eve, and the common randomness shared among the three parties, while faithfully simulating independent repeated instances of the original measurement. To achieve this, we develop a mutual covering lemma and a technique for random binning of distributed quantum measurements, and, in turn, characterize a set of sufficient communication and common randomness rates required for asymptotic simulatability in terms of single-letter quantum information quantities. Furthermore, using these results we address a distributed quantum rate-distortion problem, where we characterize the achievable rate distortion region through a single-letter inner bound. Finally, via a technique of single-letterization of multi-letter quantum information quantities, we provide an outer bound for the rate-distortion region.

## CONTENTS

<b>I</b>	<b>Introduction</b>	2
<b>II</b>	<b>Preliminaries</b>	4
<b>III</b>	<b>Simulation of Distributed POVMs with Deterministic Processing</b>	6
<b>IV</b>	<b>Proof of Theorem 2</b>	9
<b>V</b>	<b>Q-C Distributed Rate Distortion Theory</b>	14
<b>VI</b>	<b>Simulation of POVMs with Stochastic Processing</b>	20
<b>VII</b>	<b>Simulation of Distributed POVMs with Stochastic Processing</b>	24
<b>VIII</b>	<b>Proof of Theorem 6</b>	26
<b>IX</b>	<b>Conclusion</b>	31
	<b>Appendix A: Proof of Lemmas</b>	31
	<b>Appendix B: Proof of Propositions</b>	34
	<b>References</b>	41

## I. INTRODUCTION

**M**easurements interface the intricate quantum world with the perceivable macroscopic classical world by associating a classical attribute to a quantum state. 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 [1] (also elaborated in [2]) quantifies the “relevant information” as the amount of resources needed to simulate the output of a quantum measurement applied on a given state in an asymptotic sense. Imagine that an agent (Alice) performs a measurement  $M$  on a quantum state  $\rho$ , and sends a set of classical bits to a receiver (Bob). Bob intends to *faithfully* recover the outcomes of Alice’s measurements without having access to  $\rho$ . 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*, where  $R$  denotes a reference of the quantum state, and  $X$  denotes the auxiliary register corresponding to the random measurement outcome. Wilde et al. [2] extended the measurement compression problem by considering additional resources available to each of the participating parties. One such formulation allows Bob to further process the information received from Alice using local private randomness. In analogy with [3], this problem formulation is referred to as non-feedback measurement simulation, while the former is termed as simulation with feedback. This quantified the benefit of private randomness in terms of enhancing the trade-off between classical bits communicated and common random bits consumed. In particular, the use of private randomness increases the requirement of classical communication bits, while reducing the common randomness constraint.

The measurement compression theorem finds applications in several paradigms including local purity distillation [2] and private classical communication over quantum channels [4]. This theorem was later used by Datta, et al. [5] to develop a quantum-to-classical (q-c) 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 this problem, the objective is to minimize the storage of the classical outputs resulting from the measurement, while being able to recover the quantum state (from classical bits) within a fixed level of distortion as measured by an observable. To achieve this, the authors in [6] 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. Further, the problem of measurement compression in the presence of quantum side information was studied in [2]. The authors here combined the ideas from [1] and [7] to reduce the classical communication rate and common randomness needed to simulate a measurement in presence of quantum side information. Recently, authors in [8] came up with a completely different technique for analyzing the measurement simulation protocols, while considering the problem of quantum measurement compression with side information. They provide a protocol based on convex-split and position based decoding, and bound rates from above in terms of smooth max and hypothesis testing relative entropies (defined in [8]).

In this work, we consider scenarios where the quantum measurements are performed in a distributed fashion on bipartite entangled states, and quantify “relevant information” for these distributed quantum measurements in an asymptotic sense. As shown in Fig. 1, a composite bipartite quantum system  $AB$  is made available to two agents, Alice and Bob, where they have access to the 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

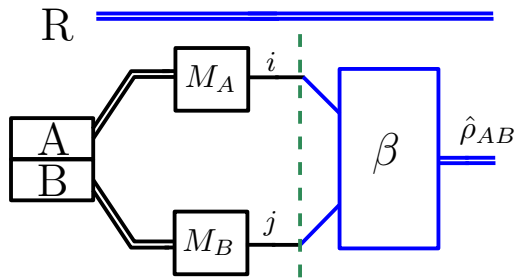


Fig. 1. The diagram of a distributed quantum measurement applied to a bipartite quantum system  $AB$ . A tensor product measurement  $M_A \otimes M_B$  is performed on many copies of the observed quantum state. The outcomes of the measurements are given by two classical bits. The receiver functions as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

Bob. Imagine that there is a third party, Eve, who is connected to Alice and Bob via two separate classical links. The objective of the three parties is to simulate the action of repeated independent measurements performed on many independent copies of the given composite state. To achieve this objective, Alice and Bob send classical bits to Eve at rate  $R_1$  and  $R_2$ , respectively. Further, common randomness at rate  $C$  is also shared amongst the three parties. Eve performs classical processing of the received bits and common randomness. We study two settings, based on whether or not Eve has access to private randomness. As an application of this quantification, we consider the quantum-to-classical distributed rate distortion problem where Eve is allowed to use classical-to-quantum channels. In this work, we focus on memoryless quantum systems in finite dimensional Hilbert spaces. We summarize the contributions of this work in the following:

- We formulate the problem of faithful simulation of distributed quantum measurements that can be decomposed as a convex-linear combination (incorporating Eve's stochastic processing) of separable measurements, as stated in Definition 1. The asymptotic performance limit for this problem is given by the set of all communication rates  $(R_1, R_2)$  and all common randomness rates  $C$ , referred to as the achievable rate region, under which the above-stated measurement is distributively simulated. We devise a distributed simulation protocol for this problem, and provide a quantum-information theoretic inner bound to the achievable rate region in terms of computable single-letter information quantities (see Theorem 6). This is the first main result of the paper.
- As an immediate application of our results on the simulation of distributed measurements, we develop an approach for a distributed quantum-to-classical rate distortion theory, where the objective is to reconstruct a quantum state at Eve, with the quality of reconstruction measured using an additive distortion observable. The asymptotic performance limit is given by the set of all communication rate pairs  $(R_1, R_2)$  at which the distortion  $D$  is achieved. For the achievability part, we characterize an inner bound in terms of single-letter quantum mutual information quantities (see Theorem 3). This is the second main result of the paper. The classical version of this result is called the Berger-Tung inner bound [9].
- We then develop a technique for deriving converse bounds based on a combination of *tensor-product* and *direct-sum* Hilbert spaces (also referred to as a multi-particle system). Using this technique, we derive a single-letter outer-bound on the optimal rate distortion region (see Theorem 4), by converting a multi-letter expression into a single-letter expression. This is the third main result of the paper.

The organization of the paper is as follows. In Section II, we set the notation and state requisite definitions. Instead of first presenting the above-stated inner bound to the performance limits of the problem of simulation of distributed measurements in its full generality, for pedagogical reasons, in Section III, we first consider a special case, where the processing at Eve is restricted to a deterministic function, and provide a simple proof based on the application of Winter's measurement compression theorem. In this setup, we compress each individual measurements  $M_A$  and  $M_B$ , comprising the decomposing of  $M_{AB}$ . As a result, faithful simulation of  $M_A$  is possible when at

least  $nI(U; R_A)$  classical bits of communication and  $nS(U|R_A)$  bits of common randomness are available between Alice and Eve. Similarly, a faithful simulation of  $M_B$  is possible with  $nI(V; R_B)$  classical bits of communication and  $nS(V|R_B)$  bits of common randomness between Eve and Bob, where  $R_A$  and  $R_B$  are purifications of the sub-systems  $A$  and  $B$ , respectively, and  $U$  and  $V$  denote the auxiliary registers corresponding to their measurement outcomes. 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 of the overall measurement,  $M_{AB}$ . To accomplish this, we develop a Mutual Covering Lemma (see Lemma 4), which also helps in converting the information quantities in terms of the reference  $R$  of the joint state  $\rho_{AB}$ .

Further, an interesting aspect about the distributed setting is that 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 state  $\rho_{AB}$ ) are not independent identically distributed (IID) sequences — rather they are codewords generated from random coding. For this we develop a proposition for mutual packing (Proposition 2), that characterizes the binning rates in term of single-letter information quantities. This issue also arises in classical distributed source coding problem which was addressed by Wyner-Ahlsvede-Körner [9] by developing Markov lemma and Mutual packing lemma. The idea of binning in quantum setting has been explored from a different perspective in [10] and [11] for quantum data compression involving side information. Toward the end of the section, we also provide an example to illustrate the inner bound to the achievable rate region.

In Section V, we apply this special setting of the distributed measurement simulation with deterministic processing to the q-c distributed rate distortion problem. Since, the proof of the inner bound of this rate distortion problem requires only the special case of distributed measurement simulation, this is another reason for providing the special case in the previous section.

In Section VI, we consider the non-feedback measurement compression problem for the point-to-point setting. The authors in [2] have discussed this formulation and provided a rate region with a proof of achievability and converse. However, the assumed equations (53) and (54) in proving the direct part (see [2]) do not appear to be true, to the best to our knowledge, but only in an average sense. A stronger version of this theorem is also developed in [12] using a different technique, wherein the authors have extended the Winter’s measurement compression for fixed independent and identically distributed inputs [1] to arbitrary inputs. Since the result is crucial for the distributed simulation problem with stochastic processing, to be described in the next section (Section VII), we formally state the problem and provide an alternative proof of the direct part for completeness (see Theorem 5).

Finally, the above proof of non-feedback simulation in the point-to-point setting provides us with necessary tools for the next task, namely, distributed quantum measurement simulation with stochastic processing. The objective of incorporating the additional processing at the decoder is to reduce the required shared randomness. Our objective in the distributed problem, considered in Section III, was to simulate  $M_A \otimes M_B$ . We achieve this by proving that a pair of POVMs that can faithfully simulate  $M_A$  and  $M_B$  individually, can also faithfully simulate  $M_A \otimes M_B$  (Lemma 4). However, it will be shown that, because of the presence of Eve’s stochastic processing, decoupling the current problem into two symmetric point-to-point problems is not feasible. Therefore, we perform a non-symmetric partitioning while being analytically tractable. Moreover, we provide a single-letter achievable inner bound that is symmetric with respect to Alice and Bob. We conclude the paper with a few remarks in Section IX.

## II. PRELIMINARIES

We here establish all our notations, briefly state few necessary definitions, and also provide Winter’s theorem on measurement compression.

**Notation:** Given any natural number  $M$ , let the finite set  $\{1, 2, \dots, M\}$  be denoted by  $[1, M]$ . 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 all unit trace positive operators acting on  $\mathcal{H}$ . Let  $I$  denote the identity operator. The trace distance between two operators  $A$  and  $B$  is defined as  $\|A - B\|_1 \triangleq \text{Tr} |A - B|$ , where for any operator  $\Lambda$  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 for a bipartite density operator  $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$  is defined as

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

Given any ensemble  $\{p_i, \rho_i\}_{i \in [1, m]}$ , the Holevo information, as in [13], is defined as

$$\chi(\{p_i, \rho_i\}) \triangleq S\left(\sum_i p_i \rho_i\right) - \sum_i p_i S(\rho_i).$$

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

$$\Lambda_x \geq 0, \forall x \in \mathcal{X}, \quad \sum_{x \in \mathcal{X}} \Lambda_x = I.$$

where  $\mathcal{X}$  is a finite set. If instead of the equality above, the inequality  $\sum_x \Lambda_x \leq I$  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  $\Psi_{RA}^\rho$  denote a purification of a density operator  $\rho \in \mathcal{D}(\mathcal{H}_A)$ . Given a POVM  $M \triangleq \{\Lambda_x^A\}_{x \in \mathcal{X}}$  acting on  $\rho$ , the post-measurement state of the reference together with the classical outputs is represented by

$$(\text{id} \otimes M)(\Psi_{RA}^\rho) \triangleq \sum_{x \in \mathcal{X}} |x\rangle\langle x| \otimes \text{Tr}_A\{(I^R \otimes \Lambda_x^A)\Psi_{RA}^\rho\}. \quad (1)$$

Consider two POVMs  $M_A = \{\Lambda_x^A\}_{x \in \mathcal{X}}$  and  $M_B = \{\Lambda_y^B\}_{y \in \mathcal{Y}}$  acting on  $\mathcal{H}_A$  and  $\mathcal{H}_B$ , respectively. Define  $M_A \otimes M_B \triangleq \{\Lambda_x^A \otimes \Lambda_y^B\}_{x \in \mathcal{X}, y \in \mathcal{Y}}$ . 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.

**Definition 1** (Joint Measurements). A POVM  $M_{AB} = \{\Lambda_z^{AB}\}_{z \in \mathcal{Z}}$ , acting on the joint state  $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ , is said to have a separable decomposition with stochastic integration if there exist POVMs  $\bar{M}_A = \{\bar{\Lambda}_u^A\}_{u \in \mathcal{U}}$  and  $\bar{M}_B = \{\bar{\Lambda}_v^B\}_{v \in \mathcal{V}}$  and a stochastic mapping  $P_{Z|U, V} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{Z}$  such that

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

where  $\mathcal{U}, \mathcal{V}$  and  $\mathcal{Z}$  are some finite sets. Further, if the mapping  $P_{Z|U, V}$  is a deterministic function then the POVM is said to have a separable decomposition with deterministic integration.

**Measurement Compression Theorem:** Here, we provide a brief overview of the measurement compression theorem [1]. 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 quantified ‘‘relevant information’’ by measuring the minimum amount of classical information bits needed to ‘‘simulate’’ the repeated action of a measurement  $M$  on a quantum state  $\rho$ . In this context, an agent (Alice) performs an approximating measurement  $\tilde{M}^{(n)}$  on a quantum state  $\rho^{\otimes n}$  and sends a set of classical bits to a receiver (Bob). In addition, Alice and Bob share some amount of common randomness. Bob intends to faithfully recover the outcomes of the original measurement  $M$  without having access to the quantum state based on the bits received from Alice and the common randomness. The objective is to minimize the rate of classical bits under the constraint that the recovered and the original outcomes be statistically indistinguishable. This is formally defined in the following.

**Definition 2** (Faithful simulation [2]). 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  $\tilde{M} \triangleq \{\tilde{\Lambda}_x\}_{x \in \mathcal{X}}$  acting on  $\mathcal{H}_A$  is said to be  $\epsilon$ -faithful to  $M$  with respect to  $\rho$ , for  $\epsilon > 0$ , if the following holds:

$$\sum_{x \in \mathcal{X}} \left\| \sqrt{\rho}(\Lambda_x - \tilde{\Lambda}_x)\sqrt{\rho} \right\|_1 + \text{Tr} \left\{ \left( I - \sum_x \tilde{\Lambda}_x \right) \rho \right\} \leq \epsilon. \quad (2)$$

The above trace norm constraint can be equivalently expressed in terms of a purification of state  $\rho$  using the following lemma.

**Lemma 1.** [2] For any state  $\rho \in \mathcal{D}(\mathcal{H})$  with any purification  $\Psi_{RA}^\rho$ , and any pair of POVMs  $M$  and  $\tilde{M}$  acting on  $\mathcal{H}$ , the following identity holds

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

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

**Theorem 1.** [1] 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  $\tilde{M}^{(n,\mu)}$  for  $\mu \in [1, N]$ , each acting on  $\mathcal{H}_A^{\otimes n}$ , and having at most  $2^{nR}$  outcomes such that  $\tilde{M}^{(n)} \triangleq \frac{1}{N} \sum_\mu \tilde{M}^{(n,\mu)}$  is  $\epsilon$ -faithful to  $M^{\otimes n}$  with respect to  $\rho^{\otimes n}$  if

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

where  $\sigma_{RU} \triangleq (\text{id} \otimes M)(\Psi_{RA}^\rho)$ , and  $\delta(\epsilon) \searrow 0$  as  $\epsilon \searrow 0$ .

**Remark 1.** A strong converse of the above result is also provided in [1].

### III. SIMULATION OF DISTRIBUTED POVMs WITH DETERMINISTIC PROCESSING

Now, we develop an extension of Winter's measurement compression [1] to quantum measurements performed in a distributed fashion with deterministic processing. Consider a bipartite composite quantum system  $(A, B)$  represented by Hilbert Space  $\mathcal{H}_A \otimes \mathcal{H}_B$ . Let  $\rho_{AB}$  be a density operator 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 these two measurements, one applied to each sub-system. The three parties share some amount of common randomness. Alice and Bob perform a measurement  $\tilde{M}_A^{(n)}$  and  $\tilde{M}_B^{(n)}$  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  $\tilde{M}_A^{(n)}$  and  $\tilde{M}_B^{(n)}$  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  $M_A^{\otimes n} \otimes M_B^{\otimes n}$ . The problem is formally defined in the following.

**Definition 3.** 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 sub-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  $|\mathcal{L}_1| \leq \Theta_1$ .
- 2) a collections of Bob's sub-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  $|\mathcal{L}_2| \leq \Theta_2$ .

3) a collection of Eve's decoding maps  $f^{(\mu)} : \mathcal{L}_1 \times \mathcal{L}_2 \rightarrow \mathcal{Z}^n$  for  $\mu \in [1, N]$ .

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

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

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

In the above definition,  $(\Theta_1, \Theta_2)$  determines the amount of classical bits communicated from Alice and Bob to Eve. The amount of common randomness is characterized by  $N$ , and  $\mu$  can be viewed as the common randomness bits distributed among the parties. The mapping  $f^{(\mu)}$  represents the action of Eve on the received classical bits.

**Definition 4.** Given a POVM  $M_{AB} \triangleq \{\Lambda_z^{AB}\}_{z \in \mathcal{Z}}$  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 sub-POVM  $\tilde{M}_{AB}$  is  $\epsilon$ -faithful to  $M_{AB}^{\otimes n}$  with respect to  $\rho_{AB}^{\otimes n}$  (see Definition 2), and

$$\frac{1}{n} \log_2 N \leq C + \epsilon, \quad \frac{1}{n} \log_2 \Theta_i \leq R_i + \epsilon, \quad i = 1, 2.$$

The set of all achievable triples  $(R_1, R_2, C)$  is called the achievable rate region.

The following theorem provides an inner bound to the achievable rate region, which is proved in Section IV,

**Theorem 2.** Given a density operator  $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$  and a POVM  $M_{AB} \triangleq \{\Lambda_z^{AB}\}_{z \in \mathcal{Z}}$  acting on  $\mathcal{H}_A \otimes \mathcal{H}_B$  having a separable decomposition with deterministic integration (as in Definition 1), a triple  $(R_1, R_2, C)$  is achievable if the following inequalities are satisfied:

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

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

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

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

$$R_2 + C \geq S(V|U)_{\sigma_3}, \quad (5e)$$

$$R_1 + R_2 + C \geq S(U, V)_{\sigma_3}, \quad (5f)$$

for some decomposition with POVMs  $\bar{M}_A = \{\bar{\Lambda}_u^A\}_{u \in \mathcal{U}}$  and  $\bar{M}_B = \{\bar{\Lambda}_v^B\}_{v \in \mathcal{V}}$  and a function  $g : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{Z}$ , where the information quantities are computed for the auxiliary states  $\sigma_1^{RUB} \triangleq (\text{id}_R \otimes \bar{M}_A \otimes \text{id}_B)(\Psi_{RAB}^{\rho_{AB}})$ ,  $\sigma_2^{RAV} \triangleq (\text{id}_R \otimes \text{id}_A \otimes \bar{M}_B)(\Psi_{RAB}^{\rho_{AB}})$ , and  $\sigma_3^{RUV} \triangleq (\text{id}_R \otimes \bar{M}_A \otimes \bar{M}_B)(\Psi_{RAB}^{\rho_{AB}})$ , with  $\Psi_{RAB}^{\rho_{AB}}$  being a purification of  $\rho_{AB}$ , and  $\mathcal{U}$  and  $\mathcal{V}$  are some finite sets.

**Remark 2.** An alternative characterization of the above rate region can be obtained in terms of Holevo information. For this, we use the canonical ensembles  $\{\lambda_u^A, \hat{\rho}_u^A\}$ ,  $\{\lambda_v^B, \hat{\rho}_v^B\}$  and  $\{\lambda_{uv}^{AB}, \hat{\rho}_{uv}^{AB}\}$  defined as

$$\begin{aligned} \lambda_u^A &\triangleq \text{Tr}\{\bar{\Lambda}_u^A \rho_A\}, & \lambda_v^B &\triangleq \text{Tr}\{\bar{\Lambda}_v^B \rho_B\}, & \lambda_{uv}^{AB} &\triangleq \text{Tr}\{(\bar{\Lambda}_u^A \otimes \bar{\Lambda}_v^B) \rho_{AB}\}, \\ \hat{\rho}_u^A &\triangleq \frac{1}{\lambda_u^A} \sqrt{\rho_A} \bar{\Lambda}_u^A \sqrt{\rho_A}, & \hat{\rho}_v^B &\triangleq \frac{1}{\lambda_v^B} \sqrt{\rho_B} \bar{\Lambda}_v^B \sqrt{\rho_B}, & \hat{\rho}_{uv}^{AB} &\triangleq \frac{1}{\lambda_{uv}^{AB}} \sqrt{\rho_{AB}} (\bar{\Lambda}_u^A \otimes \bar{\Lambda}_v^B) \sqrt{\rho_{AB}}. \end{aligned} \quad (6)$$

Using this, we get

$$I(U; RB)_{\sigma_1} = \chi(\{\lambda_u^A, \hat{\rho}_u^A\}) \quad \text{and} \quad I(V; RA)_{\sigma_2} = \chi(\{\lambda_v^B, \hat{\rho}_v^B\}).$$

Also,  $I(U; V)_{\sigma_3}$ , and  $S(U, V)_{\sigma_3}$  are equal to the classical mutual information and joint entropy with respect to the joint distribution  $\{\lambda_{uv}^{AB}\}$ , respectively.

Before providing a proof in the next section, we briefly discuss two corner points of the rate region with respect to the common randomness available. Firstly, 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  $C \geq S(U|RB)_{\sigma_1} + S(V|RA)_{\sigma_2}$ . Next, let us consider the regime where  $C = 0$ . This implies  $R_1 + R_2 \geq S(U, V)_{\sigma_3}$ . This regime corresponds to the quantum measurement  $M_A \otimes M_B$  followed by classical Slepian-Wolf compression [14]. Fig. 2 demonstrates the achievable rate region in these cases.

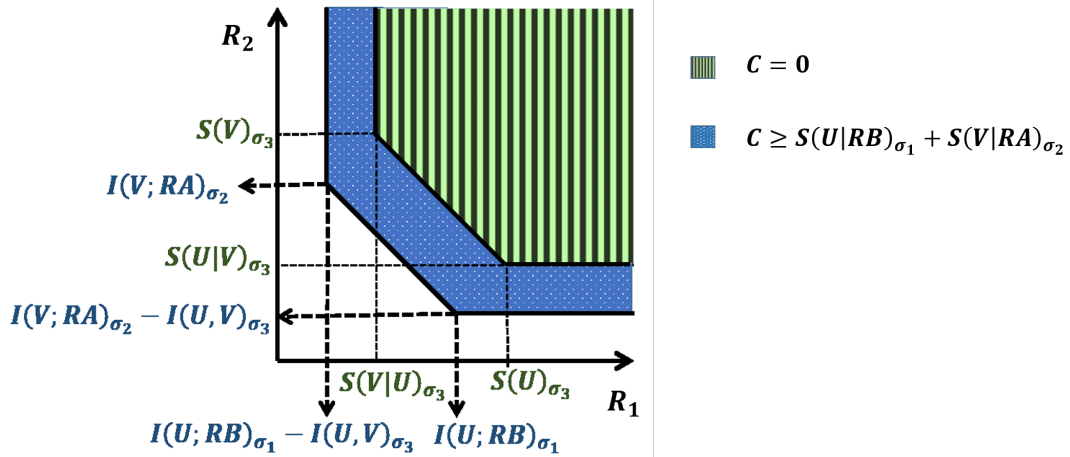


Fig. 2. The inner bound to the achievable rate region given in Theorem 2 at two planes: 1) with no common randomness, i.e.,  $C = 0$  (green color), and 2) with at least  $S(U|RB)_{\sigma_1} + S(V|RA)_{\sigma_2}$  amount of common randomness (blue color). As a result, the latter region contains the former.

We encounter two challenges in developing the single-letter inner bound to the achievable rate region as stated in Theorem 2: 1) The direct use of single-POVM compression theorem, proved using random coding arguments as in [1], for each individual POVMs,  $M_A$  and  $M_B$ , does not necessarily ensure a “distributed” faithful simulation for the overall measurement,  $M_A \otimes M_B$ . This issue is unique to the quantum settings. One of the contributions of this work is to prove this when the two sources  $A$  and  $B$  are not necessarily independent, i.e.,  $\rho_{AB} \neq \rho_A \otimes \rho_B$  (see Lemma 4).

2) The classical outputs of the approximating POVMs (operating on  $n$  copies of the source) are not *independently and identically distributed* (IID) sequences - rather they are codewords generated from random coding. The Slepian-Wolf scheme [14] (also referred to as *binning* in the literature) is developed for distributed compression of IID source sequences. Applicability of such an approach to the problem requires that the classical outputs produced from the two approximating POVMs are jointly typical with high probability. This issue also arises in classical distributed source coding problem which was addressed by Wyner-Ahlsvede-Korner by developing Markov Lemma and Mutual Packing Lemma (Lemma 12.1 and 12.2 in [15]). Building upon these ideas, we develop quantum-classical counterparts of these lemmas for the multi-user quantum measurement simulation problem (see the discussion in Section IV-B and Proposition 2). Let us consider an example to illustrate the above inner bound.

**Example 1.** Suppose the composite state  $\rho_{AB}$  is described using one of the Bell states on  $\mathcal{H}_A \otimes \mathcal{H}_B$  as

$$\rho^{AB} = \frac{1}{2}(|00\rangle_{AB} + |11\rangle_{AB})(\langle 00|_{AB} + \langle 11|_{AB}).$$

Since  $\pi^A = \text{Tr}_B \rho^{AB}$  and  $\pi^B = \text{Tr}_A \rho^{AB}$ , Alice and Bob would perceive each of their particles in maximally mixed states  $\pi^A = \frac{I^A}{2}$  and  $\pi^B = \frac{I^B}{2}$ , respectively. Upon receiving the quantum state, the two parties wish to independently measure their states, using identical POVMs  $M_A$  and  $M_B$ , given by  $\left\{ \frac{1}{2}|0\times 0\rangle, \frac{1}{2}|1\times 1\rangle, \frac{1}{2}|+\times +\rangle, \frac{1}{2}|-\times -\rangle \right\}$ . Alice



and Bob together with Eve are trying to simulate the action of  $M_A \otimes M_B$  using the classical communication and common randomness as the resources available to them (as described earlier).

We compute the constraints given in Theorem 2. Considering the first constraint from (5a), we evaluate  $\sigma_1^{UB}$  as

$$\sigma_1^{UB} = \frac{1}{4}(|0\rangle\langle 0|_U \otimes |0\rangle\langle 0|_B + |1\rangle\langle 1|_U \otimes |1\rangle\langle 1|_B + |2\rangle\langle 2|_U \otimes |+\rangle\langle +|_B + |3\rangle\langle 3|_U \otimes |-\rangle\langle -|_B),$$

where the vectors  $\{|0\rangle_U, |1\rangle_U, |2\rangle_U, |3\rangle_U\}$  denote a set of orthogonal states on the space  $\mathcal{H}_U$ . Based on this state, we get

$$S(\sigma_1^{RUB}) = S(\sigma_1^{UB}) = 2, \quad S(\sigma_1^{RB}) = S(\sigma_1^B) = 1 \quad \text{and} \quad S(\sigma_1^U) = 2.$$

This gives  $I(U; RB)_{\sigma_1}$  to be equal to 1 bit. Similarly, from the symmetry of the example, we also get  $I(V; RA)_{\sigma_2}$  to be equal to 1 bit. Similarly, we can evaluate  $\sigma_3^{UV}$  as

$$\sigma_3^{UV} = \left( \frac{1}{8} \sum_{i=0}^3 |i\rangle\langle i|_U \otimes |i\rangle\langle i|_V + \frac{1}{16} \sum_{i=0}^3 \sum_{j=i+2}^{i+4} |i\rangle\langle i|_U \otimes |\text{mod}(j, 4)\rangle\langle \text{mod}(j, 4)|_V \right),$$

which gives

$$S(U, V)_{\sigma_3} = 3.5 \quad \text{and} \quad I(U; V)_{\sigma_3} = 0.5.$$

Therefore, we can write the constraints given in Theorem 2 as

$$R_1 \geq 0.5, \quad R_2 \geq 0.5, \quad R_1 + R_2 \geq 1.5, \quad R_1 + C \geq 1.5, \quad R_2 + C \geq 1.5 \quad \text{and} \quad R_1 + R_2 + C \geq 3.5.$$

Consider the case when  $C \geq 2$  is available. By approximating  $M_A$  and  $M_B$  individually, we receive a gain of 1 bit, decreasing the rate from  $S(U)_{\sigma_1} = 2$  bits to  $I(U; RB)_{\sigma_1} = 1$  bit and similarly from  $S(V)_{\sigma_2} = 2$  bits to  $I(V; RA)_{\sigma_2} = 1$  bit. Binning of these approximating POVMs (as discussed in Section (IV-B)), gives an additional gain of half a bit, which is characterized by  $I(U; V)_{\sigma_3} = 0.5$ , thus giving us the achievable sum-rate of 1.5 bits.

In the next section we provide the proof for this inner bound.

#### IV. PROOF OF THEOREM 2

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

$$\Lambda_z^{AB} = \sum_{u,v} \mathbb{1}_{\{g(u,v)=z\}} \bar{\Lambda}_u^A \otimes \bar{\Lambda}_v^B, \quad \forall z \in \mathcal{Z}, \quad (7)$$

for some POVMs  $\bar{M}_A$  and  $\bar{M}_B$  with operators denoted by  $\{\bar{\Lambda}_u^A\}_{u \in \mathcal{U}}$  and  $\{\bar{\Lambda}_v^B\}_{v \in \mathcal{V}}$ , respectively, and for some function  $g : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{Z}$  where  $\mathcal{U}, \mathcal{V}$  and  $\mathcal{Z}$  are three finite sets. The proof follows by constructing a protocol for faithful simulation of  $\bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n}$ . We start by generating the canonical ensembles corresponding to  $\bar{M}_A$  and  $\bar{M}_B$ , as given in (6). With this notation, corresponding to each of the probability distributions, we can associate a  $\delta$ -typical set. Let us denote  $\mathcal{T}_\delta^{(n)}(U)$ ,  $\mathcal{T}_\delta^{(n)}(V)$  and  $\mathcal{T}_\delta^{(n)}(UV)$  as the  $\delta$ -typical sets defined for  $\{\lambda_u^A\}$ ,  $\{\lambda_v^B\}$  and  $\{\lambda_{uv}^{AB}\}$ , respectively.

Let  $\Pi_{\rho_A}$  and  $\Pi_{\rho_B}$  denote the  $\delta$ -typical projectors (as in [13]) for marginal density operators  $\rho_A$  and  $\rho_B$ , 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 [13]) 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

$$\Lambda_{u^n}^{A'} \triangleq \Pi_{\rho_A} \Pi_{u^n}^A \hat{\rho}_{u^n}^A \Pi_{u^n}^A \Pi_{\rho_A}, \quad \Lambda_{v^n}^{B'} \triangleq \Pi_{\rho_B} \Pi_{v^n}^B \hat{\rho}_{v^n}^B \Pi_{v^n}^B \Pi_{\rho_B}, \quad (8)$$

where  $\hat{\rho}_{u^n}^A \triangleq \otimes_i \hat{\rho}_{u_i}^A$  and  $\hat{\rho}_{v^n}^B \triangleq \otimes_i \hat{\rho}_{v_i}^B$ <sup>1</sup>.

<sup>1</sup>Note that  $\Lambda_{u^n}^{A'}$  and  $\Lambda_{v^n}^{B'}$  are not tensor products operators.

With the notation above, define  $\sigma^{A'}$  and  $\sigma^{B'}$  as

$$\sigma^{A'} \triangleq \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \frac{\lambda_{u^n}^A}{(1-\varepsilon)} \Lambda_{u^n}^{A'} \quad \text{and} \quad \sigma^{B'} \triangleq \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \frac{\lambda_{v^n}^B}{(1-\varepsilon')} \Lambda_{v^n}^{B'}, \quad (9)$$

where  $\varepsilon = \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A$  and  $\varepsilon' = \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \lambda_{v^n}^B$ . Note that  $\sigma^{A'}$  and  $\sigma^{B'}$  defined above are expectations with respect to the pruned distribution [16]. 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  $\varepsilon 2^{-n(S(\rho_A)+\delta_1)}$  and  $\varepsilon' 2^{-n(S(\rho_B)+\delta_2)}$ , for some  $\delta_1, \delta_2 \geq 0$ . Lastly, define<sup>2</sup>.

$$\Lambda_{u^n}^A \triangleq \hat{\Pi}^A \Lambda_{u^n}^{A'} \hat{\Pi}^A, \quad \text{and} \quad \Lambda_{v^n}^B \triangleq \hat{\Pi}^B \Lambda_{v^n}^{B'} \hat{\Pi}^B. \quad (10)$$

### 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  $\sigma_3$  is defined as

$$\sigma_3^{RUV} \triangleq (\text{id}^R \otimes \bar{M}_A \otimes \bar{M}_B)(\Psi_{RAB}^{\rho_{AB}}),$$

with  $\Psi_{RAB}^{\rho_{AB}}$  being any purification of  $\rho_{AB}$ <sup>3</sup>. Let  $\mu_1 \in [1, N_1]$  denote the common randomness shared between the first encoder and the decoder, and let  $\mu_2 \in [1, N_2]$  denote the common randomness shared between the second encoder and the decoder, with  $\log(N_1) + \log(N_2) \leq \log(N)$ . For each  $\mu_1 \in [1, N_1]$  and  $\mu_2 \in [1, N_2]$ , randomly and independently select  $2^{n\tilde{R}_1}$  and  $2^{n\tilde{R}_2}$  sequences  $(U^{n,(\mu_1)}(l), V^{n,(\mu_2)}(k))$  according to the pruned distributions, i.e.,

$$\mathbb{P}\left((U^{n,(\mu_1)}(l), V^{n,(\mu_2)}(k)) = (u^n, v^n)\right) = \begin{cases} \frac{\lambda_{u^n}^A}{(1-\varepsilon)} \frac{\lambda_{v^n}^B}{(1-\varepsilon')} & \text{for } u^n \in \mathcal{T}_\delta^{(n)}(U), v^n \in \mathcal{T}_\delta^{(n)}(V) \\ 0 & \text{otherwise} \end{cases}. \quad (11)$$

Construct operators

$$A_{u^n}^{(\mu_1)} \triangleq \gamma_{u^n}^{(\mu_1)} \left( \sqrt{\rho_A}^{-1} \Lambda_{u^n}^A \sqrt{\rho_A}^{-1} \right) \quad \text{and} \quad B_{v^n}^{(\mu_2)} \triangleq \zeta_{v^n}^{(\mu_2)} \left( \sqrt{\rho_B}^{-1} \Lambda_{v^n}^B \sqrt{\rho_B}^{-1} \right), \quad (12)$$

where

$$\gamma_{u^n}^{(\mu_1)} \triangleq \frac{1-\varepsilon}{1+\eta} 2^{-n\tilde{R}_1} |\{l : U^{n,(\mu_1)}(l) = u^n\}| \quad \text{and} \quad \zeta_{v^n}^{(\mu_2)} \triangleq \frac{1-\varepsilon'}{1+\eta} 2^{-n\tilde{R}_2} |\{k : V^{n,(\mu_2)}(k) = v^n\}|, \quad (13)$$

where  $\eta \in (0, 1)$  is a parameter to be determined. Then, for each  $\mu_1 \in [1, N_1]$  and  $\mu_2 \in [1, N_2]$ , construct  $M_1^{(n, \mu_1)}$  and  $M_2^{(n, \mu_2)}$  as in the following

$$M_1^{(n, \mu_1)} \triangleq \{A_{u^n}^{(\mu_1)} : u^n \in \mathcal{T}_\delta^{(n)}(U)\}, \quad M_2^{(n, \mu_2)} \triangleq \{B_{v^n}^{(\mu_2)} : v^n \in \mathcal{T}_\delta^{(n)}(V)\}. \quad (14)$$

We show in the later part of the proof (Lemma 2) that  $M_1^{(n, \mu_1)}$  and  $M_2^{(n, \mu_2)}$  form sub-POVMs, with high probability, for all  $\mu \in [1, N_1]$  and  $\mu_2 \in [1, N_2]$ , respectively. These collections  $M_1^{(n, \mu_1)}$  and  $M_2^{(n, \mu_2)}$  are completed using the operators  $I - \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\mu_1)}$  and  $I - \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\mu_2)}$ , and these operators are associated with sequences  $u_0^n$  and  $v_0^n$ , which are chosen arbitrarily from  $\mathcal{U}^n \setminus \mathcal{T}_\delta^{(n)}(U)$  and  $\mathcal{V}^n \setminus \mathcal{T}_\delta^{(n)}(V)$ , respectively.

<sup>2</sup>Note that  $\Lambda_{u^n}^A$  and  $\Lambda_{v^n}^B$  are not tensor products operators.

<sup>3</sup>The information theoretic quantities calculated with respect to  $\sigma_3^{RUV}$  remain independent of the purification used in its definition.

## 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  $(R_1, R_2)$  and choose a  $(\mu_1, \mu_2)$  pair. For each sequence  $u^n \in \mathcal{T}_\delta^{(n)}(U)$  assign an index from  $[1, 2^{nR_1}]$  randomly and uniformly, such that the assignments for different sequences are done independently. Perform a similar random and independent assignment for all  $v^n \in \mathcal{T}_\delta^{(n)}(V)$  with indices chosen from  $[1, 2^{nR_2}]$ . Repeat this assignment for every  $\mu_1 \in [1, 2^{nC_1}]$  and  $\mu_2 \in [1, 2^{nC_2}]$ . For each  $i \in [1, 2^{nR_1}]$  and  $j \in [1, 2^{nR_2}]$ , let  $\mathcal{B}_1^{(\mu_1)}(i)$  and  $\mathcal{B}_2^{(\mu_2)}(j)$  denote the  $i^{\text{th}}$  and the  $j^{\text{th}}$  bins, respectively. More precisely,  $\mathcal{B}_1^{(\mu_1)}(i)$  is the set of all  $u^n$  sequences with assigned index equal to  $i$ , and similar is  $\mathcal{B}_2^{(\mu_2)}(j)$ . Define the following operators:

$$\Gamma_i^{A,(\mu_1)} \triangleq \sum_{u^n \in \mathcal{B}_1^{(\mu_1)}(i)} A_{u^n}^{(\mu_1)}, \quad \text{and} \quad \Gamma_j^{B,(\mu_2)} \triangleq \sum_{v^n \in \mathcal{B}_2^{(\mu_2)}(j)} B_{v^n}^{(\mu_2)},$$

for all  $i \in [1, 2^{nR_1}]$  and  $j \in [1, 2^{nR_2}]$ . Using these operators, we form the following collection:

$$M_A^{(n,\mu_1)} \triangleq \{\Gamma_i^{A,(\mu_1)}\}_{i \in [1, 2^{nR_1}]}, \quad M_B^{(n,\mu_2)} \triangleq \{\Gamma_j^{B,(\mu_2)}\}_{j \in [1, 2^{nR_2}]}. \quad (15)$$

Note that if  $M_1^{(n,\mu_1)}$  and  $M_2^{(n,\mu_2)}$  are sub-POVMs, then so are  $M_A^{(n,\mu_1)}$  and  $M_B^{(n,\mu_2)}$ . This is due to the relations

$$\sum_i \Gamma_i^{A,(\mu_1)} = \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\mu_1)}, \quad \text{and} \quad \sum_j \Gamma_j^{B,(\mu_2)} = \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\mu_2)}.$$

To make  $M_A^{(n,\mu_1)}$  and  $M_B^{(n,\mu_2)}$  complete, we define  $\Gamma_0^{A,(\mu_1)}$  and  $\Gamma_0^{B,(\mu_2)}$  as  $\Gamma_0^{A,(\mu_1)} = I - \sum_i \Gamma_i^{A,(\mu_1)}$  and  $\Gamma_0^{B,(\mu_2)} = I - \sum_j \Gamma_j^{B,(\mu_2)}$ , respectively<sup>4</sup>. Now, we intend to use the completions  $[M_A^{(n,\mu_1)}]$  and  $[M_B^{(n,\mu_2)}]$  as the POVMs for each encoder. Also, 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  $A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)}$  are used to simulate  $\bar{M}_A \otimes \bar{M}_B$ . 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 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  $F^{(\mu)}$ , for  $\mu = (\mu_1, \mu_2)$ , acting on the outputs of  $[M_A^{(n,\mu_1)}] \otimes [M_B^{(n,\mu_2)}]$  as follows. Let  $\mathcal{C}^{(\mu)}$  denote the codebook containing all pairs of codewords  $(U^{n,(\mu_1)}(l), V^{n,(\mu_2)}(k))$ . On observing  $\mu$  and the classical indices  $(i, j) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$  communicated by the encoder, the decoder first deduces  $(\mu_1, \mu_2)$  from  $\mu$  and then populates,

$$D_{i,j}^{(\mu_1,\mu_2)} \triangleq \left\{ (u^n, v^n) \in \mathcal{C}^{(\mu)} : (u^n, v^n) \in \mathcal{T}_\delta^{(n)}(UV) \text{ and } (u^n, v^n) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j) \right\}. \quad (16)$$

For every  $\mu \in [1 : N]$ ,  $i \in [1 : 2^{nR_1}]$  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  $D_{i,j}^{(\mu_1,\mu_2)}$ ; otherwise  $F^{(\mu)}(i, j) = (u_0^n, v_0^n)$ . Further,  $F^{(\mu)}(i, j) = (u_0^n, v_0^n)$  for  $i = 0$  or  $j = 0$ . Finally, the decoder produces  $z^n \in \mathcal{Z}^n$  according to the map  $g^n(F^{(\mu)}(i, j))$ . With this mapping, we form the following collection of operators, denoted by  $\tilde{M}_{AB}^{(n)}$ ,

$$\tilde{M}_{u^n, v^n}^{AB} \triangleq \frac{1}{N_1 N_2} \sum_{\mu_1=1}^{N_1} \sum_{\mu_2=1}^{N_2} \sum_{(i,j): F^{(\mu)}(i,j)=(u^n, v^n)} \Gamma_i^{A,(\mu_1)} \otimes \Gamma_j^{B,(\mu_2)}, \quad \forall (u^n, v^n) \in \mathcal{U}^n \times \mathcal{V}^n.$$

<sup>4</sup>Note that  $\Gamma_0^{A,(\mu_1)} = I - \sum_i \Gamma_i^{A,(\mu_1)} = I - \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\mu_1)}$  and  $\Gamma_0^{B,(\mu_2)} = I - \sum_j \Gamma_j^{B,(\mu_2)} = I - \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\mu_2)}$

Note that for  $\tilde{\Lambda}_{u^n, v^n}^{AB} = 0$  for  $(u^n, v^n) \notin (\mathcal{T}_\delta^{(n)}(U) \times \mathcal{T}_\delta^{(n)}(V)) \cup \{(u_0^n, v_0^n)\}$ . We show that  $\tilde{M}_{AB}^{(n)}$  is a POVM that is  $\epsilon$ -faithful to the intermediate POVM  $\bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n}$ , with respect to  $\rho_{AB}^{\otimes n}$ . For faithful simulation of the original POVM  $M_{AB}$ , we apply the deterministic mapping  $g^n(u^n, v^n)$  to the classical outputs of  $\tilde{M}_{AB}^{(n)}$ . More precisely, we construct the POVM  $\hat{M}_{AB}^{(n)}$  with the following operators:

$$\hat{\Lambda}_{z^n}^{AB} = \sum_{u^n, v^n} \tilde{\Lambda}_{u^n, v^n}^{AB} \mathbb{1}_{\{g^n(u^n, v^n) = z^n\}}, \quad \forall z^n \in \mathcal{Z}^n.$$

#### D. Analysis of POVM and Trace Distance

In what follows, we show that  $\hat{M}_{AB}^{(n)}$  is a POVM, and is  $\epsilon$ -faithful with respect to  $\rho_{AB}$  (according to Definition 2) to  $M_{AB}$ , where  $\epsilon > 0$  can be made arbitrarily small for sufficiently large  $n$ . More precisely, we show that, with probability sufficiently close to 1,

$$\sum_{z^n} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \Lambda_{z^n}^{AB} - \hat{\Lambda}_{z^n}^{AB} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \leq \epsilon. \quad (17)$$

According to the decomposition of  $\Lambda_z$ , given in (7), the above inequality is equivalent to

$$\sum_{z^n} \left\| \sum_{u^n, v^n} \mathbb{1}_{\{g(u^n, v^n) = z^n\}} \left( \sqrt{\rho_{AB}^{\otimes n}} (\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \tilde{\Lambda}_{u^n, v^n}^{AB}) \sqrt{\rho_{AB}^{\otimes n}} \right) \right\|_1 \leq \epsilon.$$

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

$$\sum_{z^n} \sum_{u^n, v^n} \mathbb{1}_{\{g(u^n, v^n) = z^n\}} \left\| \sqrt{\rho_{AB}^{\otimes n}} (\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \tilde{\Lambda}_{u^n, v^n}^{AB}) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 = \sum_{u^n, v^n} \left\| \sqrt{\rho_{AB}^{\otimes n}} (\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \tilde{\Lambda}_{u^n, v^n}^{AB}) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1.$$

Hence, it is sufficient to show that the above quantity is no greater than  $\epsilon$ , with probability sufficiently close to 1. This is equivalent to showing that  $\tilde{M}_{AB}^{(n)}$  is  $\epsilon$ -faithful to  $\bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n}$  with respect to  $\rho_{AB}^{\otimes n}$ . Alternatively, using Lemma 1, we prove the following inequality

$$G \triangleq \left\| (\text{id} \otimes \bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n})(\Psi_{R^n A^n B^n}^\rho) - (\text{id} \otimes \tilde{M}_{AB}^{(n)})(\Psi_{R^n A^n B^n}^\rho) \right\|_1 \leq \epsilon. \quad (18)$$

We characterize the conditions on  $(n, N, R_1, R_2)$  under which the inequality given in (18) holds, using the following steps.

##### Step 1: $M_1^{(n, \mu_1)}$ and $M_2^{(n, \mu_2)}$ are sub-POVMs and individually approximating

As a first step, one can show that with probability sufficiently close to one,  $M_1^{(n, \mu_1)}$  and  $M_2^{(n, \mu_2)}$  form sub-POVMs for all  $\mu_1 \in [1, N_1]$  and  $\mu_2 \in [1, N_2]$ , and also individually approximate the corresponding tensor product POVMs. More precisely the following Lemma holds.

**Lemma 2.** *For any two positive integers  $N_1$  and  $N_2$ , and  $\epsilon, \epsilon', \eta \in (0, 1)$ , as in (13), and any  $\zeta \in (0, 1)$ , there exists  $n(\epsilon, \epsilon', \eta, \zeta)$  such that for all  $n \geq n(\epsilon, \epsilon', \eta, \zeta)$ , the collection of operators  $M_1^{(n, \mu_1)}$  and  $M_2^{(n, \mu_2)}$  form sub-POVMs for all  $\mu_1 \in [1, N_1]$  and  $\mu_2 \in [1, N_2]$  with probability at least  $(1 - \zeta)$ , provided that*

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

where  $\sigma_1, \sigma_2$  are defined as in the statement of the theorem. In addition, if

$$\frac{1}{n} \log_2 N_1 + \tilde{R}_1 > S(U)_{\sigma_1}, \quad \frac{1}{n} \log_2 N_2 + \tilde{R}_2 > S(V)_{\sigma_2}, \quad (19)$$

then with probability at least  $(1 - \zeta)$  the collection of average operators  $M_i^{(n)} \triangleq \frac{1}{N_i} \sum_{\mu_i} [M_i^{(n, \mu_i)}]$ ,  $i = 1, 2$  are  $\epsilon$ -faithful to  $M_A^{\otimes n}$  with respect to  $\rho_A^{\otimes n}$  and  $M_B^{\otimes n}$  with respect to  $\rho_B^{\otimes n}$ , respectively.

*Proof.* The proof uses a similar argument as in that of Theorem 2 in [1]. Hence it is omitted.  $\square$

As a result of the lemma,  $\tilde{M}_{AB}^{(n)}$  and  $\hat{M}_{AB}^{(n)}$  are valid POVMs with high probability.

### Step 2: Isolating the effect of un-binned approximating measurements

In this step, we separate out the effect of un-binned approximating measurements from  $G$  in (18). This is done by adding and subtracting an appropriate term within the trace norm and applying triangle inequality, which bounds  $G$  as  $G \leq S_1 + S_2$ , where

$$\begin{aligned} S_1 &\triangleq \left\| (\text{id} \otimes \bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n})(\Psi_{R^n A^n B^n}^\rho) - \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} (\text{id} \otimes [M_1^{(n, \mu_1)}] \otimes [M_2^{(n, \mu_2)}])(\Psi_{R^n A^n B^n}^\rho) \right\|_1, \\ S_2 &\triangleq \left\| \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} (\text{id} \otimes [M_1^{(n, \mu_1)}] \otimes [M_2^{(n, \mu_2)}])(\Psi_{R^n A^n B^n}^\rho) - (\text{id} \otimes \tilde{M}_{AB}^{(n)})(\Psi_{R^n A^n B^n}^\rho) \right\|_1, \end{aligned} \quad (20)$$

where the  $S_1$  captures the effect of using approximating POVMs  $M_1^{(n, \mu_1)}$  and  $M_2^{(n, \mu_2)}$  instead of the actual POVMs  $\bar{M}_A^{\otimes n}$  and  $\bar{M}_B^{\otimes n}$ , while  $S_2$  captures the error introduced by binning these approximating POVMs. Before we proceed further, we provide the following lemma which will be useful in the rest of the paper.

**Lemma 3.** *Given a density operator  $\rho_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$ , a sub-POVM  $M_Y \triangleq \{\Lambda_y^B : y \in \mathcal{Y}\}$  acting on  $\mathcal{H}_B$ , for some set  $\mathcal{Y}$ , and any Hermitian operator  $\Gamma^A$  acting on  $\mathcal{H}_A$ , we have*

$$\sum_{y \in \mathcal{Y}} \|\sqrt{\rho_{AB}} (\Gamma^A \otimes \Lambda_y^B) \sqrt{\rho_{AB}}\|_1 \leq \|\sqrt{\rho_A} \Gamma^A \sqrt{\rho_A}\|_1, \quad (21)$$

with equality if  $\sum_{y \in \mathcal{Y}} \Lambda_y^B = I$ , where  $\rho_A = \text{Tr}_B\{\rho_{AB}\}$ .

*Proof.* The proof is provided in Appendix A-A. □

Next, we show  $S_1$  is sufficiently small using the following Mutual Covering Lemma.

**Lemma 4 (Mutual Covering Lemma).** *Suppose the sub-POVM  $\hat{M}_X$  is  $\epsilon$ -faithful to  $M_X$  with respect to  $\rho_X$ , and the sub-POVM  $\hat{M}_Y$  is  $\epsilon$ -faithful to  $M_Y$  with respect to  $\rho_Y$ , where  $\rho_X = \text{Tr}_Y\{\rho_{XY}\}$  and  $\rho_Y = \text{Tr}_X\{\rho_{XY}\}$ . Then the sub-POVM  $\hat{M}_X \otimes \hat{M}_Y$  is  $2\epsilon$ -faithful to the POVM  $M_X \otimes M_Y$  with respect to  $\rho_{XY}$ .*

*Proof.* The proof is provided in the Appendix A-B. □

Using Lemma 4 with  $\rho_{XY} = \rho_{AB}^{\otimes n}$ ,  $\hat{M}_X = \frac{1}{N_1} \sum_{\mu_1} [M_1^{(n, \mu_1)}]$ ,  $\hat{M}_Y = \frac{1}{N_2} \sum_{\mu_2} [M_2^{(n, \mu_2)}]$ ,  $M_X = \bar{M}_A^{\otimes n}$  and  $M_Y = \bar{M}_B^{\otimes n}$ , and Lemma 2, with probability at least  $(1 - \zeta)$ , we have  $S_1 \leq 2\epsilon$ .

### Step 3: Analyzing the effect of Binning

In this step, we provide an upper bound on  $S_2$ . For  $(u^n, v^n) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j)$ , 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  $\Phi_{u^n, v^n} \triangleq |u^n, v^n\rangle\langle u^n, v^n|$ . With this notation, we simplify  $S_2$  using the following proposition.

**Proposition 1.**  $S_2$  can be simplified as

$$S_2 = \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \|\Phi_{u^n, v^n} - \Phi_{e^{(\mu)}(u^n, v^n)}\|_1 \gamma_{u^n}^{(\mu_1)} \zeta_{v^n}^{(\mu_2)} \Omega_{u^n, v^n},$$

where  $\Omega_{u^n, v^n}$  is defined as

$$\Omega_{u^n, v^n} \triangleq \text{Tr} \left\{ \sqrt{\rho_A^{\otimes n} \otimes \rho_B^{\otimes n}}^{-1} (\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B) \sqrt{\rho_A^{\otimes n} \otimes \rho_B^{\otimes n}}^{-1} \rho_{AB}^{\otimes n} \right\}.$$

*Proof.* The proof is provided in Appendix B-A. □

In the next proposition we provide a bound on  $S_2$ .

**Proposition 2 (Mutual Packing).** *There exist functions  $\epsilon_{S_2}(\delta)$  and  $\delta_{S_2}(\delta)$ , such that for all sufficiently small  $\delta$  and sufficiently large  $n$ , we have  $\mathbb{P}(S_2 > \epsilon_{S_2}(\delta)) \leq \delta_{S_2}(\delta)$ , if  $\tilde{R}_1 + \tilde{R}_2 - R_1 - R_2 < I(U; V)_{\sigma_2}$ , where  $\sigma_3$  is the auxiliary state defined in the theorem and  $\epsilon_{S_2}, \delta_{S_2} \searrow 0$  as  $\delta \searrow 0$ .*

*Proof.* The proof is provided in Appendix B-B. □

Using this, and from Step 2, with probability sufficiently close to one, we have

$$G \leq S_1 + S_2 \leq 2\epsilon + \epsilon_{S_2}.$$

### E. Rate Constraints

To sum-up, we showed that the trace distance inequality in (18) holds for sufficiently large  $n$  and with probability sufficiently close to 1, if the following bounds hold:

$$\tilde{R}_1 \geq I(U; RB)_{\sigma_1} \tag{22a}$$

$$\tilde{R}_2 \geq I(V; RA)_{\sigma_2} \tag{22b}$$

$$C_1 + \tilde{R}_1 \geq S(U)_{\sigma_3}, \quad C_2 + \tilde{R}_2 \geq S(V)_{\sigma_3}, \tag{22c}$$

$$(\tilde{R}_1 - R_1) + (\tilde{R}_2 - R_2) < I(U; V)_{\sigma_3} \tag{22d}$$

$$\tilde{R}_1 \geq R_1 \geq 0, \quad \tilde{R}_2 \geq R_2 \geq 0, \tag{22e}$$

$$C_1 + C_2 \leq C, \quad C_1 \geq 0, \quad C_2 \geq 0, \tag{22f}$$

where  $C_i \triangleq \frac{1}{n} \log_2 N_i$ ; for  $i = 1, 2$ . This implies that  $\tilde{M}_{AB}^{(n)}$  is  $\epsilon$ -faithful to  $\bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n}$  with probability sufficiently close to one, and hence,  $\hat{M}_{AB}^{(n)}$  is also  $\epsilon$ -faithful to  $M_{AB}^{\otimes n}$  with respect to  $\rho_{AB}^{\otimes n}$ , i.e., (17) is satisfied. Therefore, there exists a distributed protocol with parameters  $(n, 2^{nR_1}, 2^{nR_2}, 2^{nC})$  such that its overall POVM  $\hat{M}_{AB}^{(n)}$  is  $\epsilon$ -faithful to  $M_{AB}^{\otimes n}$  with respect to  $\rho_{AB}^{\otimes n}$ . Lastly, we complete the proof of the theorem using the following lemma.

**Lemma 5.** *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 sextuple  $(R_1, R_2, C_1, C_2, \tilde{R}_1, \tilde{R}_2)$  satisfies the inequalities in (22). 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.* The prove follows by Fourier-Motzkin elimination [17]. □

## V. Q-C DISTRIBUTED RATE DISTORTION THEORY

As an application of faithful simulation of distributed measurements (Theorem 2), we consider the distributed extension of q-c rate distortion coding [5]. This problem is a quantum counterpart of the classical distributed source coding. In this setting, consider a memoryless bipartite quantum source, characterized by  $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ . Alice and Bob have access to sub-systems  $A$  and  $B$ , characterized by  $\rho_A \in \mathcal{D}(\mathcal{H}_A)$  and  $\rho_B \in \mathcal{D}(\mathcal{H}_B)$ , respectively, where  $\rho_A = \text{Tr}_B\{\rho_{AB}\}$  and  $\rho_B = \text{Tr}_A\{\rho_{AB}\}$ . They both perform a measurement on  $n$  copies of their sub-systems and send the classical bits to Eve. Upon receiving the classical bits sent by Alice and Bob, a reconstruction state is produced by Eve. The objective of Eve is to produce a reconstruction of the source  $\rho_{AB}$  within a targeted distortion threshold which is measured by a given distortion observable.

### A. Problem Formulation

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

**Definition 5.** A q-c source coding setup is characterized by a triple  $(\Psi_{RAB}^{\rho_{AB}}, \mathcal{H}_{\hat{X}}, \Delta)$ , where  $\Psi_{RAB}^{\rho_{AB}} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_A \otimes \mathcal{H}_B)$  is a purified quantum state,  $\mathcal{H}_{\hat{X}}$  is a reconstruction Hilbert space, and  $\Delta \in \mathcal{B}(\mathcal{H}_R \otimes \mathcal{H}_{\hat{X}})$ , which satisfies  $\Delta \geq 0$ , is a distortion observable.

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

**Definition 6.** An  $(n, \Theta_1, \Theta_2)$  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  $\Theta_1$  and  $\Theta_2$  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  $\rho_{AB}$  is given by the following operation

$$\mathcal{N}_{A^n B^n \rightarrow \hat{X}^n} : \rho_{AB}^{\otimes n} \mapsto \sum_{i,j} \text{Tr}\{(\Lambda_i^A \otimes \Lambda_j^B) \rho_{AB}^{\otimes n}\} S_{i,j}, \quad (23)$$

where  $\{\Lambda_i^A\}$  and  $\{\Lambda_j^B\}$  are the operators of the POVMs  $M_A^{(n)}$  and  $M_B^{(n)}$ , respectively. With this notation and given a q-c source coding setup as in Definition 5, the distortion of a  $(n = 1, \Theta_1, \Theta_2)$  q-c protocol is measured as

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

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

$$\Delta^{(n)} = \frac{1}{n} \sum_{i=1}^n \Delta_{R_i \hat{X}_i} \otimes I_{R\hat{X}}^{\otimes [n] \setminus i}, \quad (24)$$

where  $\Delta_{R_i \hat{X}_i}$  is understood as the observable  $\Delta$  acting on the  $i$ th 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  $(n, \Theta_1, \Theta_2)$  q-c protocol is given by

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

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

The authors in [5] 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  $\Delta_{\hat{x}} \geq 0$  acts on the reference Hilbert space and  $\hat{\mathcal{X}}$  is the reconstruction alphabet. In this paper, we allow  $\Delta$  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 7.** For a q-c source coding setup  $(\Psi_{RAB}^{\rho_{AB}}, \mathcal{H}_{\hat{X}}, \Delta)$ , 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  $(n, \Theta_1, \Theta_2)$  q-c protocol satisfying

$$\begin{aligned} \frac{1}{n} \log_2 \Theta_i &\leq R_i + \epsilon, \quad i = 1, 2, \\ \bar{d}(\rho_{AB}^{\otimes n}, \mathcal{N}_{A^n B^n \rightarrow \hat{X}^n}) &\leq D + \epsilon, \end{aligned}$$

where  $\mathcal{N}_{A^n B^n \rightarrow \hat{X}^n}$  is defined as in (23). 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

We provide an inner bound to the achievable rate-distortion region which is stated in the following theorem. We employ a q-c protocol based on a randomized faithful simulation strategy involving a time sharing classical random variable  $Q$  that is independent of the quantum source. This can be viewed as a conditional version of the faithful simulation problem considered in Section III.

**Theorem 3.** *For a q-c source coding setup  $(\Psi_{RAB}^{\rho_{AB}}, \mathcal{H}_{\hat{X}}, \Delta)$ , any rate-distortion triplet  $(R_1, R_2, D)$  satisfying the following inequalities is achievable*

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

for POVM of the form  $M_{AB} = \sum_{q \in \mathcal{Q}} P_Q(q) M_A^q \otimes M_B^q$ , where for every  $q \in \mathcal{Q}$ ,  $M_A^q \triangleq \{\Lambda_u^{A,q}\}_{u \in \mathcal{U}}$  and  $M_B^q \triangleq \{\Lambda_v^{B,q}\}_{v \in \mathcal{V}}$  are POVMs acting on  $\mathcal{H}_A \otimes \mathcal{H}_B$ , and reconstruction states  $\{S_{u,v,q}\}$  with each state in  $\mathcal{D}(\mathcal{H}_{\hat{X}})$ , and some finite sets  $\mathcal{U}, \mathcal{V}$  and  $\mathcal{Q}$ . The quantum mutual information quantities are computed according to the auxiliary states  $\sigma_1^{RUBQ} \triangleq \sum_{q \in \mathcal{Q}} P_Q(q) (\text{id}_R \otimes M_A^q \otimes \text{id}_B) (\Psi_{RAB}^{\rho_{AB}}) \otimes |q\rangle\langle q|$ ,  $\sigma_2^{RAVQ} \triangleq \sum_{q \in \mathcal{Q}} P_Q(q) (\text{id}_R \otimes \text{id}_A \otimes M_B^q) \otimes |q\rangle\langle q|$ , and  $\sigma_3^{RUVQ} \triangleq \sum_{q \in \mathcal{Q}} P_Q(q) (\text{id}_R \otimes M_A^q \otimes M_B^q) (\Psi_{RAB}^{\rho_{AB}}) \otimes |q\rangle\langle q|$ , where  $(U, V)$  represents the output of  $M_{AB}$ , and  $\mathcal{N}_{AB \rightarrow \hat{X}} : \rho_{AB} \mapsto \sum_{u,v,q} P_Q(q) \text{Tr}\{(\Lambda_u^{A,q} \otimes \Lambda_v^{B,q}) \rho_{AB}\} S_{u,v,q}$ .

**Remark 3.** Note that for the auxiliary states  $\sigma_i, i = 1, 2, 3$ , we have  $I(R; Q) = 0$ .

*Proof.* In the interest of brevity, we provide the proof for the special case, when the time sharing random variable is trivial, i.e.,  $\mathcal{Q}$  is empty. An extension to the more general case is straightforward but tedious. For the special case, 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  $\mathcal{N}_{AB \rightarrow \hat{X}}$  be the mapping corresponding to these POVMs and the reconstruction states. Then,  $d(\rho_{AB}, \mathcal{N}_{AB \rightarrow \hat{X}}) \leq D$ . According to Theorem 2, for any  $\epsilon > 0$ , there exists an  $(n, 2^{nR_1}, 2^{nR_2}, N)$  distributed protocol for  $\epsilon$ -faithful simulation of  $M_A^{\otimes n} \otimes M_B^{\otimes n}$  with respect to  $\rho_{AB}^{\otimes n}$  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  $f^{(\mu)}$  be the POVMs and the deterministic decoding functions 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 quantum 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  $S_{i,j} \triangleq \sum_{u^n, v^n} \mathbb{1}\{f^{(\mu)}(i, j) = (u^n, v^n)\} S_{u^n, v^n}$ , where  $S_{u^n, v^n} = \otimes_i S_{u_i, v_i}$ . Further, let the corresponding mappings be denoted as  $\tilde{\mathcal{N}}_{A^n B^n \rightarrow \hat{X}^n}^{(\mu)}$ . With this notation, for the average of these random protocols, the following bounds hold:

$$\begin{aligned} \frac{1}{N} \sum_{\mu} \bar{d}(\rho_{AB}^{\otimes n}, \tilde{\mathcal{N}}_{A^n B^n \rightarrow \hat{X}^n}^{(\mu)}) &= \frac{1}{N} \sum_{\mu} \text{Tr} \left\{ \Delta^{(n)} (\text{id} \otimes \tilde{\mathcal{N}}_{A^n B^n \rightarrow \hat{X}^n}^{(\mu)}) \Psi_{R^n A^n B^n}^{\rho_{AB}} \right\} \\ &= \text{Tr} \left\{ \Delta^{(n)} (\text{id} \otimes \mathcal{N}_{AB \rightarrow \hat{X}}^{\otimes n}) \Psi_{R^n A^n B^n}^{\rho_{AB}} \right\} + \text{Tr} \left\{ \Delta^{(n)} (\text{id} \otimes (\mathcal{N}_{AB \rightarrow \hat{X}}^{\otimes n} - \tilde{\mathcal{N}}_{A^n B^n \rightarrow \hat{X}^n}^{(\mu)})) \Psi_{R^n A^n B^n}^{\rho_{AB}} \right\} \\ &\leq \text{Tr} \left\{ \Delta^{(n)} ((\text{id}_R \otimes \mathcal{N}_{AB \rightarrow \hat{X}})(\Psi_{RAB}^{\rho_{AB}})) \right\} + \|\Delta^{(n)} (\text{id} \otimes (\mathcal{N}_{AB \rightarrow \hat{X}}^{\otimes n} - \tilde{\mathcal{N}}_{A^n B^n \rightarrow \hat{X}^n}^{(\mu)})) \Psi_{R^n A^n B^n}^{\rho_{AB}}\|_1 \\ &\leq D + \|\Delta^{(n)}\|_{\infty} \|(\text{id} \otimes (\mathcal{N}_{AB \rightarrow \hat{X}}^{\otimes n} - \tilde{\mathcal{N}}_{A^n B^n \rightarrow \hat{X}^n}^{(\mu)})) \Psi_{R^n A^n B^n}^{\rho_{AB}}\|_1 \\ &\leq D + \|\Delta^{(n)}\|_{\infty} \|(\text{id} \otimes (M_A^{\otimes n} \otimes M_B^{\otimes n} - \tilde{M}_{AB})) \Psi_{R^n A^n B^n}^{\rho_{AB}}\|_1 \\ &\leq D + \epsilon \|\Delta\|_{\infty}, \end{aligned}$$



where  $\tilde{\mathcal{N}}_{AB \rightarrow \hat{X}}$  is the average of  $\tilde{\mathcal{N}}_{AB \rightarrow \hat{X}}^{(\mu)}$ , and  $\tilde{M}_{AB}$  is the overall POVM of the underlying distributed protocol as given in (4). The first inequality holds by the fact that  $|\text{Tr}\{A\}| \leq \|A\|_1$ . The second inequality follows by Lemma 6 given in the sequel. The third inequality is due to the monotonicity of the trace-distance [16] with respect to the quantum channel given by  $\text{id} \otimes \mathcal{L}_{UV \rightarrow \hat{X}}^{\otimes n}$ , where

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

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

**Lemma 6.** *For any operator  $A$  and  $B$  acting on a Hilbert space  $\mathcal{H}$  the following inequalities hold.*

$$\|BA\|_1 \leq \|B\|_\infty \|A\|_1, \quad \text{and} \quad \|AB\|_1 \leq \|B\|_\infty \|A\|_1.$$

*Proof.* See Exercise 12.2.1 in [16].  $\square$

One can observe that the rate region in Theorem 3 matches in form with the classical Berger-Tung region when  $\rho_{AB}$  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 [18].

### 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  $(\Psi_{RAB}^{\rho_{AB}}, \mathcal{H}_{\hat{X}}, \Delta)$ , if any triplet  $(R_1, R_2, D)$  is achievable, then the following inequalities must be satisfied*

$$R_1 \geq I(W_1; R | W_2, Q)_\sigma, \tag{25a}$$

$$R_2 \geq I(W_2; R | W_1, Q)_\sigma, \tag{25b}$$

$$R_1 + R_2 \geq I(W_1, W_2; R | Q)_\sigma, \tag{25c}$$

$$D \geq \text{Tr}\{\Delta\} \sigma^{R\hat{X}}, \tag{25d}$$

for some state  $\sigma^{W_1 W_2 R Q \hat{X}}$  which can be written as

$$\sigma^{W_1 W_2 R Q \hat{X}} = (\text{id} \otimes \mathcal{N}_{AB \rightarrow W_1 W_2 Q \hat{X}})(\Psi_{RAB}^{\rho_{AB}}),$$

where  $Q$  represents an auxiliary quantum state, and  $\mathcal{N}_{AB \rightarrow W_1 W_2 Q \hat{X}}$  is a quantum test channel with  $I(R; Q)_\sigma = 0$ .

*Proof.* Suppose the triplet  $(R_1, R_2, D)$  is achievable. Then, from Definition 7, for all  $\epsilon > 0$ , there exists an  $(n, \Theta_1, \Theta_2)$  q-c protocol satisfying the inequalities in the definition. Let  $M_A \triangleq \{\Lambda_{l_1}^A\}$ ,  $M_B \triangleq \{\Lambda_{l_2}^B\}$  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

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

where the state  $\tau$  is defined as

$$\tau^{L_1 L_2 R^n \hat{X}^n} \triangleq \sum_{l_1, l_2} |l_1, l_2\rangle\langle l_1, l_2| \otimes \text{Tr}_{A^n B^n} \left\{ (\text{id} \otimes \Lambda_{l_1}^A \otimes \Lambda_{l_2}^B) \Psi_{R^n A^n B^n}^{\rho_{AB}} \right\} \otimes S_{l_1, l_2}.$$

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 [19]. Consider a tuple of Hilbert spaces  $\mathcal{H}_k$ ,  $k = 1, 2, \dots, n$  with inner products  $\langle \cdot | \cdot \rangle_k$ . Define  $\bigoplus_{k=1}^n \mathcal{H}_k$  as the collection of tuples of vectors  $(|x\rangle_1, |x\rangle_2, \dots, |x\rangle_n)$ . The inner product of two tuples  $(|x\rangle_1, |x\rangle_2, \dots, |x\rangle_n)$  and  $(|y\rangle_1, |y\rangle_2, \dots, |y\rangle_n)$  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  $(A_1, A_2, \dots, A_n)$ , where  $A_k$  operates on  $\mathcal{H}_k$ , and  $\text{Tr}(A) = \sum_{k=1}^n \text{Tr}(A_k)$ . 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  $\sigma$  is defined below

$$\sigma^{L_1 L_2 R Q \hat{X}} \triangleq \sum_{l_1, l_2} \frac{|l_1, l_2\rangle\langle l_1, l_2|}{n} \otimes \left( \bigoplus_{j=1}^n \left( \text{Tr}_{R_{j+1}^n A^n B^n} \left\{ (\text{id} \otimes \Lambda_{l_1}^A \otimes \Lambda_{l_2}^B) \Psi_{R^n A^n B^n}^{\rho_{AB}} \right\} \otimes |j\rangle\langle j| \otimes \text{Tr}_{\hat{X}^n \sim j} \{S_{l_1, l_2}\} \right) \right), \quad (26)$$

where  $\text{Tr}_{\hat{X}^n \sim j}$  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

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

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, which is a truncated version of the so-called Fock space [20].

Similarly, for Bob's rate we have

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

For the sum-rate, the following inequalities hold

$$\begin{aligned} n(R_1 + R_2 + 2\epsilon) &\geq H(L_1, L_2) \geq I(L_1, L_2; R^n)_\tau \\ &= \sum_{j=1}^n I(L_1, L_2; R_j | R^{j-1})_\tau \\ &= nI(L_1, L_2; R | Q)_\sigma. \end{aligned}$$

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

$$D + \epsilon \geq \frac{1}{n} \sum_{j=1}^n \text{Tr} \left\{ \left( \Delta_{R_j \hat{X}_j} \otimes I_{R \hat{X}}^{\otimes [n]j} \right) (\text{id} \otimes \mathcal{N}_{A^n B^n \rightarrow \hat{X}^n}) (\Psi_{R^n A^n B^n}^{\rho_{AB}}) \right\}$$

$$\begin{aligned}
&= \frac{1}{n} \sum_{j=1}^n \text{Tr} \left\{ \left( \Delta_{R_j \hat{X}_j} \otimes I_{R_1^{j-1}} \otimes I_{R_{j+1}^n \hat{X}^{n-j}} \right) (\text{id} \otimes \mathcal{N}_{A^n B^n \mapsto \hat{X}^n}) (\Psi_{R^n A^n B^n}^{\rho_{AB}}) \right\} \\
&= \frac{1}{n} \sum_{j=1}^n \text{Tr} \left\{ \left( \Delta_{R_j \hat{X}_j} \otimes I_{R_1^{j-1}} \right) \left( \text{Tr}_{R_{j+1}^n \hat{X}^{n-j}} \left\{ (\text{id} \otimes \mathcal{N}_{A^n B^n \mapsto \hat{X}^n}) (\Psi_{R^n A^n B^n}^{\rho_{AB}}) \right\} \right) \right\} \\
&\stackrel{(a)}{=} \text{Tr} \{ (\Delta \otimes I_Q) \sigma^{RQ\hat{X}} \},
\end{aligned}$$

where the third equality holds, because of the following argument. From (26), one can show by partially tracing over  $(L_1, L_2)$ , that

$$\sigma^{RQ\hat{X}} = \text{Tr}_{L_1, L_2} \{ \sigma^{L_1 L_2 RQ\hat{X}} \} = \bigoplus_j \frac{1}{n} |j\rangle\langle j| \otimes \text{Tr}_{R_{j+1}^n \hat{X}^{n-j}} \{ (\text{id} \otimes \mathcal{N}_{A^n B^n \mapsto \hat{X}^n}) (\Psi_{R^n A^n B^n}^{\rho_{AB}}) \}, \quad (27)$$

and  $I_Q \triangleq \bigoplus_{j=1}^n (I_R^{\otimes(j-1)} \otimes |j\rangle\langle j|)$ . 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

$$\text{Tr} \{ (\Delta \otimes I_Q) \sigma^{RQ\hat{X}} \} = \text{Tr} \{ \Delta \sigma^{R\hat{X}} \}.$$

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  $\sigma^{L_1 L_2 RQ\hat{X}}$  as

$$\sigma^{L_1 L_2 RQ\hat{X}} = \left( \bigoplus_{j=1}^n \frac{1}{n} \sum_{l_1, l_2} |l_1, l_2\rangle\langle l_1, l_2| \otimes \left( \text{Tr}_{R_{j+1}^n A^n B^n} \left\{ (\text{id} \otimes \Lambda_{l_1}^A \otimes \Lambda_{l_2}^B) \Psi_{R^n A^n B^n}^{\rho_{AB}} \right\} \otimes |j\rangle\langle j| \otimes \text{Tr}_{\hat{X}^{n-j}} \{ S_{l_1, l_2} \} \right) \right).$$

Next, we identify a quantum channel  $\mathcal{N}_{AB \mapsto L_1 L_2 Q\hat{X}} : \rho_{AB} \mapsto \sigma^{L_1 L_2 Q\hat{X}}$ . For that and for any  $j$  define the following intermediate quantum channels:

$$\mathcal{N}_{AB \mapsto L_1 L_2 R^{(j-1)}\hat{X}}^{(j)}(\omega_{AB}) \triangleq \sum_{l_1, l_2} |l_1, l_2\rangle\langle l_1, l_2| \otimes \left( \text{Tr}_{R_{j+1}^n A^n B^n} \left\{ (\text{id}_{R^{n-j}} \otimes \Lambda_{l_1}^A \otimes \Lambda_{l_2}^B) (\omega_{AB} \otimes E_j) \right\} \otimes \text{Tr}_{\hat{X}^{n-j}} \{ S_{l_1, l_2} \} \right),$$

where  $E_j = \Psi_{(RAB)_{\sim j}}^{\rho_{AB}}$ . One can verify that  $\mathcal{N}_{AB \mapsto L_1 L_2 R^{(j-1)}\hat{X}}^{(j)}$  is indeed a quantum channel. With these definitions, let

$$\mathcal{N}_{AB \mapsto L_1 L_2 Q\hat{X}}(\omega_{AB}) \triangleq \bigoplus_j \frac{1}{n} \left( \mathcal{N}_{AB \mapsto L_1 L_2 R^{(j-1)}\hat{X}}^{(j)}(\omega_{AB}) \otimes |j\rangle\langle j| \right).$$

Using the property of direct-sum operation, one can verify that  $\mathcal{N}_{AB \mapsto L_1 L_2 Q\hat{X}}$  is a valid quantum channel, moreover,

$$\sigma^{L_1 L_2 RQ\hat{X}} = (\text{id} \otimes \mathcal{N}_{AB \mapsto L_1 L_2 Q\hat{X}}) (\Psi_{RAB}^{\rho_{AB}}).$$

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

$$\begin{aligned}
\sigma^{RQ} &= \text{Tr}_{L_1 L_2 \hat{X}} (\sigma^{L_1 L_2 RQ\hat{X}}) = \bigoplus_{j=1}^n \frac{1}{n} \sum_{l_1, l_2} \left( \text{Tr}_{R_{j+1}^n A^n B^n} \left\{ (\text{id} \otimes \Lambda_{l_1}^A \otimes \Lambda_{l_2}^B) \Psi_{R^n A^n B^n}^{\rho_{AB}} \right\} \right) \otimes |j\rangle\langle j| \\
&= \bigoplus_{j=1}^n \frac{1}{n} \left( \text{Tr}_{R_{j+1}^n A^n B^n} \left\{ \Psi_{R^n A^n B^n}^{\rho_{AB}} \right\} \right) \otimes |j\rangle\langle j| \\
&= \bigoplus_{j=1}^n \frac{1}{n} \left( \text{Tr}_{AB} \{ \Psi_{RAB}^{\rho_{AB}} \} \right)^{\otimes j} \otimes |j\rangle\langle j| \\
&= \text{Tr}_{AB} \{ \Psi_{RAB}^{\rho_{AB}} \} \otimes \left( \bigoplus_{j=1}^n \frac{1}{n} \left( \text{Tr}_{AB} \{ \Psi_{RAB}^{\rho_{AB}} \} \right)^{\otimes(j-1)} \otimes |j\rangle\langle j| \right),
\end{aligned}$$

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

**Remark 4.** One may question the computability of the outer bound provided in Theorem 4. The computability of this bound depends on the dimensionality of the auxiliary space  $\mathcal{H}_Q$  defined in the theorem. Currently, we are unable to bound the dimension of the Hilbert space  $\mathcal{H}_Q$ , but aim to provide one in our future work. As a matter of fact, the current outer bounds for the equivalent classical distributed rate distortion problem still suffers from the computability issue. The first outer bound to the classical problem was provided in [9] and a recent substantial improvement was made by authors in [21]. Both of these bounds suffer from the absence of cardinality bounds on at least one of the variables used and hence cannot be claimed to be computable using finite resources. □

## VI. SIMULATION OF POVMS WITH STOCHASTIC PROCESSING

We now provide an extension of the Winter's point-to-point measurement compression scheme [1] (discussed in Section II) with stochastic processing. We assume that the receiver (Bob) has access to additional private randomness, and he is allowed to use this additional resource to perform any stochastic mapping of the received classical bits. In fact, the overall effect on the quantum state can be assumed to be a measurement which is a concatenation of the POVM Alice performs and the stochastic map Bob implements. Hence, Alice in this case, does not remain aware of the measurement outcome. It is for this reason that [2] describes this as a non-feedback problem, with the sender not required to know the outcomes of the measurement. With the availability of additional resources, such a formulation is expected to help reduce the overall resources needed.

### A. Problem Formulation

**Definition 8.** For a given finite set  $\mathcal{X}$ , and a Hilbert space  $\mathcal{H}_A$ , a measurement simulation protocol with stochastic processing with parameters  $(n, \Theta, N)$  is characterized by

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

2) a Bob's classical stochastic map  $P^{(\mu)}(x^n|l)$  for all  $l \in \mathcal{L}$ ,  $x^n \in \mathcal{X}^n$  and  $\mu \in [1, N]$ .

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

$$\tilde{\Lambda}_{x^n} \triangleq \frac{1}{N} \sum_{\mu, l} P^{(\mu)}(x^n|l) \Lambda_l^{(\mu)}, \quad \forall x^n \in \mathcal{X}^n, \quad (28)$$

where  $\Lambda_l^{(\mu)}$  are the operators corresponding to the sub-POVMs  $\tilde{M}^{(\mu)}$ .

In the above definition,  $\Theta$  characterizes the amount of classical bits communicated from Alice to Bob, and the amount of common randomness is determined by  $N$ , with  $\mu$  being the common randomness bits distributed among the parties. The classical stochastic mappings induced by  $P^{(\mu)}$  represents the action of Bob on the received classical bits.

**Definition 9.** Given a POVM  $M$  acting on  $\mathcal{H}_A$ , and a density operator  $\rho \in \mathcal{D}(\mathcal{H}_A)$ , a pair  $(R, C)$  is said to be achievable, if for all  $\epsilon > 0$  and for all sufficiently large  $n$ , there exists a measurement simulation protocol with stochastic processing with parameters  $(n, \Theta, N)$  such that its overall sub-POVM  $\tilde{M}$  is  $\epsilon$ -faithful to  $M^{\otimes n}$  with respect to  $\rho^{\otimes n}$  (see Definition 2), and

$$\frac{1}{n} \log_2 \Theta \leq R + \epsilon, \quad \frac{1}{n} \log_2 N \leq C + \epsilon.$$

The set of all achievable pairs is called the achievable rate region.

The following theorem characterizes the achievable rate region.

**Theorem 5.** For any density operator  $\rho \in \mathcal{D}(\mathcal{H}_A)$  and any POVM  $M \triangleq \{\Lambda_x\}_{x \in \mathcal{X}}$  acting on the Hilbert space  $\mathcal{H}_A$ , a pair  $(R, C)$  is achievable if and only if there exist a POVM  $\bar{M}_A \triangleq \{\bar{\Lambda}_w^A\}_{w \in \mathcal{W}}$ , with  $\mathcal{W}$  being a finite set, and a stochastic map  $P_{X|W} : \mathcal{W} \rightarrow \mathcal{X}$  such that

$$R \geq I(R; W)_\sigma \quad \text{and} \quad R + C \geq I(RX; W)_\sigma,$$

$$\Lambda_x = \sum_{w \in \mathcal{W}} P_{X|W}(x|w) \bar{\Lambda}_w^A, \quad \forall x \in \mathcal{X}.$$

where  $\sigma_{RWX} \triangleq \sum_{w,x} \sqrt{\rho} \bar{\Lambda}_w^A \sqrt{\rho} \otimes P_{X|W}(x|w) |w\rangle\langle w| \otimes |x\rangle\langle x|$ .

**Remark 5.** An alternative characterization of the above rate region can also be obtained in terms of Holevo information. For this, we define the following ensemble  $\{\lambda_x, \hat{\rho}_x\}$  as

$$\lambda_x = \sum_{w \in \mathcal{W}} \lambda_w^A P_{X|W}(x|w) \quad \text{and} \quad \hat{\rho}_x = \sum_{w \in \mathcal{W}} P_{W|X} \hat{\rho}_w^A,$$

for  $\{\lambda_w^A, \hat{\rho}_w^A\}$  being the canonical ensemble associated with the POVM  $M$  and the state  $\rho$  as defined in (6). With this ensemble, we have

$$I(R; W)_\sigma = \chi(\{\lambda_w^A, \hat{\rho}_w^A\}) \quad \text{and} \quad I(RX; W)_\sigma = I(X; W)_\sigma + \chi(\{\lambda_w^A, \hat{\rho}_w^A\}) - \chi(\{\lambda_x, \hat{\rho}_x\}).$$

### B. Proof of Achievability of Theorem 5

Suppose there exist a POVM  $\bar{M}_A$  and a stochastic map  $P_{X|W} : \mathcal{W} \rightarrow \mathcal{X}$ , such that  $M$  can be decomposed as

$$\Lambda_x = \sum_w P_{X|W}(x|w) \bar{\Lambda}_w^A, \quad \forall x \in \mathcal{X}. \quad (29)$$

We begin by defining a canonical ensemble corresponding to  $\bar{M}_A$  as  $\{\lambda_w^A, \hat{\rho}_w^A\}_{w \in \mathcal{W}}$ , where

$$\lambda_w^A \triangleq \text{Tr}\{\bar{\Lambda}_w^A \rho\}, \quad \text{and} \quad \hat{\rho}_w^A \triangleq \frac{1}{\lambda_w^A} \sqrt{\rho} \bar{\Lambda}_w^A \sqrt{\rho}. \quad (30)$$

Similarly, for each  $w^n \in \mathcal{W}^n$ , we also define

$$\Lambda_{w^n}^A \triangleq \hat{\Pi} \Pi_\rho \Pi_{w^n} \hat{\rho}_{w^n}^A \Pi_{w^n} \Pi_\rho \hat{\Pi},$$

where  $\hat{\rho}_{w^n}^A \triangleq \bigotimes_i \hat{\rho}_{w_i}^A$ , and  $\hat{\Pi}, \Pi_\rho$  and  $\Pi_{w^n}$  are similar to the ones defined in Section IV. Using the above definitions, we now construct the approximating POVM.

1) *Construction of Random POVMs:* In what follows, we construct a collection of random POVMs. Fix  $R$  and  $C$  as two positive integers. Let  $\mu \in [1, 2^{nC}]$  denote the common randomness shared between the sender and receiver. For each  $\mu \in [1, 2^{nC}]$ , randomly and independently select  $2^{nR}$  sequences  $W^{n,(\mu)}(l)$  according to the pruned distributions, i.e.,

$$\mathbb{P}\left((W^{n,(\mu)}(l)) = (w^n)\right) = \begin{cases} \frac{\lambda_{w^n}^A}{(1-\varepsilon)} & \text{for } w^n \in \mathcal{T}_\delta^{(n)}(W) \\ 0 & \text{otherwise} \end{cases}. \quad (31)$$

For  $w^n \in \mathcal{T}_\delta^{(n)}(W)$ , let the operators of the POVM  $\tilde{M}_A^{(\mu)}$  be  $\{A_{w^n}^{(\mu)}; w^n \in \mathcal{W}^n\}$  for each  $\mu \in [1, 2^{nC}]$ , where  $A_{w^n}^{(\mu)}$  is defined as

$$A_{w^n}^{(\mu)} \triangleq \gamma_{w^n}^{(\mu)} \left( \sqrt{\rho_A}^{-1} \Lambda_{w^n}^A \sqrt{\rho_A}^{-1} \right) \quad \text{and} \quad \gamma_{w^n}^{(\mu)} = \frac{1}{2^{nR}} \sum_{l=1}^{2^{nR}} \frac{(1-\varepsilon)}{(1+\eta)} \mathbb{1}_{\{W^{n,(\mu)}(l)=w^n\}}, \quad (32)$$

with  $\eta \in (0, 1)$  being a parameter to be determined. Now, for each  $\mu \in [1, 2^{nC}]$  construct  $\tilde{M}_A^{(n, \mu)}$  as

$$\tilde{M}_A^{(n, \mu)} = \{A_{w^n}^{(\mu)} : w^n \in \mathcal{T}_\delta^{(n)}(W)\}.$$

Since the construction is very similar to the one used in Section IV, we make a claim similar to the one in Lemma 2. This claim gives us the first constraint on the classical rate of communication  $R$ , which ensures that the operators constructed above for all  $\mu \in [1, 2^{nC}]$  are valid sub-POVMs (characterized as the event  $E_1$ ) with high probability. The claim is as follows. If  $R > I(R; W)_\sigma$  then  $\mathbb{P}(E_1) \geq (1 - \delta_1)$  for some  $\delta_1 \in (0, 1/6)$ ; or in other words, with probability sufficiently close to one,  $\tilde{M}_A^{(n, \mu)}$  forms a sub-POVM for all  $\mu \in [1, 2^{nC}]$ . Note the definition of  $\sigma_{RWX}$  follows from the statement of theorem. From this, let  $[\tilde{M}_A^{(n, \mu)}]$  denote the completion of the corresponding sub-POVM  $\tilde{M}_A^{(n, \mu)}$  for  $\mu \in [1, 2^{nC}]$ . Let the operators completing these POVMs, given by  $I - \sum_{w^n} A_{w^n}^{(\mu)}$ , be denoted by  $A_{w_0^n}^{(\mu)}$  for some  $w_0^n \notin \mathcal{T}_\delta^{(n)}(W)$ , for all  $\mu \in [1, 2^{nC}]$ , and  $A_{w^n}^{(\mu)} = 0$  for  $w^n \notin \mathcal{T}_\delta^{(n)}(W) \cup \{w_0^n\}$ . Using this construction, we define the intermediate POVM  $\tilde{M}_A^{(n)}$  as  $\tilde{M}_A^{(n)} = \frac{1}{2^{nC}} \sum_\mu \tilde{M}_A^{(n, \mu)}$  and the operators of  $\tilde{M}_A^{(n)}$  as  $\tilde{\Lambda}_{w^n}^A = \frac{1}{2^{nC}} \sum_\mu A_{w^n}^{(\mu)}$ . Now, we define Bob's stochastic map as  $P_{X|W}^n$ , yielding the operators of the final approximating POVM as

$$\sum_{w^n \in \mathcal{W}^n} P_{X|W}^n(x^n|w^n) \tilde{\Lambda}_{w^n}^A, \quad x^n \in \mathcal{X}^n.$$

2) *Trace Distance*: Now, we compare the action of this approximating POVM on the input state  $\rho^{\otimes n}$  with that of the given POVM  $M$ , using the characterization provided in Definition 2. Specifically, we show using the expressions for canonical ensemble that, with probability close to one,

$$G \triangleq \sum_{x^n \in \mathcal{X}^n} \left\| \sum_{w^n \in \mathcal{W}^n} P_{X|W}^n(x^n|w^n) \sqrt{\rho^{\otimes n}} (\tilde{\Lambda}_{w^n}^A - \tilde{\Lambda}_{w_0^n}^A) \sqrt{\rho^{\otimes n}} \right\|_1 \leq \epsilon. \quad (33)$$

As a first step, we split and bound  $G$  as  $G \leq S_1 + S_2$ , where

$$S_1 \triangleq \sum_{x^n} \left\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) - \frac{1}{2^{nC}} \sum_{w^n \neq w_0^n} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \Lambda_{w^n}^A P_{X|W}^n(x^n|w^n) \right\|_1,$$

$$S_2 \triangleq \sum_{x^n} \left\| P_{X|W}^n(x^n|w_0^n) \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \left[ \sqrt{\rho^{\otimes n}} (I - \sum_{w^n} A_{w^n}^{(\mu)}) \sqrt{\rho^{\otimes n}} \right] \right\|_1.$$

Now we bound  $S_1$  by adding and subtracting an appropriate term and using triangle inequality as  $S_1 \leq S_{11} + S_{12}$ , where  $S_{11}$  and  $S_{12}$  are given by

$$S_{11} \triangleq \left\| \sum_{x^n} \left[ \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) \otimes |x^n\rangle\langle x^n| - \frac{1}{2^{nC}} \sum_{w^n \neq w_0^n} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) \otimes |x^n\rangle\langle x^n| \right] \right\|_1,$$

$$S_{12} \triangleq \left\| \sum_{x^n} \sum_{w^n \neq w_0^n} \left[ \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) - \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \Lambda_{w^n}^A P_{X|W}^n(x^n|w^n) \right] \otimes |x^n\rangle\langle x^n| \right\|_1.$$

Note that in the above expressions, we have used an additional triangle inequality for block operators (which is in fact an equality) to move the summation over  $\mathcal{X}^n$  inside the trace norm. Firstly, we show  $S_{11}$  is small with high probability. To simplify the notation, we define  $\sigma_{w^n} = \sum_{x^n} P_{X|W}^n(x^n|w^n) |x^n\rangle\langle x^n|$  which gives  $S_{11}$  as

$$S_{11} = \left\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A \otimes \sigma_{w^n} - \frac{1}{2^{n(R+C)}} \frac{(1-\epsilon)}{(1+\eta)} \sum_{l, \mu} \hat{\rho}_{W^{n, (\mu)}(l)}^A \otimes \sigma_{W^{n, (\mu)}(l)} \right\|_1.$$

We develop the following lemma to bound this term.

**Lemma 7.** Consider an ensemble given by  $\{\tilde{P}_{W^n}(w^n), \mathcal{T}_{w^n}\}$ , where  $\tilde{P}_{W^n}(w^n)$  is the pruned distribution as defined in (31) and  $\mathcal{T}_{w^n}$  is any tensor product state of the form  $\mathcal{T}_{w^n} = \bigotimes_{i=1}^n \mathcal{T}_w$ . Then, for any  $\epsilon_2 > 0$ ,  $\delta > 0$ , there exists functions  $\delta_s(\delta, \epsilon_2)$  and  $\delta'(\delta)$ , such that for all sufficiently large  $n$ , the inequality

$$\left\| \sum_{w^n} \lambda_{w^n}^A \mathcal{T}_{w^n} - \frac{1}{2^{n(R+C)}} \frac{(1-\epsilon)}{(1+\eta)} \sum_{l,\mu} \mathcal{T}_{W^{n,(\mu)}(l)} \right\|_1 \leq \epsilon_2, \quad (34)$$

holds with probability greater than  $1 - \delta_s(\delta, \epsilon_2)$ , if  $R+C \geq S(\sum_w \lambda_w^A \mathcal{T}_w) - \sum_w \lambda_w^A S(\mathcal{T}_w) + \delta' = \chi(\{\lambda_w^A, \mathcal{T}_w\}) + \delta'$ , where  $\{W^{n,(\mu)}(l) : l \in [1, 2^{nR}], \mu \in [1, 2^{nC}]\}$  are independent random vectors generated according to the pruned distribution given in (31), and  $\delta_s \searrow 0$ ,  $\delta' \searrow 0$  as  $\epsilon_2 \searrow 0$ ,  $\delta \searrow 0$ .

*Proof.* The proof of the lemma is provided in Appendix A-C □

Therefore, using the lemma above,  $S_{11}$  can be made arbitrarily small, for sufficiently large  $n$ , with high probability, given the constraints  $R+C \geq S(\sum_w \lambda_w^A \hat{\rho}_w^A \otimes \sigma_w) - \sum_w \lambda_w^A S(\hat{\rho}_w^A \otimes \sigma_w) + \delta' = \chi(\{\lambda_w^A, \{\hat{\rho}_w^A \otimes \sigma_w\}\}) + \delta' = I(RX; W)_\sigma + \delta'$ . Secondly, we bound  $S_{12}$  by applying expectation and using Gentle Measurement Lemma [16] as follows,

$$\begin{aligned} \mathbb{E}[S_{12}] &= \mathbb{E} \left[ \left\| \sum_{x^n} \sum_{w^n \neq w_0^n} \left[ \frac{1}{2^{nC}} \sum_{\mu=1}^M \gamma_{w^n}^{(\mu)} \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) - \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \Lambda_{w^n}^A P_{X|W}^n(x^n|w^n) \right] \otimes |x^n\rangle\langle x^n| \right\|_1 \right] \\ &\stackrel{(a)}{\leq} \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{x^n} \sum_{w^n \neq w_0^n} P_{X|W}^n(x^n|w^n) \mathbb{E} \left[ \gamma_{w^n}^{(\mu)} \|\hat{\rho}_{w^n}^A - \Lambda_{w^n}^A\|_1 \right] \\ &\stackrel{(b)}{=} \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \frac{\lambda_{w^n}^A}{(1+\eta)} \|\hat{\rho}_{w^n}^A - \Lambda_{w^n}^A\|_1 \\ &\stackrel{(c)}{=} \frac{1}{(1+\eta)} \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n}^A \|\hat{\rho}_{w^n}^A - \hat{\Pi} \Pi_\rho \Pi_{w^n} \hat{\rho}_{w^n}^A \Pi_{w^n} \Pi_\rho \hat{\Pi}\|_1 \stackrel{(d)}{\leq} \frac{(1-\epsilon)}{(1+\eta)} (2\sqrt{\epsilon'} + 2\sqrt{\epsilon''}) \triangleq \epsilon_3, \end{aligned} \quad (35)$$

where (a) is obtained by using triangle inequality and the linearity of expectation, (b) is obtained by marginalizing over  $x^n$  and using the fact that  $\mathbb{E}[\gamma_{w^n}^{(\mu)}] = \frac{\lambda_{w^n}^A}{(1+\eta)}$ , (c) is obtained by substitution, and finally (d) uses repeated application of the average gentle measurement lemma, by setting  $\epsilon_3 = \frac{(1-\epsilon)}{(1+\eta)} (2\sqrt{\epsilon'} + 2\sqrt{\epsilon''})$  with  $\epsilon_3 \searrow 0$  as  $\delta \searrow 0$  and,  $\epsilon' = \epsilon + 2\sqrt{\epsilon}$  and  $\epsilon'' = 2\epsilon + 2\sqrt{\epsilon}$  (see (35) in [2] for details). Finally, we show that the term corresponding to  $S_2$  can also be made arbitrarily small. This term can be simplified as follows

$$\begin{aligned} S_2 &\leq \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{x^n} P_{X|W}^n(x^n|w_0^n) \left\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A - \sum_{w^n \neq w_0^n} \sqrt{\rho^{\otimes n}} A_{w^n}^{(\mu)} \sqrt{\rho^{\otimes n}} \right\|_1, \\ &\leq \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \left\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A - \sum_{w^n \neq w_0^n} \gamma_{w^n}^{(\mu)} \hat{\rho}_{w^n}^A \right\|_1 + \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{w^n \neq w_0^n} \gamma_{w^n}^{(\mu)} \|\hat{\rho}_{w^n}^A - \Lambda_{w^n}^A\|_1 = S_{21} + S_{22}, \end{aligned}$$

where

$$S_{21} \triangleq \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \left\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A - \frac{(1-\epsilon)}{(1+\eta)} \frac{1}{2^{nR}} \sum_{l=1}^{2^{nR}} \hat{\rho}_{W_l^{n,(\mu)}}^A \right\|_1 \quad \text{and} \quad S_{22} \triangleq \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{w^n \neq w_0^n} \gamma_{w^n}^{(\mu)} \|\hat{\rho}_{w^n}^A - \Lambda_{w^n}^A\|_1. \quad (36)$$

Now, for the first term in (36) we use Lemma 7 and claim that for given any  $\epsilon_{w_0}, \delta_{w_0} \in (0, 1)$ , if

$$R > S \left( \sum_{w \in \mathcal{W}} \lambda_w^A \hat{\rho}_w \right) + \sum_{w \in \mathcal{W}} \lambda_w^A S(\hat{\rho}_w) = I(R; W)_\sigma,$$

then the probability of this term being greater than  $\epsilon_{w_0}$  is bounded by  $\delta_{w_0}$  for sufficiently large  $n$ , where  $\sigma$  is as defined in the statement of the theorem. Note that the requirements we obtain on  $R$  were already imposed when claiming the collection of operators  $A_{w^n}^{(\mu)}$  forms a sub-POVM. As for the second term in (36) we again use the gentle measurement Lemma and bound its expected value as

$$\mathbb{E} \left[ \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{w^n} \gamma_{w^n}^{(\mu)} \|\hat{\rho}_{w^n} - \Lambda_{w^n}\|_1 \right] = \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \frac{\lambda_{w^n}}{(1+\eta)} \|\hat{\rho}_{w^n} - \Lambda_{w^n}\|_1 \leq \epsilon_3,$$

where  $\epsilon_3$  is defined in (35).

In summary, we have performed the following sequence of steps. Firstly, we argued that  $\tilde{M}_A^{(n,\mu)}$  forms a valid sub-POVM for all  $\mu \in [1, 2^{nC}]$ , with high probability, when the rate  $R$  satisfies  $R > I(R; W)_\sigma$ . Secondly, we moved onto bounding the trace norm between the states obtained after the action for these approximating POVMs when compared with those obtained from the action of actual POVM  $M$ , characterized as  $G$  using Definition 2. As a first step in establishing this bound, we showed that  $G \leq S_1 + S_2$ . Considering  $S_1$ , we used triangle inequality and divided it into two terms:  $S_{11}$  and  $S_{12}$ . Then, using Lemma 7 we showed that for any given  $\epsilon_1 \in (0, 1)$ ,  $S_{11}$  can be made smaller than  $\epsilon_1$ , with high probability if  $R + C > I(RX; W)_\sigma$ . As for  $S_{12}$ , we showed that it goes to zero in the expected sense using (35). Finally, for the term given by  $S_2$ , we bounded this as a sum of two trace norms  $S_{21}$  and  $S_{22}$  given in (36). We showed that its first term can be made smaller than  $\epsilon_{w_0}$  with high probability if  $R > I(R; W)$  for sufficiently large  $n$ , and the second term was shown to approach zero in expected sense.

Now, using Markov inequality we argue the existence of at least one collection of POVMs that satisfies the statement of the Theorem 5 as follows. Note that  $S_{12}$  is same as  $S_{22}$ . Let  $E_1$  be the event defined earlier in the proof. Let us define  $E_2, E_3$  and  $E_4$  as the random variables corresponding to the terms  $S_{11}, S_{12}$  and  $S_{21}$ , respectively. Firstly, if  $R > I(R; W)_\sigma$ . then  $\mathbb{P}(E_1) \geq (1 - \delta_1)$ . Secondly, from Lemma 7, for all  $0 < \epsilon_2 < 1$ , and for all sufficiently large  $n$ , if

$$R + C \geq I(RX; W)_\sigma + \delta', \quad \text{and} \quad R \geq I(W; R)_\sigma + \delta',$$

then we have  $\mathbb{P}(E_2 \leq \epsilon_2) \geq 1 - \delta_s(\epsilon_2)$ ,  $\mathbb{P}(E_4 \leq \epsilon_{w_0}) \geq 1 - \delta_{w_0}$ . Thirdly, from (35) we have  $\mathbb{E}[E_3] \leq \epsilon_3$ . This implies, from the Markov inequality, that

$$\mathbb{P}(E_3 \geq 2\epsilon_3) \leq \frac{\mathbb{E}[E_3]}{2\epsilon_3} \leq \frac{1}{2}.$$

Using these bounds, we get

$$\begin{aligned} \mathbb{P}((E_1) \cap (E_2 < \epsilon_2) \cap (E_3 < 2\epsilon_3) \cap (E_4 < \epsilon_{w_0})) &\geq \mathbb{P}(E_1) + \mathbb{P}(E_2 < \epsilon_2) + \mathbb{P}(E_3 < 2\epsilon_3) + \mathbb{P}(E_4 < \epsilon_{w_0}) - 3 \\ &\geq 3 - (\delta_1 + \delta_s + \delta_{w_0}) + \frac{1}{2} - 3 > \frac{1}{4}, \end{aligned} \quad (37)$$

given that we choose  $0 < \delta_1, \delta_s, \delta_{w_0} < \frac{1}{12}$ . Note that  $G \leq E_2 + (2E_3) + E_4$ , and the inequality in (37) ensures that there exists a valid collection of sub-POVMs satisfying  $G \leq \epsilon_2 + 4\epsilon_3 + \epsilon_{w_0}$ , with non-vanishing probability. Therefore, using random coding arguments, there exists at least one collection of sub-POVMs with the above construction satisfying the statement of Theorem 5.

## VII. SIMULATION OF DISTRIBUTED POVMs WITH STOCHASTIC PROCESSING

In this section, we develop a stochastic processing variant of the distributed POVM simulation problem described in Section III. Let  $\rho_{AB}$  be a density operator acting on a composite Hilbert Space  $\mathcal{H}_A \otimes \mathcal{H}_B$ . Consider two measurements  $M_A$  and  $M_B$  on sub-systems  $A$  and  $B$ , respectively. Imagine again that we have three parties, named Alice, Bob and Eve, that are trying to collectively simulate a given measurement  $M_{AB}$  acting on the state



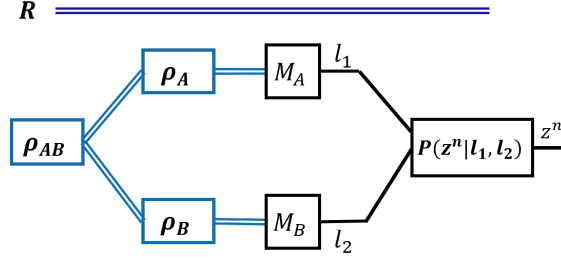


Fig. 3. The diagram depicting the distributed POVM simulation problem with stochastic processing. In this setting, Eve additionally has access to unlimited private randomness.

$\rho_{AB}$ , as shown in Fig. 3. In this version of distributed simulation, Eve additionally has access to unlimited private randomness. The problem is defined in the following.

**Definition 10.** For a given finite set  $\mathcal{Z}$ , and a Hilbert space  $\mathcal{H}_A \otimes \mathcal{H}_B$ , a distributed protocol with stochastic processing with parameters  $(n, \Theta_1, \Theta_2, N)$  is characterized by

- 1) a collections of Alice's sub-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  $|\mathcal{L}_1| \leq \Theta_1$ .
- 2) a collections of Bob's sub-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  $|\mathcal{L}_2| \leq \Theta_2$ .
- 3) Eve's classical stochastic map  $P^{(\mu)}(z^n | l_1, l_2)$  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 sub-POVM of this distributed protocol, given by  $\tilde{M}_{AB}$ , is characterized by the following operators:

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

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

In the above definition,  $(\Theta_1, \Theta_2)$  determines the amount of classical bits communicated from Alice and Bob to Eve. The amount of common randomness is determined by  $N$ . The classical stochastic maps  $P^{(\mu)}(z^n | l_1, l_2)$  represent the action of Eve on the received classical bits.

**Definition 11.** 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 triple  $(R_1, R_2, C)$  is said to be achievable, if for all  $\epsilon > 0$  and for all sufficiently large  $n$ , there exists a distributed protocol with stochastic processing with parameters  $(n, \Theta_1, \Theta_2, N)$  such that its overall sub-POVM  $\tilde{M}_{AB}$  is  $\epsilon$ -faithful to  $M_{AB}^{\otimes n}$  with respect to  $\rho_{AB}^{\otimes n}$  (see Definition 2), and

$$\frac{1}{n} \log_2 \Theta_i \leq R_i + \epsilon, \quad i = 1, 2, \quad \text{and} \quad \frac{1}{n} \log_2 N \leq C + \epsilon.$$

The set of all achievable triples  $(R_1, R_2, C)$  is called the achievable rate region.

The following theorem provides an inner bound to the achievable rate region, which is proved in Section VIII.

**Theorem 6.** Given a density operator  $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ , and a POVM  $M_{AB} = \{\Lambda_z^{AB}\}_{z \in \mathcal{Z}}$  acting on  $\mathcal{H}_A \otimes \mathcal{H}_B$  having a separable decomposition with stochastic integration (as in Definition 1), a triple  $(R_1, R_2, C)$  is achievable if the following inequalities are satisfied:

$$R_1 \geq I(U; RB)_{\sigma_1} - I(U; V)_{\sigma_3}, \quad (38a)$$

$$R_2 \geq I(V; RA)_{\sigma_2} - I(U; V)_{\sigma_3}, \quad (38b)$$

$$R_1 + R_2 \geq I(U; RB)_{\sigma_1} + I(V; RA)_{\sigma_2} - I(U; V)_{\sigma_3}, \quad (38c)$$

$$R_1 + C \geq I(U; RZV)_{\sigma_3} - I(U; V)_{\sigma_3}, \quad (38d)$$

$$R_2 + C \geq I(V; RZ)_{\sigma_3} - I(U; V)_{\sigma_3}, \quad (38e)$$

$$R_1 + R_2 + C \geq I(UV; RZ)_{\sigma_3}, \quad (38f)$$

for some decomposition with POVMs  $\bar{M}_A = \{\bar{\Lambda}_u^A\}_{u \in \mathcal{U}}$  and  $\bar{M}_B = \{\bar{\Lambda}_v^B\}_{v \in \mathcal{V}}$  and a stochastic map  $P_{Z|U,V} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{Z}$ , where  $\Psi_{RAB}^{\rho_{AB}}$  is a purification of  $\rho_{AB}$ , and the above information quantities are computed for the auxiliary states  $\sigma_1^{RUB} \triangleq (\text{id}_R \otimes \bar{M}_A \otimes \text{id}_B)(\Psi_{RAB}^{\rho_{AB}})$ ,  $\sigma_2^{RAV} \triangleq (\text{id}_R \otimes \text{id}_A \otimes \bar{M}_B)(\Psi_{RAB}^{\rho_{AB}})$ , and  $\sigma_3^{RUVZ} \triangleq \sum_{u,v,z} \sqrt{\rho_{AB}}(\bar{\Lambda}_u^A \otimes \bar{\Lambda}_v^B) \sqrt{\rho_{AB}} \otimes P_{Z|U,V}(z|u,v) |u\rangle\langle u| \otimes |v\rangle\langle v| \otimes |z\rangle\langle z|$ .

**Remark 6.** An alternative characterization of the above rate region can be obtained in terms of Holevo information. For this, we define the following ensemble  $\{\lambda_z, \hat{\rho}_z\}$  as

$$\lambda_z = \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} \lambda_{uv}^{AB} P_{Z|UV}(z|u,v) \quad \text{and} \quad \hat{\rho}_z = \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} P_{UV|Z}(u,v|z) \hat{\rho}_{uv}^{AB},$$

with  $\{\lambda_u^A, \hat{\rho}_u^A\}$ ,  $\{\lambda_v^B, \hat{\rho}_v^B\}$  and  $\{\lambda_{uv}^{AB}, \hat{\rho}_{uv}^{AB}\}$  being the canonical ensembles defined in (6), and  $P_{UV|Z}(u,v|z) = \lambda_{uv}^{AB} \cdot P_{Z|UV}(z|u,v) / \lambda_z$  for all  $(u,v,z) \in \mathcal{U} \times \mathcal{V} \times \mathcal{Z}$ . With this ensemble, we have  $I(U; RB)_{\sigma_1} = \chi(\{\lambda_u^A, \hat{\rho}_u^A\})$ ,  $I(V; RA)_{\sigma_2} = \chi(\{\lambda_v^B, \hat{\rho}_v^B\})$ , and  $I(UV; RZ)_{\sigma_3} = I(UV; Z) + \chi(\{\lambda_{uv}^{AB}, \hat{\rho}_{uv}^{AB}\}) - \chi(\{\lambda_z, \hat{\rho}_z\})$ .

## VIII. PROOF OF THEOREM 6

### A. Construction of POVMs

Suppose there exist POVMs  $\bar{M}_A \triangleq \{\bar{\Lambda}_u^A\}_{u \in \mathcal{U}}$  and  $\bar{M}_B \triangleq \{\bar{\Lambda}_v^B\}_{v \in \mathcal{V}}$  and a stochastic map  $P_{Z|UV} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{Z}$ , such that  $M_{AB}$  can be decomposed as

$$\Lambda_z^{AB} = \sum_{u,v} P_{Z|UV}(z|u,v) \bar{\Lambda}_u^A \otimes \bar{\Lambda}_v^B, \quad \forall z, \quad (39)$$

Note that the proof technique here is very different to the one used in Section IV for proving Theorem 2. Recall that in Theorem 2 we initiated the proof by constructing a protocol to faithfully simulate  $\bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n}$ . However, here we are not interested in faithfully simulating  $\bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n}$ . Instead, by carefully exploiting the private randomness Eve possesses, manifested in terms of the stochastic processing applied by her on the classical bits received, i.e.,  $P_{Z|U,V}$ , we aim to strictly reduce the sum rate constraints compared to the ones obtained in (5f) of Theorem 2. This requires a considerably different methodology. More specifically, Lemma 1 was employed in Theorem 2, which guaranteed that any two point-to-point POVMs that can individually approximate their corresponding original POVMs, can also faithfully approximate a measurement formed by the tensor product of the original POVMs performed on any state in the tensor product Hilbert space. Such a lemma cannot be developed in the setting involving a stochastic decoder. This is due to the fact that bits received from Alice and Bob are jointly perturbed by the stochastic decoder which doesn't allow a straightforward segmentation into two point-to-point problems. However, the analysis performed in the Section VIII actually modularizes the problem, using an asymmetric partitioning.

Nevertheless, we use the same POVM construction and binning operation as in the proof of Theorem 2, and hence we appeal to Section IV-A and IV-B for constructing the POVMs based on the codebook  $\mathcal{C}^{(\mu)}$  and binning them, resulting in the sub-POVMs  $M_1^{(n,\mu_1)}$  and  $M_2^{(n,\mu_2)}$  (see (14)), and  $M_A^{(n,\mu_1)}$  and  $M_B^{(n,\mu_2)}$  (see (15)), and their completions. All the notations used subsequently can be found in these sections. Therefore, the main focus of the proof hereon is to describe the decoder which is distinct from the one with deterministic mapping, in the sense that it employs the additional stochastic map, and a thorough analysis of the achievability result.

To start with, one can show by using a result similar to Lemma 2 that with probability sufficiently close to one,  $M_1^{(n,\mu_1)}$  and  $M_2^{(n,\mu_2)}$  form sub-POVMs for all  $\mu_1 \in [1, N_1]$  and  $\mu_2 \in [1, N_2]$  if  $\tilde{R}_1 > I(U; RB)_{\sigma_1}$  and

$\tilde{R}_2 > I(V; RA)_{\sigma_2}$ , where  $\sigma_1, \sigma_2$  are defined as in the statement of the theorem. Further, from  $D^{(\mu_1, \mu_2)}$  in (16) and  $F^{(\mu)}$ , as defined subsequently, we obtain the sub-POVM  $\tilde{M}_{AB}$  with the following operators.

$$\tilde{\Lambda}_{u^n, v^n}^{AB} \triangleq \frac{1}{N_1 N_2} \sum_{\mu_1=1}^{N_1} \sum_{\mu_2=1}^{N_2} \sum_{(i,j): F^{(\mu)}(i,j)=(u^n, v^n)} \Gamma_i^{A, (\mu_1)} \otimes \Gamma_j^{B, (\mu_2)}, \quad \forall (u^n, v^n) \in \mathcal{U}^n \times \mathcal{V}^n.$$

Now, we use the stochastic mapping to define the approximating sub-POVM  $\hat{M}_{AB}^{(n)} \triangleq \{\hat{\Lambda}_{z^n}\}$  as

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

### B. Trace Distance

In what follows, we show that  $\hat{M}_{AB}^{(n)}$  is  $\epsilon$ -faithful to  $M_{AB}^{\otimes n}$  with respect to  $\rho_{AB}^{\otimes n}$  (according to Definition 2), where  $\epsilon > 0$  can be made arbitrarily small. More precisely, using (39), we show that, with probability sufficiently close to 1, the following inequality holds

$$G = \sum_{z^n} \left\| \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B P_{Z|U, V}^n(z^n | u^n, v^n) - \tilde{\Lambda}_{u^n, v^n}^{AB} P_{Z|U, V}^n(z^n | u^n, v^n) \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \leq \epsilon. \quad (40)$$

#### Step 1: Isolating the effect of error induced by not covering

Consider the second term within  $G$ , which can be written as

$$\begin{aligned} & \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \tilde{\Lambda}_{u^n, v^n}^{AB} \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U, V}^n(z^n | u^n, v^n) \\ &= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{i, j} \sqrt{\rho_{AB}^{\otimes n}} \left( \Gamma_i^{A, (\mu_1)} \otimes \Gamma_j^{B, (\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U, V}^n(z^n | F^{(\mu)}(i, j)) \underbrace{\sum_{u^n, v^n} \mathbb{1}_{\{F^{(\mu)}(i, j) = (u^n, v^n)\}}}_{=1} \\ &= T + \tilde{T}, \end{aligned}$$

where

$$\begin{aligned} T &\triangleq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{\{i>0\} \cap \{j>0\}} \sqrt{\rho_{AB}^{\otimes n}} \left( \Gamma_i^{A, (\mu_1)} \otimes \Gamma_j^{B, (\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U, V}^n(z^n | F^{(\mu)}(i, j)), \\ \tilde{T} &\triangleq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{\{i=0\} \cup \{j=0\}} \sqrt{\rho_{AB}^{\otimes n}} \left( \Gamma_i^{A, (\mu_1)} \otimes \Gamma_j^{B, (\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U, V}^n(z^n | u_0^n, v_0^n). \end{aligned}$$

Hence, we have

$$G \leq S + \tilde{S}, \quad (41)$$

where

$$S \triangleq \sum_{z^n} \left\| \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B P_{Z|U, V}^n(z^n | u^n, v^n) \right) \sqrt{\rho_{AB}^{\otimes n}} - T \right\|_1, \quad (42)$$

and  $\tilde{S} \triangleq \sum_{z^n} \|\tilde{T}\|_1$ . Note that  $\tilde{S}$  captures the error induced by not covering the state  $\rho_{AB}^{\otimes n}$ . For the term corresponding to  $\tilde{S}$ , we prove the following result.

**Proposition 3.** *There exist functions  $\epsilon_{\tilde{S}}(\delta)$ , and  $\delta_{\tilde{S}}(\delta)$ , such that for all sufficiently small  $\delta$  and sufficiently large  $n$ , we have  $\mathbb{P}(\tilde{S} > \epsilon_{\tilde{S}}(\delta)) \leq \delta_{\tilde{S}}(\delta)$ , if  $\tilde{R}_1 > I(U; RB)_{\sigma_1}$  and  $\tilde{R}_2 > I(V; RA)_{\sigma_2}$ , where  $\sigma_1$  and  $\sigma_2$  are auxiliary states defined in the theorem and  $\epsilon_{\tilde{S}}, \delta_{\tilde{S}} \searrow 0$  as  $\delta \searrow 0$ .*

*Proof.* The proof is provided in Appendix B-C. □

**Remark 7.** The terms corresponding to the operators that complete the sub-POVMs  $M_A^{(n,\mu_1)}$  and  $M_B^{(n,\mu_2)}$ , i.e.,  $I - \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\mu_1)}$  and  $I - \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\mu_2)}$  are taken care in  $\tilde{T}$ . The expression  $T$  excludes the completing operators. Therefore, we use  $A_{u^n}^{(\mu_1)}$  and  $B_{v^n}^{(\mu_2)}$  to denote the operators corresponding to  $u^n \in \mathcal{T}_\delta^{(n)}(U)$  and  $v^n \in \mathcal{T}_\delta^{(n)}(V)$ , respectively.

### Step 2: Isolating the effect of error induced by binning

Recall the definition of  $e^{(\mu)}(u^n, v^n)$  as  $e^{(\mu)}(u^n, v^n) \triangleq F^{(\mu)}(i, j)$ , for each  $(u^n, v^n) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j)$  and  $(u^n, v^n) \in \mathcal{C}^{(\mu)}$ . For any  $(u^n, v^n) \notin \mathcal{C}^{(\mu)}$  let  $e^{(\mu)}(u^n, v^n) = (u_0^n, v_0^n)$ . This simplifies  $T$  as

$$\begin{aligned} T &= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{\substack{i>0, \\ j>0}} \sqrt{\rho_{AB}^{\otimes n}} \left( \sum_{u^n \in \mathcal{B}_1^{(\mu_1)}(i)} A_{u^n}^{(\mu_1)} \otimes \sum_{v^n \in \mathcal{B}_2^{(\mu_2)}(j)} B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | F^{(\mu_1, \mu_2)}(i, j)) \\ &= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \sum_{\substack{i>0, \\ j>0}} \mathbb{1}_{\{u^n \in \mathcal{B}_1^{(\mu_1)}(i), v^n \in \mathcal{B}_2^{(\mu_2)}(j)\}} P_{Z|U,V}^n(z^n | e^{(\mu)}(u^n, v^n)) \\ &= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | e^{(\mu)}(u^n, v^n)), \end{aligned}$$

where we have used the fact that  $\sum_{u^n \in \mathcal{B}_1^{(\mu_1)}(i)} A_{u^n}^{(\mu_1)} = \sum_{u^n} A_{u^n}^{(\mu_1)} \mathbb{1}_{\{u^n \in \mathcal{B}_1^{(\mu_1)}(i)\}}$  and  $\sum_{i>0} \mathbb{1}_{\{u^n \in \mathcal{B}_1^{(\mu_1)}(i)\}} = 1$  for all  $u^n \in \mathcal{T}_\delta^{(n)}(U)$ , and similar holds for the POVM  $\{B_{v^n}^{(\mu_2)}\}$ . Note that the  $(u^n, v^n)$  that appear in the above summation is confined to  $(\mathcal{T}_\delta^{(n)}(U) \times \mathcal{T}_\delta^{(n)}(V))$ , however for ease of notation, we do not make this explicit. We substitute the above expression into  $S$  as in (42) to obtain

$$S = \sum_{z^n} \left\| \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | e^{(\mu)}(u^n, v^n)) \right\|_1.$$

We add and subtract an appropriate term within  $S$  and apply triangle inequality to isolate the effect of binning as  $S \leq S_1 + S_2$ , where

$$\begin{aligned} S_1 &\triangleq \sum_{z^n} \left\| \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | u^n, v^n) \right\|_1 \quad \text{and} \\ S_2 &\triangleq \sum_{z^n} \left\| \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \left( P_{Z|U,V}^n(z^n | u^n, v^n) - P_{Z|U,V}^n(z^n | e^{(\mu)}(u^n, v^n)) \right) \right\|_1. \end{aligned} \tag{43}$$

Note that the term  $S_1$  characterizes the error introduced by approximation of the original POVM with the collection of approximating sub-POVMs  $M_1^{(n,\mu_1)}$  and  $M_2^{(n,\mu_2)}$ , and the term  $S_2$  characterizes the error caused by binning of these approximating sub-POVMs. In this step, we analyze  $S_2$  and prove the following proposition.

**Proposition 4 (Mutual Packing).** *There exist functions  $\epsilon_{S_2}(\delta)$  and  $\delta_{S_2}(\delta)$ , such that for all sufficiently small  $\delta$  and sufficiently large  $n$ , we have  $\mathbb{P}(S_2 > \epsilon_{S_2}(\delta)) \leq \delta_{S_2}(\delta)$ , if  $\tilde{R}_1 + \tilde{R}_2 - R_1 - R_2 < I(U; V)_{\sigma_3}$ , where  $\sigma_3$  is the auxiliary state defined in the theorem and  $\epsilon_{S_2}, \delta_{S_2} \searrow 0$  as  $\delta \searrow 0$ .*

*Proof.* The proof is provided in Appendix B-D □

### Step 3: Isolating the effect of Alice's approximating measurement

In this step, we separately analyze the effect of approximating measurements at the two distributed parties in the term  $S_1$ . For that, we split  $S_1$  as  $S_1 \leq Q_1 + Q_2$ , where

$$Q_1 \triangleq \sum_{z^n} \left\| \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} A_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | u^n, v^n) \right\|_1,$$

$$Q_2 \triangleq \sum_{z^n} \left\| \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( A_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B - \frac{1}{N_2} \sum_{\mu_2=1}^{N_2} A_{u^n}^{(\mu_2)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1.$$

With this partition, the terms within the trace norm of  $Q_1$  differ only in the action of Alice's measurement. And similarly, the terms within the norm of  $Q_2$  differ only in the action of Bob's measurement. Showing that these two terms are small forms a major portion of the achievability proof.

**Analysis of  $Q_1$ :** To show  $Q_1$  is small, we compute rate constraints which ensure that an upper bound to  $Q_1$  can be made to vanish in an expected sense. Furthermore, this upper bound becomes convenient in obtaining a single-letter characterization for the rate needed to make the term corresponding to  $Q_2$  vanish. For this, we define  $J$  as

$$J \triangleq \sum_{z^n, v^n} \left\| \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} A_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1. \quad (44)$$

By defining  $J$  and using triangle inequality for block operators (which holds with equality), we add the sub-system  $V$  to  $RZ$ , resulting in the joint system  $RZV$ , corresponding to the state  $\sigma_3$  as defined in the theorem. Then we approximate the joint system  $RZV$  using an approximating sub-POVM  $M_A^{(n)}$  producing outputs on the alphabet  $\mathcal{U}^n$ . To make  $J$  small for sufficiently large  $n$ , we expect the sum of the rate of the approximating sub-POVM and common randomness, i.e.,  $\tilde{R}_1 + C_1$ , to be larger than  $I(U; RZV)_{\sigma_3}$ . We seek to prove this in the following.

Note that from triangle inequality, we have  $Q_1 \leq J$ . Further, we add and subtract an appropriate term within  $J$  and use triangle inequality obtain  $J \leq J_1 + J_2$ , where

$$J_1 \triangleq \sum_{z^n, v^n} \left\| \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \frac{\gamma_{u^n}^{(\mu_1)}}{\lambda_{u^n}^A} \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1 \quad \text{and}$$

$$J_2 \triangleq \sum_{z^n, v^n} \left\| \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \frac{\gamma_{u^n}^{(\mu_1)}}{\lambda_{u^n}^A} \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} A_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1.$$

Now with the intention of employing Lemma 7, we express  $J_1$  as

$$J_1 = \left\| \sum_{z^n, u^n, v^n} \lambda_{u^n, v^n}^{AB} \hat{\rho}_{u^n, v^n}^{AB} \otimes P_{Z|U,V}^n(z^n|u^n, v^n) |v^n\rangle\langle v^n| \otimes |z^n\rangle\langle z^n| \right. \\ \left. - \frac{(1-\varepsilon)}{(1+\eta)} \frac{1}{2^{n(\tilde{R}_1+C_1)}} \sum_{\mu_1, l} \sum_{z^n, u^n, v^n} \mathbb{1}_{\{U^{n, (\mu_1)}(l)=u^n\}} \frac{\lambda_{u^n, v^n}^{AB}}{\lambda_{u^n}^A} \hat{\rho}_{u^n, v^n}^{AB} \otimes P_{Z|U,V}^n(z^n|u^n, v^n) |v^n\rangle\langle v^n| \otimes |z^n\rangle\langle z^n| \right\|_1,$$

where the equality above is obtained by using the definitions of  $\gamma_{u^n}^{(\mu_1)}$  and  $\hat{\rho}_{u^n, v^n}^{AB}$ , followed by using the triangle inequality for the block diagonal operators, which in fact becomes an equality.

Let us define  $\mathcal{T}_{u^n}$  as

$$\mathcal{T}_{u^n} = \sum_{z^n, v^n} \frac{\lambda_{u^n, v^n}^{AB}}{\lambda_{u^n}^A} \hat{\rho}_{u^n, v^n}^{AB} \otimes P_{Z|U,V}^n(z^n|u^n, v^n) |v^n\rangle\langle v^n| \otimes |z^n\rangle\langle z^n|.$$

Note that the above definition of  $\mathcal{T}_{u^n}$  contains all the elements in product form, and thus it can be written as  $\mathcal{T}_{u^n} = \bigotimes_{i=1}^n \mathcal{T}_{u_i}$ . This simplifies  $J_1$  as

$$J_1 = \left\| \sum_{u^n} \lambda_{u^n}^A \mathcal{T}_{u^n} - \frac{(1-\varepsilon)}{(1+\eta)} \frac{1}{2^{n(\tilde{R}_1+C_1)}} \sum_{\mu_1, l} \mathcal{T}_{U^{n, (\mu_1)}(l)} \right\|_1.$$

Now, using Lemma 7 we get the following bound. For any  $\epsilon_{J_1}, \delta_{J_1} \in (0, 1)$ , if

$$\tilde{R}_1 + C_1 > S \left( \sum_{u \in \mathcal{U}} \lambda_u^A \mathcal{T}_u \right) + \sum_{u \in \mathcal{U}} \lambda_u^A S(\mathcal{T}_u) = I(U; RZV)_{\sigma_3}, \quad (45)$$

then  $\mathbb{P}(J_1 \geq \epsilon_{J_1}) \leq \delta_{J_1}$  for sufficiently large  $n$ , where  $\sigma_3 = \sum_{u \in \mathcal{U}} \lambda_u^A \mathcal{T}_u \otimes |u\rangle\langle u|$ .

Now, we consider the term corresponding to  $J_2$  and prove that its expectation with respect to the Alice's codebook is small. Recalling  $J_2$ , we get

$$\begin{aligned} J_2 &\leq \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sum_{u^n, v^n} \sum_{z^n} P_{Z|U,V}^n(z^n|u^n, v^n) \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \frac{\gamma_{u^n}^{(\mu_1)}}{\lambda_{u^n}^A} \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - A_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1, \\ &= \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sum_{u^n, v^n} \gamma_{u^n}^{(\mu_1)} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \left( \frac{1}{\lambda_{u^n}^A} \bar{\Lambda}_{u^n}^A - \sqrt{\rho_A^{\otimes n}}^{-1} \Lambda_{u^n}^A \sqrt{\rho_A^{\otimes n}}^{-1} \right) \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1, \end{aligned}$$

where the inequality is obtained by using triangle and the next equality follows from the fact that  $\sum_{z^n} P_{Z|U,V}^n(z^n|u^n, v^n) = 1$  for all  $u^n \in \mathcal{U}^n$  and  $v^n \in \mathcal{V}^n$  and using the definition of  $A_{u^n}^{(\mu_1)}$ . By applying expectation of  $J_2$  over the Alice's codebook, we get

$$\mathbb{E}[J_2] \leq \frac{1}{(1+\eta)} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A \sum_{v^n} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \left( \frac{1}{\lambda_{u^n}^A} \bar{\Lambda}_{u^n}^A - \sqrt{\rho_A^{\otimes n}}^{-1} \Lambda_{u^n}^A \sqrt{\rho_A^{\otimes n}}^{-1} \right) \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1,$$

where we have used the fact that  $\mathbb{E}[\gamma_{u^n}^{(\mu_1)}] = \frac{\lambda_{u^n}^A}{(1+\eta)}$ . To simplify the above equation, we employ Lemma 3 from Section IV-D that completely discards the effect of Bob's measurement. Since  $\sum_{v^n} \bar{\Lambda}_{v^n}^B = I$ , from Lemma 3 we have for every  $u^n \in \mathcal{T}_\delta^{(n)}(A)$ ,

$$\begin{aligned} &\sum_{v^n} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \left( \frac{1}{\lambda_{u^n}^A} \bar{\Lambda}_{u^n}^A - \sqrt{\rho_A^{\otimes n}}^{-1} \Lambda_{u^n}^A \sqrt{\rho_A^{\otimes n}}^{-1} \right) \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\ &= \left\| \sqrt{\rho_A^{\otimes n}} \left( \frac{1}{\lambda_{u^n}^A} \bar{\Lambda}_{u^n}^A - \sqrt{\rho_A^{\otimes n}}^{-1} \Lambda_{u^n}^A \sqrt{\rho_A^{\otimes n}}^{-1} \right) \sqrt{\rho_A^{\otimes n}} \right\|_1. \end{aligned}$$

This simplifies  $\mathbb{E}[J_2]$  as

$$\begin{aligned} \mathbb{E}[J_2] &\leq \frac{1}{(1+\eta)} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A \left\| \sqrt{\rho_A^{\otimes n}} \left( \frac{1}{\lambda_{u^n}^A} \bar{\Lambda}_{u^n}^A - \sqrt{\rho_A^{\otimes n}}^{-1} \Lambda_{u^n}^A \sqrt{\rho_A^{\otimes n}}^{-1} \right) \sqrt{\rho_A^{\otimes n}} \right\|_1 \\ &= \frac{1}{(1+\eta)} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A \|\hat{\rho}_{u^n}^A - \Lambda_{u^n}^A\|_1 \leq \frac{(1-\varepsilon)}{(1+\eta)} (2\sqrt{\varepsilon'_A} + 2\sqrt{\varepsilon''_A}) = \epsilon_{J_2}, \end{aligned}$$

where the last inequality is obtained by the repeated usage of the average gentle measurement lemma by setting  $\epsilon_{J_2} = \frac{(1-\varepsilon)}{(1+\eta)} (2\sqrt{\varepsilon'_A} + 2\sqrt{\varepsilon''_A})$  with  $\epsilon_{J_2} \searrow 0$  as  $\delta \searrow 0$  and  $\varepsilon'_A = \varepsilon + 2\sqrt{\varepsilon}$  and  $\varepsilon''_A = 2\varepsilon + 2\sqrt{\varepsilon}$  ( see (35) in [2] for details). Since  $Q_1 \leq J \leq J_1 + J_2$ , hence  $J$ , and consequently  $Q_1$ , can be made arbitrarily small for sufficiently large  $n$ , if  $\tilde{R}_1 + C_1 > I(U; RZV)_{\sigma_3}$ . Now we move on to bounding  $Q_2$ .

#### Step 4: Analyzing the effect of Bob's approximating measurement

Step 3 ensured that the sub-system  $RZV$  is close to a tensor product state in trace-norm. In this step, we approximate the state corresponding to the sub-system  $RZ$  using the approximating POVM  $M_B^{(n)}$ , producing outputs on the alphabet  $\mathcal{V}^n$ . We proceed with the following proposition.

**Proposition 5.** *There exist functions  $\epsilon_{Q_2}(\delta)$  and  $\delta_{Q_2}(\delta)$ , such that for all sufficiently small  $\delta$  and sufficiently large  $n$ , we have  $\mathbb{P}(Q_2 > \epsilon_{Q_2}(\delta)) \leq \delta_{Q_2}(\delta)$ , if  $\tilde{R}_2 + C_2 > I(V; RZ)_{\sigma_3}$ , where  $\sigma_3$  is the auxiliary state defined in the theorem and  $\epsilon_{Q_2}, \delta_{Q_2} \searrow 0$  as  $\delta \searrow 0$ .*

*Proof.* The proof is provided in Appendix B-E. □

### C. Rate Constraints

To sum-up, we showed that the trace distance inequality in (40) holds for sufficiently large  $n$  and with probability sufficiently close to 1, if the following bounds hold:

$$\tilde{R}_1 \geq I(U; RB)_{\sigma_1}, \quad (46a)$$

$$\tilde{R}_2 \geq I(V; RA)_{\sigma_2}, \quad (46b)$$

$$\tilde{R}_1 + C_1 \geq I(U; RZV)_{\sigma_3}, \quad (46c)$$

$$\tilde{R}_2 + C_2 \geq I(V; RZ)_{\sigma_3}, \quad (46d)$$

$$(\tilde{R}_1 - R_1) + (\tilde{R}_2 - R_2) < I(U; V)_{\sigma_3}, \quad (46e)$$

$$\tilde{R}_1 \geq R_1 \geq 0, \quad \tilde{R}_2 \geq R_2 \geq 0, \quad (46f)$$

$$C_1 + C_2 \leq C, \quad C \geq 0, \quad (46g)$$

where  $C \triangleq \frac{1}{n} \log_2 N$ . 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  $\epsilon$ -faithful to  $M_{AB}^{\otimes n}$  with respect to  $\rho_{AB}^{\otimes n}$ . Lastly, we complete the proof of the theorem using the following lemma.

**Lemma 8.** *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 septuple  $(R_1, R_2, C, \tilde{R}_1, \tilde{R}_2, \tilde{C}_1, \tilde{C}_2)$  satisfies the inequalities in (46). Let,  $\mathcal{R}_2$  denote the set of all triples  $(R_1, R_2, C)$  that satisfies the inequalities in (38) given in the statement of the theorem. Then,  $\mathcal{R}_1 = \mathcal{R}_2$ .*

*Proof.* This follows from Fourier-Motzkin elimination [17]. □

## IX. CONCLUSION

We have developed a distributed measurement compression protocol where we introduced the technique of mutual covering and random binning of distributed measurements. Using these techniques, a set of communication rate-pairs and common randomness rate is characterized for faithful simulation of distributed measurements. We further developed an approach for a distributed quantum-to-classical rate-distortion theory, and provided single-letter inner and outer bounds. As a part of future work, we intend to improve the outer bound by providing a dimensionality bound on the auxiliary Hilbert space involved in the expression. Further, we also desire to improve the achievable rate region by using structured POVMs based on algebraic codes.

**Acknowledgement:** We thank Mark Wilde for his valuable inputs on techniques needed to prove Theorem 5 and for referring us to the additional work performed in [12] and [22]. We are also grateful to Arun Padakandla for his inputs on the classical analogue of the current work [23], which was very helpful in developing the proof techniques here.

## APPENDIX A PROOF OF LEMMAS

### A. Proof of Lemma 3

Consider the LHS of (21). We define an operator  $\Lambda_{y_0}$  which completes the sub-POVM  $\{\Lambda_y\}_{y \in \mathcal{Y}}$  as  $\Lambda_{y_0} \triangleq I - \sum_{y \in \mathcal{Y}} \Lambda_y$ . Further, let the set  $\mathcal{Y}^+ \triangleq \mathcal{Y} \cup \{y_0\}$ . Since trace norm is invariant to transposition with respect to  $\rho_{AB}$ , we can write for any  $y \in \mathcal{Y}^+$ ,

$$\|\sqrt{\rho_{AB}} (\Gamma^A \otimes \Lambda_y^B) \sqrt{\rho_{AB}}\|_1 = \left\| \left[ \sqrt{\rho_{AB}} (\Gamma^A \otimes \Lambda_y^B) \sqrt{\rho_{AB}} \right]^T \right\|_1 = \left\| \sqrt{\rho_{AB}} \left( (\Gamma^A)^T \otimes (\Lambda_y^B)^T \right) \sqrt{\rho_{AB}} \right\|_1.$$

One can easily prove for any  $\Gamma_A$  (not necessarily positive) that

$$\left(\sqrt{\rho_{AB}} \left( (\Gamma^A)^T \otimes (\Lambda_y^B)^T \right) \sqrt{\rho_{AB}}\right)^R = \text{Tr}_{AB} \left\{ (\text{id} \otimes \Gamma^A \otimes \Lambda_y^B) \Psi_{RAB} \right\}, \quad (47)$$

where  $\Psi_{RAB}$  is the canonical purification of  $\rho_{AB}$  defined as  $\Psi_{RAB} \triangleq \sum_{x,x'} \sqrt{\lambda_x \lambda_{x'}} |x\rangle\langle x|_{AB} \otimes |x'\rangle\langle x'|_R$  for the spectral decomposition of  $\rho_{AB}$  given as  $\rho_{AB} = \sum_x \lambda_x |x\rangle\langle x|_{AB}$ . Now, using (47) we perform the following simplification

$$\begin{aligned} \sum_{y \in \mathcal{Y}} \left\| \sqrt{\rho_{AB}} (\Gamma^A \otimes \Lambda_y^B) \sqrt{\rho_{AB}} \right\|_1 &\leq \sum_{y \in \mathcal{Y}^+} \left\| \sqrt{\rho_{AB}} (\Gamma^A \otimes \Lambda_y^B) \sqrt{\rho_{AB}} \right\|_1 \\ &= \sum_{y \in \mathcal{Y}^+} \left\| \text{Tr}_{AB} \left\{ (\text{id}_R \otimes \Gamma^A \otimes \Lambda_y^B) \Psi_{RAB} \right\} \right\|_1 \\ &= \left\| \sum_{y \in \mathcal{Y}^+} \text{Tr}_{AB} \left\{ (\text{id}_{RB} \otimes \Gamma^A) (\text{id}_{RA} \otimes \Lambda_y^B) \Psi_{RAB} \otimes |y\rangle\langle y| \right\} \right\|_1 \\ &= \left\| \text{Tr}_A \left\{ (\text{id}_{RY} \otimes \Gamma^A) \left( \sum_{y \in \mathcal{Y}^+} |y\rangle\langle y| \otimes \text{Tr}_B \left\{ (\text{id}_{RA} \otimes \Lambda_y^B) \Psi_{RAB} \right\} \right) \right\} \right\|_1 \\ &= \left\| \text{Tr}_A \left\{ (\text{id}_{RY} \otimes \Gamma^A) \sigma_{RAY} \right\} \right\|_1 \\ &= \left\| \text{Tr}_{AZ} \left\{ (\text{id}_{RY} \otimes \Gamma^A \otimes \text{id}_Z) \Phi_{RAYZ}^{\sigma_{RAY}} \right\} \right\|_1, \end{aligned} \quad (48)$$

where the second equality uses the triangle inequality for block diagonal operators, the third equality first uses the property that  $\text{Tr}_{XY} \{ \cdot \} = \text{Tr}_X \{ \text{Tr}_Y \{ \cdot \} \}$ , followed by the definition of partial trace and its linearity, the fourth equality uses  $\sigma_{RAY}$  defined as

$$\sigma_{RAY} = \sum_{y \in \mathcal{Y}^+} |y\rangle\langle y| \otimes \text{Tr}_B \left\{ (\text{id}_{RA} \otimes \Lambda_y^B) \Psi_{RAB} \right\},$$

and finally, the last one uses  $\Phi_{RAYZ}^{\sigma_{RAY}}$  defined as the canonical purification of  $\sigma_{RAY}$ . Note that the above inequality becomes an equality when  $\sum_{y \in \mathcal{Y}} \Lambda_y = I$ . Using similar sequence of arguments as used in (47), we have

$$\left\| \text{Tr}_{AZ} \left\{ (\text{id}_{RY} \otimes \Gamma^A \otimes \text{id}_Z) \Phi_{RAYZ}^{\sigma_{RAY}} \right\} \right\|_1 = \left\| \sqrt{\text{Tr}_{RYZ} \left\{ \Phi_{RAYZ}^{\sigma_{RAY}} \right\}} \Gamma^A \sqrt{\text{Tr}_{RYZ} \left\{ \Phi_{RAYZ}^{\sigma_{RAY}} \right\}} \right\|_1 = \left\| \sqrt{\rho_A} \Lambda_x^A \sqrt{\rho_A} \right\|_1.$$

This completes the proof.

### B. Proof of Lemma 4

Let the operators of  $\hat{M}_X$  and  $\hat{M}_Y$  be denoted by  $\{\hat{\Lambda}_i^X\}_{i \in \mathcal{I}}$  and  $\{\hat{\Lambda}_j^Y\}_{j \in \mathcal{J}}$ , respectively, and let the operators of  $M_X$  and  $M_Y$  be denoted by  $\{\Lambda_i^X\}$  and  $\{\Lambda_j^Y\}$ , respectively, for some finite sets  $\mathcal{I}$  and  $\mathcal{J}$ . With this notation, we need to show the following inequality

$$G \triangleq \sum_{i,j} \left\| \sqrt{\rho_{XY}} (\Lambda_i^X \otimes \Lambda_j^Y - \hat{\Lambda}_i^X \otimes \hat{\Lambda}_j^Y) \sqrt{\rho_{XY}} \right\|_1 + \text{Tr} \left\{ \left( I - \sum_{i,j} \hat{\Lambda}_i^X \otimes \hat{\Lambda}_j^Y \right) \rho_{XY} \right\} \leq 2\epsilon,$$

where  $\Psi_{RXY}^\rho$  is a purification of  $\rho_{XY}$ . Next, by adding and subtracting appropriate terms, we get

$$\begin{aligned} G &\leq \sum_{i,j} \left\| \sqrt{\rho_{XY}} (\Lambda_i^X \otimes \Lambda_j^Y - \hat{\Lambda}_i^X \otimes \Lambda_j^Y) \sqrt{\rho_{XY}} \right\|_1 + \text{Tr} \left\{ \left( I - \sum_i \hat{\Lambda}_i^X \right) \rho_X \right\} \\ &\quad + \sum_{i,j} \left\| \sqrt{\rho_{XY}} (\hat{\Lambda}_i^X \otimes \Lambda_j^Y - \hat{\Lambda}_i^X \otimes \hat{\Lambda}_j^Y) \sqrt{\rho_{XY}} \right\|_1 + \text{Tr} \left\{ \left( I - \sum_j \hat{\Lambda}_j^Y \right) \rho_Y \right\} \\ &\quad + \text{Tr} \left\{ \left( I - \sum_{i,j} \hat{\Lambda}_i^X \otimes \hat{\Lambda}_j^Y \right) \rho_{XY} \right\} - \text{Tr} \left\{ \left( I - \sum_i \hat{\Lambda}_i^X \right) \rho_X \right\} - \text{Tr} \left\{ \left( I - \sum_j \hat{\Lambda}_j^Y \right) \rho_Y \right\} \end{aligned}$$



$$\begin{aligned}
&\leq \sum_i \left\| \sqrt{\rho_X} (\Lambda_i^X - \hat{\Lambda}_i^X) \sqrt{\rho_X} \right\|_1 + \text{Tr} \left\{ \left( I - \sum_i \hat{\Lambda}_i^X \right) \rho_X \right\} \\
&\quad + \sum_j \left\| \sqrt{\rho_Y} (\Lambda_j^Y - \hat{\Lambda}_j^Y) \sqrt{\rho_Y} \right\|_1 + \text{Tr} \left\{ \left( I - \sum_j \hat{\Lambda}_j^Y \right) \rho_Y \right\} \\
&\quad + \text{Tr} \left\{ \left( I - \sum_{i,j} \hat{\Lambda}_i^X \otimes \hat{\Lambda}_j^Y \right) \rho_{XY} \right\} - \text{Tr} \left\{ \left( I - \sum_i \hat{\Lambda}_i^X \right) \rho_X \right\} - \text{Tr} \left\{ \left( I - \sum_j \hat{\Lambda}_j^Y \right) \rho_Y \right\} \\
&\leq 2\epsilon + \text{Tr} \left\{ \left( \sum_i \hat{\Lambda}_i^X \otimes \left( I - \sum_j \hat{\Lambda}_j^Y \right) \right) \rho_{XY} \right\} - \text{Tr} \left\{ \left( I - \sum_j \hat{\Lambda}_j^Y \right) \rho_Y \right\} \leq 2\epsilon,
\end{aligned}$$

where the second inequality follows by applying Lemma 3 twice, the third inequality follows from the hypotheses of the lemma, and the final inequality uses the fact that  $\hat{M}^X$  and  $\hat{M}^Y$  are sub-POVMs. This completes the proof of the lemma.

### C. Proof of Lemma 7

*Proof.* Consider the trace norm expression given in (34). This expression can be upper bounded using triangle inequality as

$$\begin{aligned}
&\left\| \sum_{w^n} \lambda_{w^n} \mathcal{T}_{w^n} - \frac{1}{2^{n(R+C)}} \frac{(1-\epsilon)}{(1+\eta)} \sum_{l,\mu} \mathcal{T}_{W^{n,(\mu)}(l)} \right\|_1 \\
&\leq \left\| \sum_{w^n} \lambda_{w^n} \mathcal{T}_{w^n} - \frac{(1-\epsilon)}{(1+\eta)} \sum_{\substack{w^n \in \\ \mathcal{T}_\delta^{(n)}(W)}} \tilde{P}_{W^n}(w^n) \mathcal{T}_{w^n} \right\|_1 + \frac{(1-\epsilon)}{(1+\eta)} \left\| \sum_{\substack{w^n \in \\ \mathcal{T}_\delta^{(n)}(W)}} \tilde{P}_{W^n}(w^n) \mathcal{T}_{w^n} - \frac{1}{2^{n(R+C)}} \sum_{l,\mu} \mathcal{T}_{W^{n,(\mu)}(l)} \right\|_1. \quad (49)
\end{aligned}$$

The first term in the right-hand side is bounded from above as

$$\begin{aligned}
\left\| \sum_{w^n} \lambda_{w^n} \mathcal{T}_{w^n} - \frac{(1-\epsilon)}{(1+\eta)} \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \tilde{P}_{W^n}(w^n) \mathcal{T}_{w^n} \right\|_1 &\leq \left\| \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \left( 1 - \frac{1}{(1+\eta)} \right) \mathcal{T}_{w^n} \right\|_1 + \left\| \sum_{w^n \notin \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \mathcal{T}_{w^n} \right\|_1 \\
&\leq \left( \frac{\eta}{1+\eta} \right) \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \underbrace{\|\mathcal{T}_{w^n}\|_1}_{=1} + \sum_{w^n \notin \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \underbrace{\|\mathcal{T}_{w^n}\|_1}_{=1} \\
&\leq \left( \frac{\eta}{1+\eta} \right) + \epsilon \leq \eta + \epsilon = \epsilon_1, \quad (50)
\end{aligned}$$

where  $\epsilon_1$  can be made arbitrarily small for all sufficiently large  $n$ . Now consider the second term in (49). Using covering lemma from [16], this can be bounded as follows. For  $w^n \in \mathcal{T}_\delta^{(n)}(W)$ , let  $\Pi$  and  $\Pi_{w^n}$  denote the projectors onto the typical subspace of  $\mathcal{T}^{\otimes n}$  and  $\mathcal{T}_{w^n}$ , respectively, where  $\mathcal{T} = \sum_{w^n} \lambda_{w^n} \mathcal{T}_{w^n}$ . From the definition of typical projectors, for any  $\epsilon_1 \in (0, 1)$  we have for sufficiently large  $n$ , the following inequalities satisfied for all  $w^n \in \mathcal{T}_\delta^{(n)}(W)$ :

$$\begin{aligned}
\text{Tr}\{\Pi \mathcal{T}_{w^n}\} &\geq 1 - \epsilon_1, \\
\text{Tr}\{\Pi_{w^n} \mathcal{T}_{w^n}\} &\geq 1 - \epsilon_1, \\
\text{Tr}\{\Pi\} &\leq D, \\
\Pi_{w^n} \mathcal{T}_{w^n} \Pi_{w^n} &\leq \frac{1}{d} \Pi_{w^n}, \quad (51)
\end{aligned}$$

where  $D = 2^{n(S(\mathcal{T}) + \delta_1)}$  and  $d = 2^{n[(\sum_w \lambda_w S(\mathcal{T}_w)) + \delta_2]}$ , and  $\delta_1 \searrow 0, \delta_2 \searrow 0$  as  $\epsilon_1 \searrow 0$ . From the statement of the covering lemma, we know that for an ensemble  $\{\tilde{P}_{W^n}(w^n), \mathcal{T}_{w^n}\}_{w^n \in \mathcal{W}^n}$ , if there exists projectors  $\Pi$  and  $\Pi_{w^n}$  such

that they satisfy the set of inequalities in (51), then for any  $\epsilon_2 > 0$ , sufficiently large  $n$  and  $n(R + C) > \log_2 \frac{D}{d}$ , the obfuscation error, defined as

$$\left\| \sum_{w^n} \tilde{P}_W^n(w^n) \mathcal{T}_{w^n} - \frac{1}{2^{R+C}} \sum_{l, \mu} \mathcal{T}_{W^{n, (\mu)}(l)} \right\|_1,$$

can be made smaller than  $\epsilon_2$  with high probability. This gives us the the following rate constraints  $R + C \geq S(\sum_w \lambda_w \mathcal{T}_w) - \sum_w \lambda_w S(\mathcal{T}_w) + \delta' = \chi(\{\lambda_w\}, \{\hat{\rho}_w \otimes \sigma_w\}) + \delta'$ . Using this constraint and the bound from (50), the result follows.  $\square$

## APPENDIX B PROOF OF PROPOSITIONS

### A. Proof of Proposition 1

The second term in the trace distance in  $S_2$  can be expressed as

$$\begin{aligned} (\text{id} \otimes \tilde{M}_{AB}^{(n)}) (\Psi_{R^n A^n B^n}^\rho) &= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{i, j} \Phi_{F^{(\mu)}(i, j)} \otimes \text{Tr}_{AB} \left\{ (\text{id} \otimes \Gamma_i^{A, (\mu_1)} \otimes \Gamma_j^{B, (\mu_2)}) \Psi_{R^n A^n B^n}^\rho \right\} \\ &= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{i, j \geq 1} \sum_{(u^n, v^n) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j)} \Phi_{e^{(\mu)}(u^n, v^n)} \otimes \text{Tr}_{AB} \left\{ (\text{id} \otimes A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)}) \Psi_{R^n A^n B^n}^\rho \right\} \\ &+ \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{j \geq 1} \sum_{v^n \in \mathcal{B}_2^{(\mu_2)}(j)} \Phi_{(u_0^n, v_0^n)} \otimes \text{Tr}_{AB} \left\{ (\text{id} \otimes (I - \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\mu_1)}) \otimes B_{v^n}^{(\mu_2)}) \Psi_{R^n A^n B^n}^\rho \right\} \\ &+ \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{i \geq 1} \sum_{u^n \in \mathcal{B}_1^{(\mu_1)}(i)} \Phi_{(u_0^n, v_0^n)} \otimes \text{Tr}_{AB} \left\{ (\text{id} \otimes A_{u^n}^{(\mu_1)} \otimes (I - \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\mu_2)})) \Psi_{R^n A^n B^n}^\rho \right\} \\ &+ \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \Phi_{(u_0^n, v_0^n)} \otimes \text{Tr}_{AB} \left\{ (\text{id} \otimes (I - \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\mu_1)}) \otimes (I - \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\mu_2)})) \Psi_{R^n A^n B^n}^\rho \right\}. \quad (52) \end{aligned}$$

Similarly, for the first term in the trace distance in  $S_2$ , we have

$$\begin{aligned} &\frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} (\text{id} \otimes [M_1^{(n, \mu_1)}] \otimes [M_2^{(n, \mu_2)}]) (\Psi_{R^n A^n B^n}^\rho) \\ &= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \Phi_{(u^n, v^n)} \otimes \text{Tr}_{AB} \left\{ (\text{id} \otimes A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)}) \Psi_{R^n A^n B^n}^\rho \right\} \\ &+ \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \Phi_{(u_0^n, v_0^n)} \otimes \text{Tr}_{AB} \left\{ (\text{id} \otimes (I - \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\mu_1)}) \otimes B_{v^n}^{(\mu_2)}) \Psi_{R^n A^n B^n}^\rho \right\} \\ &+ \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \Phi_{(u_0^n, v_0^n)} \otimes \text{Tr}_{AB} \left\{ (\text{id} \otimes A_{u^n}^{(\mu_1)} \otimes (I - \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\mu_2)})) \Psi_{R^n A^n B^n}^\rho \right\} \\ &+ \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \Phi_{(u_0^n, v_0^n)} \otimes \text{Tr}_{AB} \left\{ (\text{id} \otimes (I - \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\mu_1)}) \otimes (I - \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\mu_2)})) \Psi_{R^n A^n B^n}^\rho \right\}. \quad (53) \end{aligned}$$

By replacing the terms in  $S_2$  using the corresponding expansions from (52) and (53), we observe that the second, third and fourth terms on the right hand side of (52) get canceled with the corresponding terms on the right hand side of (53). This simplifies  $S_2$  as

$$\begin{aligned} &\frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \left\| (\Phi_{(u^n, v^n)} - \Phi_{e^{(\mu)}(u^n, v^n)}) \otimes \text{Tr}_{AB} \left\{ (\text{id} \otimes A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)}) \Psi_{R^n A^n B^n}^\rho \right\} \right\|_1 \\ &= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \left\| \Phi_{(u^n, v^n)} - \Phi_{e^{(\mu)}(u^n, v^n)} \right\|_1 \times \left\| \text{Tr}_{AB} \left\{ (\text{id} \otimes A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)}) \Psi_{R^n A^n B^n}^\rho \right\} \right\|_1 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \|\Phi_{(u^n, v^n)} - \Phi_{e^{(\mu)}(u^n, v^n)}\|_1 \times \text{Tr} \left\{ (\text{id} \otimes A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)}) \Psi_{R^n A^n B^n}^\rho \right\} \\
&= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \|\Phi_{u^n, v^n} - \Phi_{e^{(\mu)}(u^n, v^n)}\|_1 \gamma_{u^n}^{(\mu_1)} \zeta_{v^n}^{(\mu_2)} \Omega_{u^n, v^n},
\end{aligned}$$

where the first two equalities are obtained by using the definition of trace norm and the last equality follows from the definition of  $A_{u^n}^{(\mu_1)}$  and  $B_{v^n}^{(\mu_2)}$  as in (12), with  $\Omega_{u^n, v^n} \triangleq \text{Tr} \left\{ \sqrt{\rho_A^{\otimes n} \otimes \rho_B^{\otimes n}}^{-1} (\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B) \sqrt{\rho_A^{\otimes n} \otimes \rho_B^{\otimes n}}^{-1} \rho_{AB}^{\otimes n} \right\}$ . This completes the proof.

### B. Proof of Proposition 2

From Proposition 1, and using the definitions  $\gamma_{u^n}^{(\mu_1)}$  and  $\zeta_{v^n}^{(\mu_2)}$ ,  $S_2$  can be simplified as

$$S_2 = \frac{(1-\varepsilon)(1-\varepsilon')}{(1+\eta)^2 2^{n(\tilde{R}_1 + \tilde{R}_2)}} \sum_{u^n, v^n} \sum_{\mu_1, \mu_2} \sum_{l, k} \mathbb{1}\{U^{n, (\mu_1)}(l) = u^n, V^{n, (\mu_2)}(k) = v^n\} \Omega_{u^n, v^n} \|\Phi_{u^n, v^n} - \Phi_{e^{(\mu)}(u^n, v^n)}\|_1.$$

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

$$\begin{aligned}
&2 \times \mathbb{1} \left\{ \exists (\tilde{u}^n, \tilde{v}^n, i, j) : (u^n, v^n) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j), \right. \\
&\quad \left. (\tilde{u}^n, \tilde{v}^n) \in \mathcal{C}^{(\mu)} \cap \mathcal{T}_\delta^{(n)}(UV), (\tilde{u}^n, \tilde{v}^n) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j), (\tilde{u}^n, \tilde{v}^n) \neq (u^n, v^n) \right\}.
\end{aligned}$$

Denoting such an indicator function by  $\mathbb{1}^{(\mu_1, \mu_2)}(u^n, v^n)$ ,  $S_2$  can be bounded from above as  $S_2 \leq S_3$ , where

$$S_3 \triangleq \frac{(1-\varepsilon)(1-\varepsilon')}{(1+\eta)^2 2^{n(\tilde{R}_1 + \tilde{R}_2)}} \sum_{l, k} \sum_{(u^n, v^n)} \Omega_{u^n, v^n} \frac{2}{N_1 N_2} \sum_{\mu_1, \mu_2} \mathbb{1}^{(\mu_1, \mu_2)}(u^n, v^n) \mathbb{1}\{U^{n, (\mu_1)}(l) = u^n, V^{n, (\mu_2)}(k) = v^n\}.$$

Next, we use the Markov inequality to show that  $S_3 \leq \varepsilon$  with probability sufficiently close to 1. We first show that the expectation of  $S_3$  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 the pruned distribution, defined in (11). This gives us, for  $u^n \in \mathcal{T}_\delta^{(n)}(U)$  and  $v^n \in \mathcal{T}_\delta^{(n)}(V)$ ,

$$\begin{aligned}
&\mathbb{E} \left[ \mathbb{1}^{(\mu_1, \mu_2)}(u^n, v^n) \mathbb{1} \left\{ U^{n, (\mu_1)}(l) = u^n, V^{n, (\mu_2)}(k) = v^n \right\} \right] \\
&\leq \sum_{\substack{(\tilde{u}^n, \tilde{v}^n) \in \mathcal{T}_\delta^{(n)}(UV) \\ (\tilde{u}^n, \tilde{v}^n) \neq (u^n, v^n)}} \sum_{i, j} \sum_{(\tilde{l}, \tilde{k}) \neq (l, k)} \mathbb{E} \left[ \mathbb{1} \left\{ (u^n, v^n) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j) \right\} \mathbb{1} \left\{ (\tilde{u}^n, \tilde{v}^n) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j) \right\} \right] \\
&\quad \times \mathbb{1} \left\{ U^{n, (\mu_1)}(l) = u^n, V^{n, (\mu_2)}(k) = v^n \right\} \mathbb{1} \left\{ U^{n, (\mu_1)}(\tilde{l}) = \tilde{u}^n, V^{n, (\mu_2)}(\tilde{k}) = \tilde{v}^n \right\} \\
&\leq \frac{\lambda_{u^n}^A \lambda_{v^n}^B}{(1-\varepsilon)^2 (1-\varepsilon')^2} 2^{-n(I(U;V) - \delta_1)} \left[ 2^{n(\tilde{R}_1 - R_1)} 2^{n(\tilde{R}_2 - R_2)} \right. \\
&\quad \left. + 2^{n(\tilde{R}_1 - R_1)} + 2^{n(\tilde{R}_2 - R_2)} + 2^{-n(S(U) - \delta_1)} 2^{n\tilde{R}_1} 2^{n(\tilde{R}_2 - R_2)} + 2^{-n(S(V) - \delta_1)} 2^{n\tilde{R}_2} 2^{n(\tilde{R}_1 - R_1)} \right] \\
&\leq 5 \frac{\lambda_{u^n}^A \lambda_{v^n}^B}{(1-\varepsilon)^2 (1-\varepsilon')^2} 2^{-n(I(U;V) - 2\delta_1)} 2^{n(\tilde{R}_1 - R_1)} 2^{n(\tilde{R}_2 - R_2)}, \tag{54}
\end{aligned}$$

where  $\delta_1 \searrow 0$  as  $\delta \searrow 0$ . The first inequality follows from the union bound. The second inequality follows by evaluating the expectation of the indicator functions and the last inequality follows from the inequalities  $\tilde{R}_1 < S(U)$  and  $\tilde{R}_2 < S(V)$ . This implies

$$\mathbb{E}[S_3] \leq 10 \frac{2^{-n(I(U;V) - 2\delta_1)} 2^{n(\tilde{R}_1 - R_1)} 2^{n(\tilde{R}_2 - R_2)}}{(1+\eta)^2 (1-\varepsilon)(1-\varepsilon')} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \Omega_{u^n, v^n} \lambda_{u^n}^A \lambda_{v^n}^B.$$

We proceed using the following lemma.

**Lemma 9.** For  $\lambda_{u^n}^A$  and  $\lambda_{v^n}^B$  as defined in (6) and  $\Omega_{u^n, v^n}$  defined above, we have

$$\sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \Omega_{u^n, v^n} \lambda_{u^n}^A \lambda_{v^n}^B \leq 2^{n\delta_{AB}},$$

for some  $\delta_{AB} \searrow 0$  as  $\delta \searrow 0$ .

*Proof.* Firstly, note that

$$\sum_{\substack{u^n \in \mathcal{T}_\delta^{(n)}(U) \\ v^n \in \mathcal{T}_\delta^{(n)}(V)}} \Omega_{u^n, v^n} \lambda_{u^n}^A \lambda_{v^n}^B = \text{Tr} \left\{ \left[ \sqrt{\rho_A^{\otimes n}}^{-1} \left( \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A \Lambda_{u^n}^A \right) \sqrt{\rho_A^{\otimes n}} \otimes \sqrt{\rho_B^{\otimes n}}^{-1} \left( \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \lambda_{v^n}^B \Lambda_{v^n}^B \right) \sqrt{\rho_B^{\otimes n}} \right] \rho_{AB}^{\otimes n} \right\}. \quad (55)$$

Consider,

$$\begin{aligned} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A \Lambda_{u^n}^A &= \hat{\Pi}^A \Pi_{\rho_A} \left( \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A \Pi_{u^n}^A \hat{\rho}_{u^n}^A \Pi_{u^n}^A \right) \Pi_{\rho_A} \hat{\Pi}^A \\ &\leq \hat{\Pi}^A \Pi_{\rho_A} \left( \sum_{u^n} \lambda_{u^n}^A \hat{\rho}_{u^n}^A \right) \Pi_{\rho_A} \hat{\Pi}^A \\ &= \hat{\Pi}^A \Pi_{\rho_A} \rho_A^{\otimes n} \Pi_{\rho_A} \hat{\Pi}^A \leq 2^{-n(S(\rho_A) - \delta_A)} \hat{\Pi}^A \Pi_{\rho_A} \hat{\Pi}^A = 2^{-n(S(\rho_A) - \delta_A)} \Pi_{\rho_A} \hat{\Pi}^A \Pi_{\rho_A}. \end{aligned}$$

where the first inequality is obtained by using  $\Pi_{u^n}^A \hat{\rho}_{u^n}^A \Pi_{u^n}^A \leq \hat{\rho}_{u^n}^A$  for all  $u^n \in \mathcal{T}_\delta^{(n)}(U)$  and then by adding terms belonging to  $\mathcal{U}^n \setminus \mathcal{T}_\delta^{(n)}(U)$  into the summation. The subsequent inequality and the equality, follows from the properties of a typical projector with  $\delta_A \searrow 0$  as  $\delta \searrow 0$ , and the commutativity of  $\hat{\Pi}^A$  and  $\Pi_{\rho_A}$ , respectively. This implies,

$$\sqrt{\rho_A^{\otimes n}}^{-1} \left( \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A \Lambda_{u^n}^A \right) \sqrt{\rho_A^{\otimes n}} \leq 2^{-n(S(\rho_A) - \delta_A)} \sqrt{\rho_A^{\otimes n}}^{-1} \Pi_{\rho_A} \hat{\Pi}^A \Pi_{\rho_A} \sqrt{\rho_A^{\otimes n}} \leq 2^{2n\delta_A} \hat{\Pi}^A, \quad (56)$$

where the last inequality again appeals to the fact that  $\Pi_{\rho_A}$  and  $\hat{\Pi}^A$  commute. Similarly, using the same arguments above for the operators acting on  $\mathcal{H}_B$ , we have

$$\sqrt{\rho_B^{\otimes n}}^{-1} \left( \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \lambda_{v^n}^B \Lambda_{v^n}^B \right) \sqrt{\rho_B^{\otimes n}} \leq 2^{2n\delta_B} \hat{\Pi}^B, \quad (57)$$

where  $\delta_B \searrow 0$  as  $\delta \searrow 0$ . Using (56) and (57) in (55), gives

$$\sum_{u^n, v^n} \Omega_{u^n, v^n} \lambda_{u^n}^A \lambda_{v^n}^B \leq 2^{2n(\delta_A + \delta_B)} \text{Tr} \left\{ \left( \hat{\Pi}^B \otimes \hat{\Pi}^A \right) \rho_{AB}^{\otimes n} \right\} \leq 2^{2n(\delta_A + \delta_B)} \text{Tr} \left\{ \rho_{AB}^{\otimes n} \right\} = 2^{2n(\delta_A + \delta_B)},$$

substituting  $\delta_{AB} = 2(\delta_A + \delta_B)$  gives the result.  $\square$

As a result, given any  $\epsilon \in (0, 1)$ , the above expectation can be made less than  $\frac{10\epsilon}{(1-\epsilon)(1-\epsilon')}$  for large enough  $n$  provided that  $(\tilde{R}_1 - R_1) + (\tilde{R}_2 - R_2) \leq I(U; V) - 2\delta_1 - \delta_{AB} - \delta$ . From Markov-inequality this implies that  $S_3 \leq \sqrt{\epsilon}$  with probability at least  $1 - \frac{10\sqrt{\epsilon}}{(1-\epsilon)(1-\epsilon')}$ .

### C. Proof of Proposition 3

We bound  $\tilde{S}$  as  $\tilde{S} \leq \tilde{S}_2 + \tilde{S}_3 + \tilde{S}_4$ , where

$$\begin{aligned}\tilde{S}_2 &\triangleq \left\| \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{i>0} \sqrt{\rho_{AB}^{\otimes n}} \left( \Gamma_i^{A,(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | u_0^n, v_0^n) \right\|_1, \\ \tilde{S}_3 &\triangleq \left\| \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{j>0} \sqrt{\rho_{AB}^{\otimes n}} \left( \Gamma_0^{A,(\mu_1)} \otimes \Gamma_j^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | u_0^n, v_0^n) \right\|_1, \\ \tilde{S}_4 &\triangleq \left\| \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sqrt{\rho_{AB}^{\otimes n}} \left( \Gamma_0^{A,(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | u_0^n, v_0^n) \right\|_1.\end{aligned}$$

**Analysis of  $\tilde{S}_2$ :** We have

$$\begin{aligned}\tilde{S}_2 &\leq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{z^n} P_{Z|U,V}^n(z^n | u_0^n, v_0^n) \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \sum_{i>0} \Gamma_i^{A,(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\ &\leq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( A_{u^n}^{(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\ &\stackrel{w.h.p.}{\leq} \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \left\| \sqrt{\rho_B^{\otimes n}} \Gamma_0^{B,(\mu_2)} \sqrt{\rho_B^{\otimes n}} \right\|_1 \\ &= \frac{1}{N_2} \sum_{\mu_2} \left\| \sum_{v^n} \lambda_{v^n}^B \hat{\rho}_{v^n}^B - \sum_{v^n} \sqrt{\rho_B^{\otimes n}} B_{v^n}^{(\mu_2)} \sqrt{\rho_B^{\otimes n}} \right\|_1 \\ &\leq \frac{1}{N_2} \sum_{\mu_2} \left\| \sum_{v^n} \lambda_{v^n}^B \hat{\rho}_{v^n}^B - \frac{(1-\varepsilon')}{(1+\eta)} \frac{1}{2^{n\tilde{R}_2}} \sum_{k=1}^{2^{n\tilde{R}_2}} \hat{\rho}_{V_k^{n,(\mu_2)}}^B \right\|_1 + \underbrace{\frac{1}{N_2} \sum_{\mu_2} \sum_{v^n} \zeta_{v^n}^{(\mu_2)} \|\hat{\rho}_{v^n}^B - \Lambda_{v^n}^B\|_1}_{\tilde{S}_{22}},\end{aligned}\tag{58}$$

where the first inequality uses triangle inequality, the second uses the fact that  $\sum_{i>0} \Gamma_i^{A,(\mu_1)} = \sum_{u^n} A_{u^n}^{(\mu_1)}$ . The next inequality follows by using Lemma 3 where we use the result that with high probability (letting  $E_1$  denote this event) we have  $\sum_{u^n} A_{u^n}^{(\mu_1)} \leq I$ , given that  $\tilde{R}_1 \geq I(U; RB)_{\sigma_2}$ . Finally, the last inequality follows again from triangle inequality.

Regarding the first term in (58) using Lemma 7 we claim that for all sufficiently large  $n$ , the term can be made arbitrarily small with high probability (letting  $E_2$  denote this event), given the rate  $\tilde{R}_2$  satisfies  $\tilde{R}_2 \geq I(V; RA)_{\sigma_2}$  where  $\sigma_2$  is as defined in the statement of the theorem. Note that the requirements we obtain on  $\tilde{R}_1$  and  $\tilde{R}_2$  here were already imposed earlier in Section VIII-A. And as for the second term we use the gentle measurement lemma (as in (61)) and bound its expected value as

$$\mathbb{E} \left[ \frac{1}{N_2} \sum_{\mu_2} \sum_{v^n} \zeta_{v^n}^{(\mu_2)} \|\hat{\rho}_{v^n}^B - \Lambda_{v^n}^B\|_1 \right] = \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \frac{\lambda_{v^n}^B}{(1+\eta)} \|\hat{\rho}_{v^n}^B - \Lambda_{v^n}^B\|_1 \leq \epsilon_{\tilde{S}_2},$$

where the inequality is based on the repeated usage of the average gentle measurement lemma by setting  $\epsilon_{\tilde{S}_2} = \frac{(1-\varepsilon')}{(1+\eta)} (2\sqrt{\varepsilon'_B} + 2\sqrt{\varepsilon''_B})$  with  $\epsilon_{\tilde{S}_2} \searrow 0$  as  $\delta \searrow 0$  and  $\varepsilon'_B = \varepsilon' + 2\sqrt{\varepsilon'}$  and  $\varepsilon''_B = 2\varepsilon' + 2\sqrt{\varepsilon'}$  (see (35) in [2] for more details). Now, by using Markov inequality  $\mathbb{P}(E_3) \leq \sqrt{\epsilon_{\tilde{S}_2}}$ , where  $E_3 \triangleq \{\tilde{S}_{22} \geq \sqrt{\epsilon_{\tilde{S}_2}}\}$ . Hence, using union bound on the three events  $E_1, E_2$  and  $E_3$ ,  $\tilde{S}_2$  can be made arbitrarily small, for sufficiently large  $n$ , with high probability.

**Analysis of  $\tilde{S}_3$ :** Due to the symmetry in  $\tilde{S}_2$  and  $\tilde{S}_3$ , the analysis of  $\tilde{S}_3$  follows very similar arguments as that of  $\tilde{S}_2$  and hence we skip it.

**Analysis of  $\tilde{S}_4$ :** We have

$$\begin{aligned}\tilde{S}_4 &\leq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{z^n} P_{Z|U,V}^n(z^n | u_0^n, v_0^n) \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \Gamma_0^{A,(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\ &\leq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \Gamma_0^{A,(\mu_1)} \otimes I \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 + \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{v^n} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \Gamma_0^{A,(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1, \quad (59)\end{aligned}$$

where the inequalities above are obtained by a straight forward substitution and use of triangle inequality. With the above constraints on  $\tilde{R}_1$  and  $\tilde{R}_2$ , we have  $0 \leq \Gamma_0^{A,(\mu_1)} \leq I$  and  $0 \leq \Gamma_0^{B,(\mu_2)} \leq I$ . This simplifies the first term in (59) as

$$\frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \Gamma_0^{A,(\mu_1)} \otimes I \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 = \frac{1}{N_1} \sum_{\mu_1} \left\| \sqrt{\rho_A^{\otimes n}} \left( \Gamma_0^{A,(\mu_1)} \right) \sqrt{\rho_A^{\otimes n}} \right\|_1.$$

Similarly, the second term in (59) simplifies using Lemma 3 as

$$\frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{v^n} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \Gamma_0^{A,(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \leq \frac{1}{N_1} \sum_{\mu_1} \left\| \sqrt{\rho_A^{\otimes n}} \left( \Gamma_0^{A,(\mu_1)} \right) \sqrt{\rho_A^{\otimes n}} \right\|_1.$$

Using these simplifications, we have

$$\tilde{S}_4 \leq \frac{2}{N_1} \sum_{\mu_1} \left\| \sqrt{\rho_A^{\otimes n}} \left( \Gamma_0^{A,(\mu_1)} \right) \sqrt{\rho_A^{\otimes n}} \right\|_1.$$

The above expression is similar to the one obtained in the simplification of  $\tilde{S}_2$  and hence we can bound  $\tilde{S}_4$  using the same constraints as  $\tilde{S}_2$ , for sufficiently large  $n$ .

#### D. Proof of Proposition 4

Recalling  $S_2$ , we have

$$\begin{aligned}S_2 &\leq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{z^n} \sum_{u^n, v^n} \left| P_{Z|U,V}^n(z^n | u^n, v^n) - P_{Z|U,V}^n(z^n | e^{(\mu)}(u^n, v^n)) \right| \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\ &\leq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n, v^n} \sum_{z^n} \left| P_{Z|U,V}^n(z^n | u^n, v^n) - P_{Z|U,V}^n(z^n | e^{(\mu)}(u^n, v^n)) \right| \gamma_{u^n}^{(\mu_1)} \zeta_{v^n}^{(\mu_2)} \Omega_{u^n, v^n} \\ &\leq \frac{(1-\varepsilon)(1-\varepsilon')}{(1+\eta)^2} \frac{2}{N_1 N_2} 2^{-n(\tilde{R}_1 + \tilde{R}_2)} \sum_{\mu_1, \mu_2} \sum_{l, k} \sum_{u^n, v^n} \Omega_{u^n, v^n} \mathbb{1}^{(\mu_1, \mu_2)}(u^n, v^n) \mathbb{1}_{\{U^n, (\mu_1)(l) = u^n\}} \mathbb{1}_{\{V^n, (\mu_2)(k) = v^n\}},\end{aligned}$$

where  $\Omega_{u^n, v^n}$  and  $\mathbb{1}^{(\mu_1, \mu_2)}(u^n, v^n)$  are defined as

$$\begin{aligned}\Omega_{u^n, v^n} &\triangleq \text{Tr} \left\{ \sqrt{\rho_A^{\otimes n} \otimes \rho_B^{\otimes n}}^{-1} \left( \Lambda_{u^n}^A \otimes \Lambda_{v^n}^B \right) \sqrt{\rho_A^{\otimes n} \otimes \rho_B^{\otimes n}}^{-1} \rho_{AB}^{\otimes n} \right\}, \\ \mathbb{1}^{(\mu_1, \mu_2)}(u^n, v^n) &\triangleq \mathbb{1} \left\{ \exists (\tilde{u}^n, \tilde{v}^n, i, j) : (u^n, v^n) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j), \right. \\ &\quad \left. (\tilde{u}^n, \tilde{v}^n) \in \mathcal{C}^{(\mu)} \cap \mathcal{T}_\delta^{(n)}(UV), (\tilde{u}^n, \tilde{v}^n) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j), (\tilde{u}^n, \tilde{v}^n) \neq (u^n, v^n) \right\}.\end{aligned}$$

We know from the simplification in (54) that

$$\mathbb{E} \left[ \mathbb{1}^{(\mu_1, \mu_2)}(u^n, v^n) \mathbb{1}_{\{U^n, (\mu_1)(l) = u^n\}} \mathbb{1}_{\{V^n, (\mu_2)(k) = v^n\}} \right] \leq \frac{5 \lambda_{u^n}^A \lambda_{v^n}^B}{(1-\varepsilon)^2 (1-\varepsilon')^2} 2^{-n(I(U;V) - 2\delta_1)} 2^{n(\tilde{R}_1 - R_1)} 2^{n(\tilde{R}_2 - R_2)}.$$

Substituting this in the expression for  $S_{12}$  gives

$$\mathbb{E}[S_2] \leq 10 \frac{2^{-n(I(U;V) - 2\delta_1)} 2^{n(\tilde{R}_1 - R_1)} 2^{n(\tilde{R}_2 - R_2)}}{(1+\eta)^2 (1-\varepsilon)^2 (1-\varepsilon')^2} \sum_{u^n, v^n} \Omega_{u^n, v^n} \lambda_{u^n}^A \lambda_{v^n}^B$$

$$\leq 10 \frac{2^{-n(I(U;V)-2\delta_1-\delta_{AB})} 2^{n(\tilde{R}_1-R_1)} 2^{n(\tilde{R}_2-R_2)}}{(1+\eta)^2(1-\varepsilon)^2(1-\varepsilon')^2},$$

where the second inequality above uses Lemma 9. Therefore, if  $\tilde{R}_1 + \tilde{R}_2 - R_1 - R_2 \leq I(U;V)_{\sigma_3} - 2\delta_1 - \delta_{AB} - \delta$ , then we have  $\mathbb{E}[S_2] \leq 10 \frac{2^{-n\delta}}{(1+\eta)^2(1-\varepsilon)(1-\varepsilon')}$ . The proposition follows from Markov Inequality.

### E. Proof of Proposition 5

We start by adding and subtracting the following terms in  $Q_2$

$$\begin{aligned} (i) & \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} (\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \\ (ii) & \sum_{u^n, v^n} \frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \sqrt{\rho_{AB}^{\otimes n}} \left( \bar{\Lambda}_{u^n}^A \otimes \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \\ (iii) & \sum_{u^n, v^n} \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sqrt{\rho_{AB}^{\otimes n}} \left( A_{u^n}^{(\mu_1)} \otimes \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n). \end{aligned}$$

This gives us  $Q_2 \leq Q_{21} + Q_{22} + Q_{23} + Q_{24}$ , where

$$\begin{aligned} Q_{21} & \triangleq \sum_{z^n} \left\| \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \left( \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} A_{u^n}^{(\mu_1)} \right) \otimes \bar{\Lambda}_{v^n}^B - \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1, \\ Q_{22} & \triangleq \sum_{z^n} \left\| \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \bar{\Lambda}_{u^n}^A \otimes \left( \frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \bar{\Lambda}_{v^n}^B \right) \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1, \\ Q_{23} & \triangleq \sum_{z^n} \left\| \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \bar{\Lambda}_{u^n}^A \otimes \left( \frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \bar{\Lambda}_{v^n}^B \right) - \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} A_{u^n}^{(\mu_1)} \otimes \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1, \\ Q_{24} & \triangleq \sum_{z^n} \left\| \sum_{u^n, v^n} \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sqrt{\rho_{AB}^{\otimes n}} \left( A_{u^n}^{(\mu_1)} \otimes \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \bar{\Lambda}_{v^n}^B - A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1. \end{aligned}$$

We start by analyzing  $Q_{21}$ . Note that  $Q_{21}$  is exactly same as  $Q_1$  and hence using the same rate constraints as  $Q_1$ , this term can be bounded. Next, consider  $Q_{22}$ . Substitution of  $\zeta_{v^n}^{(\mu_2)}$  gives

$$\begin{aligned} Q_{22} & = \left\| \sum_{u^n, v^n, z^n} \lambda_{u^n, v^n}^{AB} \hat{\rho}_{u^n, v^n}^{AB} \otimes P_{Z|U,V}^n(z^n|u^n, v^n) |z^n\rangle\langle z^n| \right. \\ & \quad \left. - \frac{(1-\varepsilon')}{(1+\eta)} \frac{1}{2^{n(\tilde{R}_2+C_2)}} \sum_{\mu_2, k} \sum_{u^n, v^n, z^n} \mathbb{1}_{\{V^{n, (\mu_2)}(k)=v^n\}} \frac{\lambda_{u^n, v^n}^{AB}}{\lambda_{v^n}^B} \hat{\rho}_{u^n, v^n}^{AB} \otimes P_{Z|U,V}^n(z^n|u^n, v^n) |z^n\rangle\langle z^n| \right\|_1, \end{aligned}$$

where the equality uses the triangle inequality for block operators. From here on, we use Lemma 7 to bound  $Q_{22}$ .

$$\mathcal{T}_{v^n} = \sum_{u^n, z^n} \frac{\lambda_{u^n, v^n}^{AB}}{\lambda_{v^n}^B} \hat{\rho}_{u^n, v^n}^{AB} \otimes P_{Z|U,V}^n(z^n|u^n, v^n) |z^n\rangle\langle z^n|,$$

Note that  $\mathcal{T}_{v^n}$  can be written in tensor product form as  $\mathcal{T}_{v^n} = \otimes_{i=1}^n \mathcal{T}_{v_i}$ . This simplifies  $Q_{22}$  as

$$Q_{22} = \left\| \sum_{v^n} \lambda_{v^n}^B \mathcal{T}_{v^n} - \frac{(1-\varepsilon')}{(1+\eta)} \frac{1}{2^{n(\tilde{R}_2+C_2)}} \sum_{\mu_2, k} \mathcal{T}_{V^{n, (\mu_2)}(k)} \right\|_1.$$

Lemma 7 gives us, for any given  $\epsilon_{Q_{22}}, \delta_{Q_{22}} \in (0, 1)$ , if

$$\tilde{R}_2 + C_2 > S \left( \sum_{v \in \mathcal{V}} \lambda_v^B \mathcal{T}_v \right) - \sum_{v \in \mathcal{V}} \lambda_v^B S(\mathcal{T}_v) = I(RZ; V)_{\sigma_4}, \quad (60)$$

then  $\mathbb{P}(Q_{22} \geq \epsilon_{Q_{22}}) \leq \delta_{Q_{22}}$  for sufficiently large  $n$ .

Now, we move on to consider  $Q_{23}$ . Taking expectation with respect to the codebook  $\mathcal{C}^{(\mu)} = (\mathcal{C}_1^{(\mu_1)}, \mathcal{C}_2^{(\mu_2)})$  gives

$$\begin{aligned}
\mathbb{E}[Q_{23}] &\leq \mathbb{E}_{\mathcal{C}} \left[ \sum_{z^n, v^n} \frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \left\| \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} (\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right. \right. \\
&\quad \left. \left. - \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \frac{1}{N_1} \sum_{\mu_1} A_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1 \right] \\
&= \mathbb{E}_{\mathcal{C}_1} \left[ \sum_{z^n, v^n} \frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \frac{\mathbb{E}_{\mathcal{C}_2} [\zeta_{v^n}^{(\mu_2)}]}{\lambda_{v^n}^B} \left\| \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} (\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right. \right. \\
&\quad \left. \left. - \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \frac{1}{N_1} \sum_{\mu_1} A_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1 \right] \\
&= \mathbb{E}_{\mathcal{C}_1} \left[ \sum_{z^n, v^n} \frac{1}{(1+\eta)} \left\| \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} (\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right. \right. \\
&\quad \left. \left. - \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} \left( \frac{1}{N_1} \sum_{\mu_1} A_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1 \right] \\
&= \mathbb{E} \left[ \frac{J}{(1+\eta)} \right],
\end{aligned}$$

where the inequality above is obtained by using the triangle inequality, and the first equality follows as  $\mathcal{C}_1^{(\mu_1)}$  and  $\mathcal{C}_2^{(\mu_2)}$  are generated independently. The last equality follows from the definition of  $J$  as in (44). Hence, we use the result obtained in bounding  $\mathbb{E}[J]$ .

Finally, we consider  $Q_{24}$ .

$$\begin{aligned}
Q_{24} &\leq \sum_{u^n, v^n} \sum_{z^n} P_{Z|U,V}^n(z^n|u^n, v^n) \left\| \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sqrt{\rho_{AB}^{\otimes n}} \left( A_{u^n}^{(\mu_1)} \otimes \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} \right. \\
&\quad \left. - \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sqrt{\rho_{AB}^{\otimes n}} \left( A_{u^n}^{(\mu_1)} \otimes \zeta_{v^n}^{(\mu_2)} (\sqrt{\rho_B}^{-1} \Lambda_{v^n}^B \sqrt{\rho_B}^{-1}) \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\
&\leq \frac{1}{N_2} \sum_{\mu_2} \sum_{u^n, v^n} \zeta_{v^n}^{(\mu_2)} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \frac{1}{N_1} \sum_{\mu_1} A_{u^n}^{(\mu_1)} \otimes \frac{1}{\lambda_{v^n}^B} \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} \right. \\
&\quad \left. - \sqrt{\rho_{AB}^{\otimes n}} \left( \frac{1}{N_1} \sum_{\mu_1} A_{u^n}^{(\mu_1)} \otimes (\sqrt{\rho_B}^{-1} \Lambda_{v^n}^B \sqrt{\rho_B}^{-1}) \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1,
\end{aligned}$$

where the inequalities above are obtained by substituting in the definition of  $B_{v^n}^{(\mu_2)}$  and using multiple triangle inequalities. Taking expectation of  $Q_{24}$  with respect to the second codebook generation, we get

$$\begin{aligned}
\mathbb{E}_{\mathcal{C}_2}[Q_{24}] &\leq \sum_{u^n} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \frac{\lambda_{v^n}^B}{(1+\eta)} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left( \frac{1}{N_1} \sum_{\mu_1} A_{u^n}^{(\mu_1)} \otimes \frac{1}{\lambda_{v^n}^B} \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} \right. \\
&\quad \left. - \sqrt{\rho_{AB}^{\otimes n}} \left( \frac{1}{N_1} \sum_{\mu_1} A_{u^n}^{(\mu_1)} \otimes (\sqrt{\rho_B}^{-1} \Lambda_{v^n}^B \sqrt{\rho_B}^{-1}) \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\
&\stackrel{w.h.p}{\leq} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \frac{\lambda_{v^n}^B}{(1+\eta)} \left\| \sqrt{\rho_B^{\otimes n}} \left( \frac{1}{\lambda_{v^n}^B} \bar{\Lambda}_{v^n}^B - \sqrt{\rho_B}^{-1} \Lambda_{v^n}^B \sqrt{\rho_B}^{-1} \right) \sqrt{\rho_B^{\otimes n}} \right\|_1
\end{aligned}$$



$$= \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \frac{\lambda_{v^n}^B}{(1+\eta)} \|\hat{\rho}_{v^n}^B - \Lambda_{v^n}^B\|_1 \leq \frac{(1-\varepsilon')}{(1+\eta)} (2\sqrt{\varepsilon'_B} + 2\sqrt{\varepsilon''_B}) = \epsilon_{Q_{24}}, \quad (61)$$

where the second inequality above follows by using Lemma 3 and the fact that  $\sum_{u^n} \frac{1}{N_1} \sum_{\mu_1} A_{u^n}^{(\mu_1)} \leq I$ , with high probability, and the last inequality uses the result based on the average gentle measurement lemma by setting  $\epsilon_{Q_{24}} = \frac{(1-\varepsilon')}{(1+\eta)} (2\sqrt{\varepsilon'_B} + 2\sqrt{\varepsilon''_B})$  with  $\epsilon_{Q_{24}} \searrow 0$  as  $\delta \searrow 0$  and  $\varepsilon'_B = \varepsilon' + 2\sqrt{\varepsilon'}$  and  $\varepsilon''_B = 2\varepsilon' + 2\sqrt{\varepsilon'}$  (see (35) in [2] for more details ). This completes the proof for  $Q_{24}$  and hence for all the terms corresponding to  $Q_2$ .

## REFERENCES

- [1] 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.
- [2] 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.
- [3] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter, “Quantum reverse shannon theorem,” *arXiv preprint arXiv:0912.5537*, 2009.
- [4] 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.
- [5] 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.
- [6] 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.
- [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, “Convex-split and hypothesis testing approach to one-shot quantum measurement compression and randomness extraction,” *IEEE Transactions on Information Theory*, vol. 65, no. 9, pp. 5905–5924, 2019.
- [9] T. Berger, “Multiterminal source coding,” *The Inform. Theory Approach to Communications*, G. Longo, Ed., New York: Springer-Verlag, 1977.
- [10] I. Devetak and A. Winter, “Classical data compression with quantum side information,” *Physical Review A*, vol. 68, no. 4, p. 042301, 2003.
- [11] 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.
- [12] M. Berta, J. M. Renes, and M. M. Wilde, “Identifying the information gain of a quantum measurement,” *IEEE Transactions on Information Theory*, vol. 60, no. 12, pp. 7987–8006, 2014.
- [13] A. S. Holevo, *Quantum systems, channels, information: a mathematical introduction*. Walter de Gruyter, 2012, vol. 16.
- [14] D. S. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” *IEEE Transactions on Information Theory*, vol. 19, no. 4, p. 471–480, July 1973.
- [15] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge university press, 2011.
- [16] M. M. Wilde, “From classical to quantum shannon theory,” *arXiv preprint arXiv:1106.1445*, 2011.
- [17] G. M. Ziegler, *Lectures on polytopes*. Springer Science & Business Media, 2012, vol. 152.
- [18] 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.
- [19] J. B. Conway, *A Course in Functional Analysis*. Springer New York, 1985.
- [20] P. A. Meyer, *Quantum probability for probabilists*. Springer, 2006.
- [21] A. B. Wagner and V. Anantharam, “An improved outer bound for multiterminal source coding,” *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 1919–1937, 2008.
- [22] H. Martens and W. M. de Muynck, “Nonideal quantum measurements,” *Foundations of Physics*, vol. 20, no. 3, pp. 255–281, 1990.
- [23] T. A. Atif, A. Padakandla, and S. S. Pradhan, “Source coding for synthesizing correlated randomness,” *arXiv preprint arXiv:2004.03651*, 2020.