From Compression to Communication: Performance Limits for Quantum Networks

by

Touheed Anwar Atif

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical and Computer Engineering)
in The University of Michigan
2023

Doctoral Committee:

Professor S. Sandeep Pradhan, Chair
Associate Professor Mahdi Cheraghchi
Associate Professor Hessam Mahdavifar
Research Professor Andreas Winter, Universitat Autònoma de Barcelona
Professor Jun Zhang

Touheed Anwar Atif

touheed@umich.edu

ORCID iD: 0000-0002-8424-2589

© Touheed Anwar Atif 2023

To my parents and my wife

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to my professor and thesis advisor, Prof. S. Sandeep Pradhan, for his unwavering support, guidance, and encouragement throughout my academic journey. His expertise and knowledge in information theory have been invaluable in shaping my research and shaping me into the scholar I am today. His trust and confidence in my abilities allowed me to push the boundaries of my understanding and pursue avenues which otherwise seemed impossible. His support and inspiration were always there when I needed it, and I am deeply grateful for his contributions to my academic journey. Thank you for being such a wonderful advisor and mentor.

I would also like to extend my thanks to my seniors, especially Arun Padakandla and Mohsen Heidari for sharing their experiences, insights, and vital advice throughout my academic journey. Mohsen's support and mentorship have been invaluable in helping me navigate the challenges of graduate school and pursue my goals. Arun's constant collaboration has been a true testament to the power of collaboration and the value of teamwork. His insights and expertise have been invaluable in helping me to produce high-quality research and advance our collective understanding of quantum information theory. Their examples of excellence and dedication have inspired me to strive for the highest standards in my own work, and I am grateful for their positive impact on my academic and professional development.

I would also like to extend my heartfelt thanks to all my thesis committee members, Prof. Hessam Mahdavifar , Prof. Mahdi Cheraghchi, Prof. Jun Zhang and Prof. Andreas Winter, for their time, effort, and insightful feedback during the development of my thesis. Prof. Mahdavifar's courses, EECS 554 and EECS 650, have been extremely beneficial to me. Attending Prof. Zhang's course on information geometry greatly enhanced my knowledge in the geometric perspective of information.

I am truly grateful to Prof. Winter, whose problem formulation served as the foundation of my thesis research. His insightful contribution to the quantum information field has been an inspiration, and his ideas have greatly informed and influenced my work. I still remember his extremely pleasant gesture of initiating a conversation after my presentation at my first ever conference presentation, ISIT 2019. I am grateful for the opportunity to build upon his work and to contribute to the advancement of our understanding in this area.

I would also like to express my deepest gratitude to my wife, who have been a constant source of love, support, and encouragement throughout my academic journey. Despite facing challenges and surviving on limited resources, she has been an unwavering source of strength and motivation. Her sacrifices and dedication have made this achievement possible, and I am deeply grateful for her love and support. Her belief in me and my abilities has been a constant source of inspiration, and I am forever grateful for their presence in my life.

I would also like to extend my sincerest gratitude to my parents, who have made great sacrifices, including the pain of separation, to support my education and academic pursuits. Their unwavering love and support have been a constant source of inspiration and motivation, and their sacrifices have not gone unnoticed. I am deeply grateful for their investment in my future, and for their belief in my abilities and potential. Their support has been instrumental in helping me to achieve this important milestone, and I am forever grateful for their presence in my life. Thank you, Mummy and Pappa, for everything.

I would like to thank my family: my grandparents, brothers, and in-laws. Their prayers and presence fill my life with a happiness that is unquantifiable.

Finally, I would like to acknowledge the support and encouragement of my friends, who have been with me through thick and thin, and have always believed in me. Thank you all for your invaluable contributions to my academic journey.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF APPENDICES

Appendix

LIST OF ABBREVIATIONS

CQ classical quantum

QC quantum-to-classical

PtP point-to-point

IID independent and identically distributed

ABSTRACT

Quantum algorithms require quantum computers performing logical operations on a sufficiently large number of entangled qubits. Unfortunately, state-of-the-art quantum computers can only operate on tens of qubits. A solution to this scalability challenge is to employ the distributed paradigm, where a network of small-scale quantum computers are used in a distributed manner. It is the aim of my work to study the fundamental limits of distributed quantum problems and to further enhance their performance by exploiting the structure inherent to these problems employing asymptotically good Algebraic codes. This thesis consists of two parts.

The first part studies the task of faithfully simulating a distributed quantum measurement, wherein we provide a protocol for the three parties, Alice, Bob and Charlie, 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 Charlie. 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 Charlie, and the common randomness shared among the three parties, while faithfully simulating independent repeated instances of the original measurement. We characterize a set of sufficient communication and common randomness rates required for asymptotic simulatability in terms of single-letter quantum information quantities. We further improve the results obtained in the above by exploiting the structure present in the Charlie's stochastic bivariate mapping using random structured POVMs based on asymptotically good algebraic codes. The algebraic structure of these codes is matched to that of the bivariate function that models the action of Charlie. This leads to the computation being performed on the fly, thus obviating the need to reconstruct individual measurement outcomes at Charlie. We provide examples to illustrate the information-theoretic gains attained by endowing POVMs with algebraic structure. As an application of the distributed measurement compression problem, we also demonstrate a multi-party purity distillation protocol.

Concluding this part, we consider the lossy quantum source coding problem where the objective

is to compress a given quantum source below its von Neumann entropy. Inspired by the duality connections between the rate-distortion and channel coding problems in the classical setting, we propose a new formulation for the lossy quantum source coding problem. We require that the reconstruction of the compressed quantum source fulfill a global error constraint and employ the notion of a "posterior reference map" to measure the reconstruction error. Using these, we characterize the asymptotic performance limit of this problem in terms of single-letter coherent information of the given posterior reference map.

In the second part of this thesis, we study the advantage algebraic structured codes can provide, to the class of the classical-quantum network problems. In particular, we investigate two problems. Firstly, we consider the problem of communicating a general bivariate function of two classical sources observed at the encoders of a classical-quantum multiple access channel. We propose and analyze a coding scheme based on algebraic structured coset codes that enables the decoder to recover the desired function without recovering the sources themselves. We derive a new set of sufficient conditions that are weaker than the current known for identified examples. In addition, we analyze the performance of these algebraic codes toward studying the capacity of a special class of 3-user classical-quantum interference channel.

CHAPTER I

Introduction

1.1 Quantum information theory

Quantum information theory is a branch of information theory that deals with the study of representation, transmission, manipulation, and processing of information using quantum mechanics. It considers the fundamental differences between classical and quantum information, as well as the potential benefits and limitations of quantum systems for information processing tasks. Based on the number of agents (or parties) involved in the system, the information processing tasks can be broadly classified as (i) point-to-point (PtP) and (ii) multi-terminal. We refer to PtP as problems involving two parties. Problems that involve more than two parties are referred to (in this thesis) as multi-terminal ones. Since these multi-terminal setups can be varying in terms of how each agent is connected to the other, we describe them formally in the respective chapters. Our main focus in this thesis is to characterize performance limits of different multi-terminal problems (involving three or four agents) from an information theoretic approach. At times, we also consider PtP problems when they can serve as providing insights to the underlying idea without getting impeded by the technical details.

1.2 Quantum network information theory

A quantum network is a system that consists of multiple quantum systems connected by either quantum or classical channels (aka links) that are perhaps noiseless or noisy, allowing for transfer of information between different systems. Quantum network information theory, or multi-terminal quantum information theory, investigates the performance limits of these problems from an funda-

mental and theoretical perspective, and provides a way to study the distribution and manipulation of information between multiple parties. These problems are important for the development of quantum communication and quantum networking technologies, as they allow for the study of quantum information processing in more complex and realistic scenarios.

From the analogous works in classical information theory pertaining to multi-terminal problems, we observe that analysis of multi-terminal problems do not trivially follow from the results on PtP problems. The former are known to necessitate new fundamental techniques, and many of them still remain unsolved in terms of exact characterization of performance limits. A similar phenomena is observed in the quantum setting prompting researchers to comprehensively investigate a variety of multi-terminal quantum information theoretic problems.

Secondly, these problems are also valuable in their own right. In multi-terminal quantum information problems, the parties involved may have different goals or objectives, and their interactions can result in non-trivial quantum correlations. There are several examples where the study of multi-terminal quantum information problems can also lead to a deeper understanding of the foundations of quantum mechanics and the nature of quantum entanglement. One example is the phenomenon of quantum nonlocality, which refers to the non-classical correlations that can exist between distant quantum systems. The study of multi-terminal quantum information problems, has provided evidence for the existence of quantum nonlocality (beyond the standard Bell non-locality) *Šupić et al.* (2022).

Finally, there is another reason, specific to the quantum setting and relating to this thesis, why a detailed study of such multi-terminal problems is useful. A generic feature witnessed in the classical setting is that the results obtained for the point-to-point problems often provide insights towards analyzing the multi-terminal problems. But in the quantum setting, the opposite is also sometimes true: a study of multi-terminal problem can help in the characterizing fundamental limits of a PtP problem. A prime example of this is the Devetak's proof *Devetak* (2005b) which characterizes the capacity of a PtP quantum communication problem. Devetak's inspiration was the theorem he had developed to characterize the private classical capacity of a quantum channel, a multi-terminal problem involving three distributed agents.

The question now is whether there are any more instances of such intriguing behavior. In this thesis, we provide yet another such example. By examining the multi-terminal variants of the

measurement compression and classical communication over quantum channel problems, we characterize the asymptotic performance of the lossy quantum source compression problem. Details on this contribution are provided in this chapter's contribution section, and complete results are outlined in Chapter V. These developments serve as inspirations for further advancements in quantum network information theory, highlighting how multi-terminal quantum information problems provide a challenging and promising area of study that can lead to both theoretical and practical progress in the field of quantum information.

1.3 Algebraic structure in multi-terminal information theory

A significant advancement toward improving the understanding of multi-terminal problems was the use of algebraic structured codes. In the stationary memoryless scenario, there are two typical ways for determining the performance limits of communication problems. One is based on random coding and involves so-called independent and identically distributed (IID) code ensembles, with performance measured in terms of single-letter information quantities. The other is based on random coding in a one-shot context, with performance measured using smooth entropic values. Because the codes in these ensembles lack global structure, we refer to them as random coding approaches based on unstructured code ensembles. Unstructured code ensembles may not always achieve optimality for distributed multi-terminal settings. One of the early works by Korner and Marton *Korner and Marton* (1979) demonstrated this sub-optimality for a multi-terminal setup involving three agents. Motivated by this, a framework for constructing structured coding ensembles and improving the performance limits for classical problems have been the focus of many works. A few prominent among them include *Korner and Marton* (1979); *Nazer and Gastpar* (2007); *Padakandla and Pradhan* (2013); *Padakandla et al.* (2016); *Pradhan et al.* (2021). For the quantum setting, such a framework is still under development with a few recent works like *Hayashi and Vázquez-Castro* (2021). Inspired by this, this thesis aim at developing structured coding ensembles for the task for performing communication and compression for the quantum multi-terminal problems. It builds a broad theoretical framework, but it also predicts very specific effects.

1.4 Organization and contributions of this thesis

This thesis is divided into two parts. The first part deals with the quantum compression problems: from quantum-to-classical (QC) compression to pure quantum compression. The second part delves into quantum communication problems. In Section 1.5, we begin by formally stating the point-to-point measurement compression problem as was first introduced by Winter in *Winter* (2004). We highlight here the relevant work that preceded this work and also discuss the results that follow it. These include the different variants with respect resources and the problem setups, discussed in the literature and also the various applications, including the famous quantum-to-classical rate distortion theory *Datta et al.* (2013b), where the measurement compression problem is utilized.

In the next chapter, Chapter II, we introduce the setup of a distributed measurement compression problem involving three parties, and then formulate a protocol, in terms of available resources, allowed actions, and the objective the protocol needs to achieve, for the same *Atif et al.* (2022a). We build this chapter with the aim of providing an inner bound to the asymptotic performance limit of non-feedback distributed measurement compression. Toward this, we first provide two result for the feedback distributed compression problem: a one-shot and an n-letter characterization. We then formulate a distributed quantum-to-classical rate distortion problem ,and provide an inner and an outer bound to the rate distortion function (region) using the n-letter result of the feedback problem. Finally, we conclude this chapter with the main result of the non-feedback problem. A mutual covering lemma and a mutual packing lemma also form contributions of this chapter.

In Chapter III, we set out with an aim of constructing an algebraic framework to further improve on the results obtained in Chapter II. We introduce the notion of algebraic structured quantum measurements that exploits the structure that may be inherently present with a distributed problem formulation *Atif and Pradhan* (2021). Toward this, we first recall Unionized Coset Codes (UCC) *Pradhan et al.* (2021) and then construct approximating quantum measurements build using these code ensembles. A peculiar feature of these measurements is that their ensembles are only pairwise independent. Considering this constraint, we develop an alternative covering lemma which facilitates us to first achieve the point-to-point measurement compression result of *Winter* (2004) using the structured measurements. Subsequently, by developing a mutual packing lemma,

we demonstrate the improvement in the rate region compared to the result of Chapter II. We also illustrate these improvements using toy examples to establish the significance of the results.

Chapter IV moves on to analyzing an application of the measurement compression called purity distillation. It discusses the history of purity distillation problem, and details into the interesting advancement made by Devetak in *Devetak* (2005a). Devetak expanded the horizon of LOCC protocols by allowing the use of catalytic pure qubits with the promise of returning these at the end of the protocol. This assumption allowed a measurement to be used by one party and as a result increased the overall rate of purity distillation. The work was soon followed by *Krovi and Devetak* (2007), where the measurement compression protocol was employed to further reduce rate of classical communication required by the protocol. Our objective in this chapter is to devise a protocol that can extract purity from a three party setup. Similar to *Devetak* (2005a), we allow for the use of catalytic pure qubits, and also incorporate the constraint from *Krovi and Devetak* (2007) to minimize the classical communication used by the protocol. As will be discussed in detail, although the protocol is an application of the distributed measurement compression problem (developed in Chapter II), the latter cannot simply be used as a black box. For instance, measurement compression protocols focuses on the post-measured reference state, while the current protocols relies on the post-measured state of the system being measured to extract purity. Similarly, the former allows an additional resource of common randomness while the latter does not *Atif and Pradhan* (2022).

The next chapter, Chapter V, deals with an important problem in quantum information theory: the lossy quantum source coding problem. The objective here is to compress a given quantum source below its von Neumann entropy. Motivated by the results obtained so far, in this chapter, we provide a new formulation for the lossy quantum source coding problem. Another inspiration for this formulation is drawn from the profound duality results and the development of theory revolving around the relation between backward test channels and good source codes for the classical setting *Pradhan* (2004); *Cuff et al.* (2010). A well known existing formulation of this problem is the quantum rate distortion theory, which has been extensively studied in the literature. The formulation we develop differs from the existing quantum rate-distortion theory in two aspects. Firstly, we require that the reconstruction of the compressed quantum source fulfill a global error constraint as opposed to the sample-wise local error criterion used in the standard rate-distortion

setting. A global error criterion is a condition defined on the entire block of data. Secondly, instead of a distortion observable, we employ the notion of a backward quantum channel, which we refer to as a "posterior reference map", to measure the reconstruction error. As the name suggests, this map operates on the reference system of the output of a given channel to produce the reference of the input. The asymptotic performance limit of the lossy quantum source coding problem is characterized in terms of single-letter coherent information of the posterior reference map. We do so by demonstrating a protocol to encode and decode at the specified rate, achieving the asymptotic performance limit while satisfying the global error criterion. The protocol is constructed by decomposing coherent information as a difference of two Holevo information quantities, inspired from prior works in quantum communication problems. This protocol stands on the two basic but fundamental principles of information theory: covering and packing, and demonstrates an exquisite duality with the proof of quantum channel capacity problem constructed in *Devetak* (2005b). Then, we provide a single-letter converse in terms of the above stated coherent information. We also provide various examples to further motivate the formulation, and shed light on its connection to the standard rate-distortion formulation wherever possible.

In the next part of the thesis, we revisit the algebraic structured measurements and investigate their performance for Classical-Quantum (CQ) channel coding problems. In particular, in Chapter VI, we consider the problem of communicating a generic bivariate function of two classical sources observed at the encoders of a classical-quantum multiple access channel. An inner bound to the asymptotic performance characterization of a classical-quantum multiple access channel was first provided in *Winter* (2001). If we employ the approach developed in the latter to compute the given bivariate function, it would require recovering both the sources. However, we ask here if we could recover the function without completely recovering both the sources, and as a result enlarge the rate region for this problem. Building on the techniques developed for the case of a classical channel *Padakandla and Pradhan* (2013), in this chapter, we aim to propose and analyze a coding scheme based on coset codes that enables the decoder recover the desired function without recovering the sources themselves. Toward this, we first develop a Nested Coset Code (NCC) *Pradhan et al.* (2021) based communication scheme for a CQ PTP channel and analyze its performance. Leveraging this building block, we design and analyze the performance of an NCC-based coding scheme for computing sum over a general CQ-MAC *Atif et al.* (2021c). Going further we generalize this idea

for computing arbitrary functions over a general CQ-MAC. We also identify examples where the new set of sufficient conditions are weaker than the current known for this problem.

In the next chapter (Chapter VII), we consider the problem of characterizing an inner bound to the capacity region of a 3-user classical-quantum interference channel (3-CQIC) *Atif et al.* (2021b). The best known coding scheme for communicating over CQICs is based on unstructured random codes and employs the techniques of message splitting and superposition coding. For classical 3-user interference channels (ICs), it has been proven that coding techniques based on coset codes - codes possessing algebraic closure properties - strictly outperform all coding techniques based on unstructured codes. Motivated by this, we consider a special subclass of 3-user CQICs, and analyze their performance employing techniques developed in the previous chapter. We derive a new inner bound to the capacity region of 3to1-CQICs that subsume the current known largest. We also identify examples where we observe strict improvement demonstrating the efficacy of the approach.

During my Phd, I had the opportunity to collaborate on few other problems which are not part of this thesis. These include (i) classical formulation of the measurement compression problem *Atif et al.* (2022b) and (ii) further improvement in the inner bounds for the problem of computing arbitrary bivariate functions over CQ-MAC *Sohail et al.* (2022).

1.5 The Quantum Measurement Compression Problem

Measurements 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 of a quantum state.

Winter's measurement compression theorem *Winter* (2004) (also elaborated in *Wilde et al.* (2012)) quantifies the "relevant information" as the amount of resources needed to faithfully simulate the output of a quantum measurement applied on a given state in an asymptotic sense. In this chapter, we aim to introduce the measurement compression problem *Winter* (2004), provide a brief overview of related work in this regards, and then formally state the problem and the result.

1.5.1 Introduction

The measurement compression problem formulated in *Winter* (2004) is as follows. Imagine that an agent (Alice) performs a measurement M on a quantum state ρ, and sends a set of classical bits to a receiver (Bob). Bob intends to *faithfully* recover the outcomes of Alice's measurements without having access to ρ, while preserving the correlation with the post-measured state of Alice's reference. The major contribution of Winter's work (as elaborated in *Wilde et al.* (2012)) was in specifying an optimal rate region in terms of classical communication and common randomness needed to *faithfully simulate* the action of repeated independent measurements performed on many independent copies of the given quantum state. One of the salient features of the measurement compression theorem is that it achieves the following asymptotic performance. If at least quantum mutual information ($I(X;R)$) amount of classical information and conditional entropy ($S(X|R)$) amount of common shared randomness are available, then one can achieve *faithful simulation* of the measurement M with respect to the quantum state ρ, where R denotes a reference of the quantum state, and X denotes the auxiliary register corresponding to the random measurement outcome.

The measurement compression theorem *Winter* (2004) finds its applications in several quantum paradigms. It is a predecessor to the quantum reverse Shannon theorem *Bennett et al.* (2002); *Berta et al.* (2011); *Bennett et al.* (2014), useful in determining the communication cost of the local purity distillation protocol *Horodecki et al.* (2003a, 2005a); *Devetak* (2005a); *Krovi and Devetak* (2007), and also helpful in the first step of the so-called grandmother protocol *Devetak et al.* (2008) which involves distillation of entanglement from noisy bipartite states.

The measurement compression theorem was also used by Datta, et al. *Datta et al.* (2013b) to develop a QC 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 *Datta et al.* (2013b) advocated the use of the 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.

1.5.1.1 Related Work

Wilde et al. *Wilde et al.* (2012) 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 *Bennett et al.* (2014), 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.

Further, the problem of measurement compression in the presence of quantum side information was also studied in *Wilde et al.* (2012). The authors here combined the ideas from *Winter* (2004) and *Devetak and Winter* (2003) to reduce the classical communication rate and common randomness needed to simulate a measurement in presence of quantum side information.

The problem of quantifying the information gain of a measurement has been studied extensively. Early works include *Groenewold* (1971); *Lindblad* (1972); *Ozawa* (1986). Later on, Buscemi et al. *Buscemi et al.* (2008); *Luo* (2010); *Shirokov* (2011) proposed the quantum mutual information with respect to a classical-quantum state as the measure to characterize the corresponding information gain. Subsequently, Berta et al. *Berta et al.* (2014) provided a universal measurement compression theorem, generalizing the Winter's measurement compression theorem for arbitrary inputs. They identified the quantum mutual information of a measurement as the information gained by performing the measurement, independent of the input state on which it is performed. The proof was based on a new "classically coherent state merging protocol" - a variation of the quantum state merging protocol *Horodecki et al.* (2005b, 2007), and the post-selection technique for quantum channels *Christandl et al.* (2009). Recently, Anshu et al. *Anshu et al.* (2019) considered the problem of measurement compression with side information in the one-shot setting. They presented a protocol employing convex-split lemma for classical-quantum states *Anshu et al.* (2017a, 2014) and position based decoding *Anshu et al.* (2018b), and bounded the communication in terms of smooth max

and hypothesis testing relative entropies. On a similar note, Renes and Renner *Renes and Renner* (2012) studied the problem sending classical messages in the presence of quantum side information in the one-shot setting. We direct an interested reader to *Tomamichel* (2015); *Khatri and Wilde* (2020) for a detailed discussion and results pertaining to one-shot quantum information theory.

Recently, authors in *Anshu et al.* (2019) 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 *Anshu et al.* (2019)).

1.5.2 Preliminaries

We here establish all our notations, and briefly state few necessary definitions. **Notation:** Given any natural number n, let the finite set $\{1, 2, \cdots, n\}$ be denoted by $[1, n]$. 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 $\|A - B\|_1 \triangleq \operatorname{Tr}|A - B|$, where for any operator Λ we define $|\Lambda| \triangleq \sqrt{\Lambda^\dagger \Lambda}$. The von Neumann entropy of a density operator $\rho \in \mathcal{D}(\mathcal{H})$ is denoted by $S(\rho)$. The quantum mutual information 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 *Holevo* (2012), is defined as

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

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:

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

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 Ψ^{ρ}_{RA} denote a purification of a density operator $\rho \in D(\mathcal{H}_A)$. Given a POVM $M \triangleq \{\Lambda^A_x\}_{x\in\mathcal{X}}$ acting on $\rho \in D(\mathcal{H}_A)$, the post-measurement state of the reference together with the classical outputs is represented by

$$(\mathrm{id} \otimes M)(\Psi^{\rho}_{RA}) \triangleq \sum_{x\in\mathcal{X}} |x\rangle\langle x| \otimes \mathrm{Tr}_A\{(I^R \otimes \Lambda^A_x)\Psi^{\rho}_{RA}\}. \tag{1.1}$$

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

1.5.3 Problem Statement

Measurement compression theorem quantifies the "relevant information" of a measurement M by measuring the minimum amount of classical information bits needed to "simulate" the repeated action of M on a quantum state ρ. 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 approximating measurement $\tilde{M}^{(n)}$ is faithful to the actual measurement $M^{\otimes n}$ with respect to the state $\rho^{\otimes n}$. This is formally defined in the following.

Definition I.1 (Faithful simulation *Wilde et al.* (2012))**.** Given a sub-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 ϵ-faithful to M with respect to ρ, for $\epsilon > 0$, if the following holds:

$$\begin{aligned}\Xi_\rho(M, \tilde{M}) \triangleq &\sum_{x\in\mathcal{X}} \left\| \sqrt{\rho}(\Lambda_x - \tilde{\Lambda}_x)\sqrt{\rho} \right\|_1 \\ &+ \mathrm{Tr}\left\{(I - \sum_x \Lambda_x)\rho\right\} + \mathrm{Tr}\left\{(I - \sum_x \tilde{\Lambda}_x)\rho\right\} \leq \epsilon.\end{aligned} \tag{1.2}$$

Alternatively, one can complete the POVMs M and $\tilde{M}$ by associating $I - \sum_{x\in\mathcal{X}} \Lambda_x$ and $I - \sum_{x\in\mathcal{X}} \tilde{\Lambda}_x$ with additional symbols 0 and $\tilde{0}$, respectively, and thus obtaining POVMs $[M]$ and $[\tilde{M}]$, defined on $\mathcal{X} \bigcup \{0, \tilde{0}\}$. Stating the above definition for $[M]$ and $[\tilde{M}]$ gives the same as in (*Wilde et al.*, 2012, Definition 3). Further, the above trace norm constraint can be equivalently expressed in terms of a purification of state ρ using the following lemma.

Lemma I.2. (*Wilde et al.*, 2012, Lemma 4) *For any state $\rho \in \mathcal{D}(\mathcal{H})$ with any purification Ψ^{ρ}_{RA}, and any pair of POVMs M and $\tilde{M}$ acting on $\mathcal{H}$, the following identity holds*

$$\|(\mathrm{id} \otimes M)(\Psi^{\rho}_{RA}) - (\mathrm{id} \otimes \tilde{M})(\Psi^{\rho}_{RA})\|_1 = \sum_x \|\sqrt{\rho}(\Lambda_x - \tilde{\Lambda}_x)\sqrt{\rho}\|_1, \tag{1.3}$$

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

Theorem I.3. (*Winter*, 2004, Theorem 2) *For any $\epsilon > 0$, any density operator $\rho \in \mathcal{D}(\mathcal{H}_A)$, any POVM M acting on the Hilbert space $\mathcal{H}_A$, and for all sufficiently large n, there exists 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 ϵ-faithful to $M^{\otimes n}$ with respect to $\rho^{\otimes n}$ if*

$$R > I(U;R)_\sigma, \quad \text{and} \quad \frac{1}{n}\log_2 N + R > S(U)_\sigma,$$

where $\sigma_{RU} \triangleq (\mathrm{id} \otimes M)(\Psi^{\rho}_{RA})$.

Remark I.4. A strong converse of the above result is also provided in (*Winter*, 2004, Theorem 8).

1.5.4 Conclusion

This section briefly introduced the measurement compression problem. The aim of the next chapters is to consider the problem in distributed settings and obtain achievable rate regions toward faithful simulation of distributed quantum measurements. The subsequent chapters form the contribution of this thesis.

Part I

Quantum Measurement Compression

CHAPTER II

Distributed Measurement Compression

In this chapter, 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. 2.1, 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 Bob. Imagine that there is a third party, Charlie, 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 Charlie at rate R_1 and R_2, respectively. Further, pairwise common randomness at rates C_1 and C_2 are also shared between Alice and Charlie, and Bob and Charlie, respectively. Charlie performs classical processing of the received bits and common randomness. We study two settings, based on whether or not Charlie has access to private randomness, namely the feedback case and the non-feedback case, respectively. The private randomness can enable Charlie to employ random stochastic decoders (aka probabilistic decoders) for decoding the measurement outcomes. As an application of this quantification, we consider the QC distributed rate distortion problem where Charlie is allowed to use classical quantum (CQ) channels. In this chapter, we focus on memoryless quantum systems in finite-dimensional Hilbert spaces.

The contributions of this chapter can be summarized as follows:

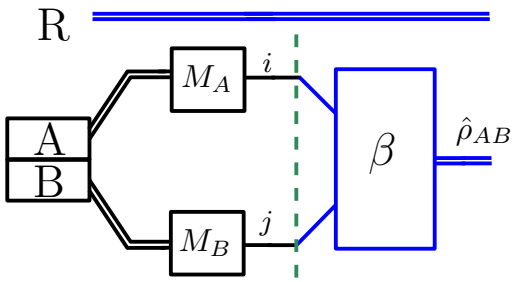

Figure 2.1: The diagram of a distributed quantum measurement applied to a bipartite quantum system AB. A tensor product measurement $M_A \otimes M_B$ is performed on many copies of the observed quantum state. The outcomes of the measurements are given by two classical bit streams. The receiver functions as a classical-to-quantum channel β mapping the classical data to a quantum state.

- We formulate the problem of faithful simulation of distributed quantum measurements that can be decomposed as a convex-linear combination (incorporating Charlie's stochastic processing) of separable measurements, as stated in Definition II.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_1 and C_2, 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 II.33). This is the first main result of the paper.

- In the special case of the above problem formulation, where the Charlie's action is restricted to a deterministic mapping, we develop a one-shot performance characterization of the distributed faithful simulation problem (see Theorem II.4). This characterization is based on a modular approach. As a corollary to this result, we develop a characterization of an inner bound to the asymptotic performance limit (see Theorem II.6).

- 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 Charlie, 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 II.20). This is the second main result of the paper. The classical version of this result is called the Berger-Tung inner bound *Berger* (1977).

- 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 II.22), by converting a multi-letter expression into a single-letter expression. This is the third main result of the paper.

As was pointed out in *Krovi and Devetak* (2007), the measurement compression theorem *Winter* (2004) is a generalization of the classical reverse Shannon theorem [14] and can be viewed as a quantum-to-classical channel simulation problem. Similarly, the distributed measurement compression problem addressed in this chapter can be viewed as a distributed multi-party quantum-to-classical channel simulation problem. This can pave the way to considering the multi-party extensions of problems such as entanglement distillation and remote state preparation. In fact, as will be shown in IV, we use the distributed measurement compression protocol to construct a multi-party purity distillation protocol. Further, this work also develops new tools such as the mutual covering lemma and the mutual packing lemma which can be promising tools for many emerging quantum network applications. Moreover, in the recent applications of the distributed paradigms, a network of limited qubit-capacity quantum computers, connected through classical and quantum channels, are used to solve problems in a distributed manner by casting known centralized algorithms into their distributed versions *Yimsiriwattana and Lomonaco Jr* (2004); *Beals et al.* (2013); *Van Meter and Devitt* (2016); *Denchev and Pandurangan* (2008).

The organization of this chapter is as follows. In Section 1.5.2, we set the notation and state requisite definitions. Toward developing the main result of this chapter, for pedagogical reasons, we first consider a special case in Section 2.2.2. For this special case, we restrict the processing at Charlie to a deterministic function and characterize the performance of a faithful simulation protocol in a one-shot setting. We achieve this by first obtaining a one-shot measurement compression theorem in a point-to-point setting (Theorem II.9), wherein Bob is absent. Then we employ this

result on the individual components (M_A and M_B) of the joint measurement M_{AB}, separately, to obtain a theorem characterizing the performance of a distributed measurement compression protocol (see Theorem II.4). As a corollary, we further provide an asymptotic quantum information-theoretic inner bound to the achievable rate region of the distributed measurement compression problem (see Theorem II.6). 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 Charlie. 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 Charlie 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 II.12), which also helps in converting the information quantities in terms of the reference R of the joint state ρ_{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 ρ_{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 II.15), 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-Ahlswede-Körner *Berger* (1977) by developing the Markov lemma and the Mutual packing lemma. The idea of binning in quantum setting has been explored from a different perspective in *Devetak and Winter* (2003) and *Anshu et al.* (2018a) 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 2.3.3, we apply this special setting of the distributed measurement simulation with deterministic processing to the QC 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 2.4, we consider the non-feedback measurement compression problem for the point-to-point setting. The authors in *Wilde et al.* (2012) have discussed this formulation and provided a rate region with a proof of achievability and converse. However, in their proof, the authors assume two inequalities (*Wilde et al.*, 2012, Eq. 53 and 54), which may not necessarily be true *Wilde* (Aug 2019) (further details are provided in Section 2.4). A stronger version of this theorem is also developed in *Berta et al.* (2014) using a different technique, wherein the authors have extended the Winter's measurement compression for fixed independent and identically distributed inputs *Winter* (2004) to arbitrary inputs. Since the result is crucial for the distributed simulation problem with stochastic processing, to be proved in the next section (Section 2.5.2), we formally state the problem and provide an alternative proof of the direct part for completeness (see Theorem II.27).

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 2.2, 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 II.12). However, it will be shown that, because of the presence of Charlie'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. Toward this we develop a non-product covering lemma (see Proposition II.41). Moreover, we provide a single-letter achievable inner bound that is symmetric with respect to Alice and Bob. We conclude the chapter with a few remarks in Section 2.6.

2.1 Notations and Definitions

Definition II.1 (Joint Measurements)**.** A POVM $M_{AB} \triangleq \{\Lambda_z^{AB}\}_{z\in\mathcal{Z}}$, 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 $M_A \triangleq \{\Lambda_u^A\}_{u\in\mathcal{U}}$ and $M_B \triangleq \{\Lambda_v^B\}_{v\in\mathcal{V}}$ and a stochastic mapping $P_{Z|U,V} : \mathcal{U} \times \mathcal{V} \to \mathcal{Z}$ such that

$$\Lambda_z^{AB} \triangleq \sum_{u,v} P_{Z|U,V}(z|u,v)\Lambda_u^A \otimes \Lambda_v^B, \quad \forall z \in \mathcal{Z},$$

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.

2.2 One Shot Distributed Simulation of POVMS with deterministic processing

We begin by considering the simulation of distributed POVMs with deterministic processing. Recall from the discussion in Section 1.5.1 that the motivation behind the restriction to deterministic processing is that the proof becomes modular, and also forms a first pedagogical step towards the distributed simulation with stochastic processing (Theorem II.33). Due to the modularity of the proof, we were able to develop a one-shot version of the proof. In the following, we state the problem formulation and provide the theorem statement.

2.2.1 Problem Formulation

In this formulation, Charlie's processing is restricted to a deterministic mapping. More precisely,

Definition II.2 (Distributed Protocol)**.** For a given finite set $\mathcal{Z}$, and a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, a distributed protocol with deterministic processing with parameters $(n, \Theta_1, \Theta_2, N_1, N_2)$ is characterized by

1) a collection of Alice's sub-POVMs $\tilde{M}_A^{(\mu_1)}, \mu_1 \in [1, N_1]$ each acting on $\mathcal{H}_A^{\otimes n}$ and with outcomes in $[1, \Theta_1]$.

2) a collection of Bob's sub-POVMs $\tilde{M}_B^{(\mu_2)}, \mu_2 \in [1, N_2]$ each acting on $\mathcal{H}_B^{\otimes n}$ and with outcomes in $[1, \Theta_2]$.

3) a collection of Charlie's classical decoding maps $f^{(\mu_1,\mu_2)} : [1, \Theta_1] \times [1, \Theta_2] \to \mathcal{Z}^n$ for $\mu_1 \in [1, N_1], \mu_2 \in [1, N_2]$.

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

$$\tilde{\Lambda}_{z^n} \triangleq \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{\substack{l_1 \in [1,\Theta], \\ l_2 \in [1,\Theta]}} \mathbb{1}_{\{f^{(\mu_1,\mu_2)}(l_1,l_2)=z^n\}} \Lambda_{l_1}^{A,(\mu_1)} \otimes \Lambda_{l_2}^{B,(\mu_2)} \tag{2.1}$$

$\forall z^n \in \mathcal{Z}^n$, where $\Lambda_{l_1}^{A,(\mu_1)}$ and $\Lambda_{l_2}^{B,(\mu_2)}$ are the operators corresponding to the sub-POVMs $\tilde{M}_A^{(\mu_1)}$ and $\tilde{M}_B^{(\mu_2)}$, respectively.

In the above definition, (Θ_1, Θ_2) determines the amount of classical bits communicated from Alice and Bob to Charlie, respectively. N_1 and N_2 denote the amount of pairwise common randomness. The classical maps $f^{(\mu_1,\mu_2)}$ represent the action of Charlie on the received classical bits.

Definition II.3 (Achievability)**.** 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 quadruple (R_1, R_2, C_1, C_2) is said to be achievable, if for all $\epsilon > 0$ and for all sufficiently large n, there exists a distributed protocol with deterministic processing with parameters $(n, \Theta_1, \Theta_2, N_1, N_2)$ such that its overall sub-POVM $\tilde{M}_{AB}$ is ϵ-faithful to $M_{AB}^{\otimes n}$ with respect to $\rho_{AB}^{\otimes n}$ (see Definition I.1), and

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

The set of all achievable quadruples (R_1, R_2, C_1, C_2) is called the achievable rate region.

2.2.2 Inner Bounds

We now provide two theorems characterizing the performance of faithful simulation protocols, one in a one-shot and the other in an asymptotic quantum information theoretic settings for faithful simulation of distributed measurements with deterministic processing. The proofs of these theorems are provided in Section 2.2.4 and 2.2.5.

Theorem II.4. *(One-shot Distributed Faithful Simulation). Consider a density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and a sub-POVM $M_{AB} \triangleq \{\Lambda_u^A \otimes \Lambda_v^B\}_{u\in\mathcal{U},v\in\mathcal{V}}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$. Suppose there exists total subspace projectors Π_{ρ_A}, Π_{ρ_B}, and codeword subspace projectors $\{\Pi_u^A\}_{u\in\mathcal{U}}$, $\{\Pi_v^A\}_{v\in\mathcal{V}}$, acting on $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, satisfying:*

$$\mathrm{Tr}\{\Pi_{\rho_A}\hat{\rho}_u^A\} \geq 1-\epsilon_1,\ \mathrm{Tr}\{\Pi_{\rho_B}\hat{\rho}_v^B\} \geq 1-\epsilon_2,\ \mathrm{Tr}\{\Pi_u^A\hat{\rho}_u^A\} \geq 1-\epsilon_1,\ \mathrm{Tr}\{\Pi_v^B\hat{\rho}_v^B\} \geq 1-\epsilon_2, \tag{2.2a}$$

$$\mathrm{Tr}\{\Pi_{\rho_A}\} \leq D_1, \quad \mathrm{Tr}\{\Pi_{\rho_B}\} \leq D_2, \quad \Pi_u^A\hat{\rho}_u^A\Pi_u^A \leq \frac{1}{d_1}\Pi_u^A, \quad \Pi_v^B\hat{\rho}_v^B\Pi_v^B \leq \frac{1}{d_2}\Pi_v^B, \tag{2.2b}$$

$$\Pi_{\rho_A}\rho_A\Pi_{\rho_A} \leq \rho_A, \quad \Pi_{\rho_B}\rho_B\Pi_{\rho_B} \leq \rho_B, \quad \Pi_u^A\hat{\rho}_u^A\Pi_u^A \leq \hat{\rho}_u^A, \quad \Pi_v^B\hat{\rho}_v^B\Pi_v^B \leq \hat{\rho}_v^B, \tag{2.2c}$$

$$\Pi_{\rho_A}\rho_A\Pi_{\rho_A} \leq \frac{1}{F_1}\Pi_{\rho_A},\ \sqrt{\rho_A}^{-1}\Pi_{\rho_A}\sqrt{\rho_A}^{-1} \leq f_1\Pi_{\rho_A},$$

$$\Pi_{\rho_B}\rho_B\Pi_{\rho_B} \leq \frac{1}{F_2}\Pi_{\rho_B},\ \sqrt{\rho_B}^{-1}\Pi_{\rho_B}\sqrt{\rho_B}^{-1} \leq f_2\Pi_{\rho_B}, \tag{2.2d}$$

where $\epsilon_i \in (0, \frac{1}{2})$, $0 < d_i < D_i$, *and* $f_i, F_i > 0$ *for* $i = 1, 2$, *and* $\hat{\rho}_u^A$ *and* $\hat{\rho}_v^B$ *are defined in* (2.37). *Let* $\mathcal{W} \subseteq \mathcal{U} \times \mathcal{V}$ *be an arbitrary set. Let* K_1 *and* K_2 *be arbitrary positive integers such that* $|\mathcal{U}| \geq K_1$ *and* $|\mathcal{V}| \geq K_2$. *Then for any* $T_i \leq K_i$ *for* $i = 1, 2$, *there exists a distributed protocol with deterministic processing for the finite set* $\mathcal{U} \times \mathcal{V}$ *and the Hilbert space* $\mathcal{H}_A \otimes \mathcal{H}_B$, *with parameters* $(1, T_1, T_2, N_1, N_2)$ *such that*

$$\Xi_{\rho^{AB}}(M_{AB}, \tilde{M}_{AB}) \leq \alpha_A + \alpha_B + \alpha_P,$$

where

$$\alpha_A(\epsilon_1, N_1, K_1) \triangleq \frac{2}{(1+\epsilon_1)\sqrt{N_1 K_1}} \sum_{u \in \mathcal{U}} \sqrt{\lambda_u^A} + \frac{2\epsilon_1}{\epsilon_1 + 1} + f(\epsilon_1, \theta_1) + 4 D_1 N_1 \exp\left[-\frac{K_1 \epsilon_1^3 d_1 D_1^{-1}}{4 \ln 2} \right] + 2\theta_1, \tag{2.3}$$

$$\alpha_B(\epsilon_2, N_2, K_2) \triangleq \frac{2}{(1+\epsilon_2)\sqrt{N_2 K_2}} \sum_{v \in \mathcal{V}} \sqrt{\lambda_v^B} + \frac{2\epsilon_2}{\epsilon_2 + 1} + f(\epsilon_2, \theta_2) + 4 D_2 N_2 \exp\left[-\frac{K_2 \epsilon_2^3 d_2 D_2^{-1}}{4 \ln 2} \right] + 2\theta_2, \tag{2.4}$$

$$\begin{aligned} \alpha_P(\epsilon_1, \epsilon_2, K_1, K_2, N_1, N_2, T_1, T_2, \mathcal{W}) &\triangleq 2\alpha_A + 2\alpha_B + 2\lambda^{AB}(\mathcal{W}^c) + \frac{2}{(1+\epsilon_1)(1+\epsilon_2)} \\ &\times \left[\frac{\lambda_m^A \lambda_m^B |\mathcal{W}| K_1 K_2}{(1-\theta_1)(1-\theta_2) T_1 T_2} + \frac{K_1 W_A \lambda_m^A}{(1-\theta_1) T_1} \left(1 + \frac{\lambda_m^B K_2}{(1-\theta_2)} \right) + \frac{K_2 W_B \lambda_m^B}{(1-\theta_2) T_2} \left(1 + \frac{\lambda_m^A K_1}{(1-\theta_1)} \right) \right] \frac{f_1 f_2}{F_1 F_2}, \end{aligned} \tag{2.5}$$

and $\lambda_{u,v}^{AB} \triangleq \mathrm{Tr}\big(\rho^{AB}(\Lambda_u \otimes \Lambda_v)\big)$ *with marginals* $(\lambda_u^A, \lambda_v^B)$, $W_A \triangleq \max_{v \in \mathcal{V}} |\{u : (u, v) \in \mathcal{W}\}|$, *and* $W_B \triangleq \max_{u \in \mathcal{U}} |\{v : (u, v) \in \mathcal{W}\}|$, $\lambda_m^A \triangleq \max_u \lambda_u^A$, *and* $\lambda_m^B \triangleq \max_v \lambda_v^B$, $\theta_1 \triangleq 1 - \sum_{u \in \mathcal{U}} \lambda_u^A$, $\theta_2 \triangleq 1 - \sum_{v \in \mathcal{V}} \lambda_v^B$, $f(\epsilon, \theta) \triangleq \left[4\sqrt{\epsilon} + 4\sqrt{\epsilon + 2\sqrt{\epsilon}} + 4\sqrt{2}(1-\theta)\sqrt{\epsilon + \sqrt{\epsilon}} \right] / (1+\epsilon)$, *and* $\lambda^{AB}(\mathcal{W}^c) \triangleq \sum_{(u,v) \in \mathcal{W}^c} \lambda_{u,v}^{AB}$.

Remark II.5. Note that the terms α_A and α_B can be identified as the one-shot expressions for the errors induced in approximating each of the sub-POVMs $\{\Lambda_u^A\}_{u \in \mathcal{U}}$ and $\{\Lambda_v^B\}_{v \in \mathcal{V}}$, using their respective approximations. This approximation employs the one-shot version of the measurement compression theorem (Theorem II.9), which is developed as a part of the proof in Section 2.2.4. Within α_A, the exponential term corresponds to the error probability that the approximating operators do not constitute a valid sub-POVMs in random coding, the term involving square-root of the probabilities corresponds to the classical soft covering error, and the term $f(\epsilon, \theta)$ corresponds

to the error incurred because of the use of gentle measurement lemma with regard to the total subspace and codeword subspace projectors. Likewise, the term α_P captures the additional error introduced by compressing the classical outcomes of the above distributed measurement using the technique of binning. The binning is used to reduce the rate of transmission by exploiting the classical correlations present in the measurement outcomes, using a many-to-one transformation. The information lost in this transformation is recovered at the receiver using a relation modeled by a bipartite sub-graph $\mathcal{W}$ of $\mathcal{U} \times \mathcal{V}$. The twice of $\alpha_A + \alpha_B$ within α_P captures the effect of binning on the event corresponding to not being able to cover the sources using the approximating sub-POVMs. $\lambda^{AB}(\mathcal{W}^c)$ captures the event where under the original sub-POVM, the measurement outcomes do not satisfy the above set relation. The final term captures the error due to binning of the approximating sub-POVMs.

As a corollary to the above theorem, we obtain the following asymptotic inner bound to the achievable rate region.

Theorem II.6 (Distributed Faithful Simulation)**.** *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 $\mathcal{H}_A \otimes \mathcal{H}_B$, and having a separable decomposition with deterministic integration (as in Definition II.1), a quadruple (R_1, R_2, C_1, C_2) is achievable if the following inequalities are satisfied:*

$$R_1 \geq I(U;RB)_{\sigma_1} - I(U;V)_{\sigma_3}, \tag{2.6a}$$

$$R_2 \geq I(V;RA)_{\sigma_2} - I(U;V)_{\sigma_3}, \tag{2.6b}$$

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

$$R_1 + C_1 \geq S(U|V)_{\sigma_3}, \tag{2.6c}$$

$$R_2 + C_2 \geq S(V|U)_{\sigma_3}, \tag{2.6d}$$

$$R_1 + R_2 + C_1 \geq I(V;RA)_{\sigma_2} + S(U|V)_{\sigma_3}, \tag{2.6e}$$

$$R_1 + R_2 + C_2 \geq I(U;RB)_{\sigma_1} + S(V|U)_{\sigma_3}, \tag{2.6f}$$

$$R_1 + R_2 + C_1 + C_2 \geq S(U,V)_{\sigma_3}, \tag{2.6g}$$

for some decomposition with POVMs $M_A = \{\Lambda_u^A\}_{u\in\mathcal{U}}$ and $M_B = \{\Lambda_v^B\}_{v\in\mathcal{V}}$ and a function $g : \mathcal{U} \times \mathcal{V} \to \mathcal{Z}$, where the information quantities are computed for the auxiliary states $\sigma_1^{RUB} \triangleq (\mathrm{id}_R \otimes$

$M_A \otimes \mathrm{id}_B)(\Psi_{RAB}^{\rho_{AB}})$, $\sigma_2^{RAV} \triangleq (\mathrm{id}_R \otimes \mathrm{id}_A \otimes M_B)(\Psi_{RAB}^{\rho_{AB}})$, *and* $\sigma_3^{RUV} \triangleq (\mathrm{id}_R \otimes M_A \otimes M_B)(\Psi_{RAB}^{\rho_{AB}})$, *with* $\Psi_{RAB}^{\rho_{AB}}$ *being a purification of* ρ_{AB}.

Remark II.7. An alternative characterization of the above rate region can be obtained in terms of Holevo information. Using the canonical ensemble, we obtain

$$\begin{aligned} I(U;RB)_{\sigma_1} &= S(RB)_{\sigma_1} - S(RB|U)_{\sigma_1} \\ &= S(\sum_{u\in\mathcal{U}} \lambda_u^A \hat{\rho}_u^A) - \sum_{u\in\mathcal{U}} \lambda_u^A S(\hat{\rho}_u^A) = \chi\left(\{\lambda_u^A, \hat{\rho}_u^A\}\right), \end{aligned}$$

where the second equality follows by noting $S(RB)_{\sigma_1} = S(\rho_A)$, $\rho_A = \sum_{u\in\mathcal{U}} \lambda_u^A \hat{\rho}_u^A$, and using the result from (*Wilde*, 2013a, Eq. 11.54). Similarly, we get $I(V;RA)_{\sigma_2} = \chi\left(\{\lambda_v^B, \hat{\rho}_v^B\}\right)$. 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.

2.2.3 Overview of Proof Technique and an Illustrative Example

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. To reduce the number of free parameters, let $C \triangleq C_1 + C_2$. Firstly, consider the regime where the sum rate $(R_1 + R_2)$ is at its minimum achievable, i.e., equation (2.6c) 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 *Slepian and Wolf* (July 1973). Fig. 2.2 demonstrates the achievable rate region in these cases.

We encounter two challenges in developing the single-letter inner bound to the achievable rate region as stated in Theorem II.6: 1) The direct use of single-POVM compression theorem, proved using random coding arguments as in *Winter* (2004), 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 II.12).

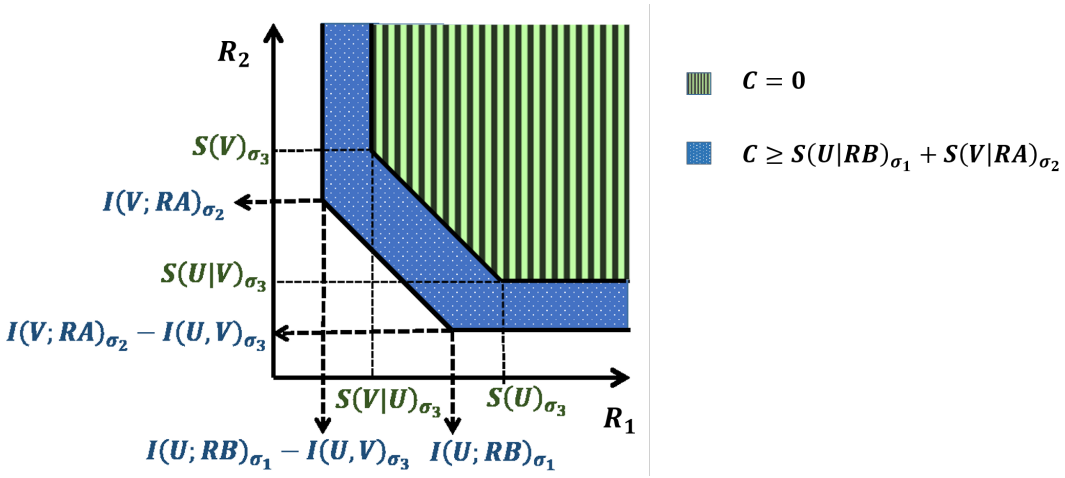

Figure 2.2: The inner bound to the achievable rate region given in Theorem II.6 at two planes: 1) with no common randomness, i.e., $C = 0$ (green color), and 2) with at least $S(U|RB)_{\sigma_1} + S(V|RA)_{\sigma_2}$ amount of common randomness (blue color). As a result, the latter region contains the former.

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 *Slepian and Wolf* (July 1973) (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-Ahlswede-Korner by developing the Markov Lemma and the Mutual Packing Lemma (Lemma 12.1 and 12.2 in *El Gamal and Kim* (2011)). 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 2.5.3.3 and Proposition II.15).

Let us consider an example to illustrate the above inner bound.

Example II.8. Suppose the composite state ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$ is described as

$$\rho^{AB} \triangleq \frac{1}{2}\left(|00\rangle\langle 00|_{AB} + |11\rangle\langle 11|_{AB}\right).$$

Since $\pi^A = \mathrm{Tr}_B\, \rho^{AB}$ and $\pi^B = \mathrm{Tr}_A\, \rho^{AB}$, 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 $\left\{\frac{1}{2}|0\rangle\langle 0|, \frac{1}{2}|1\rangle\langle 1|, \frac{1}{2}|+\rangle\langle +|, \frac{1}{2}|-\rangle\langle -|\right\}$. Alice and Bob together with Charlie 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 II.6. Let Ψ_ρ^{RAB}, the purification of ρ^{AB}, be defined as

$$\Psi_\rho^{RAB} \triangleq \left(\frac{|000\rangle_{RAB} + |111\rangle_{RAB}}{\sqrt{2}}\right)\left(\frac{\langle 000|_{RAB} + \langle 111|_{RAB}}{\sqrt{2}}\right).$$

Considering the first constraint from (2.6a), we evaluate σ_1^{RUB} as

$$\begin{aligned}\sigma_1^{RUB} = \frac{1}{4}\Bigg(& |0\rangle\langle 0|_U \otimes |00\rangle\langle 00|_{RB} + |1\rangle\langle 1|_U \otimes |11\rangle\langle 11|_{RB} \\ & + |2\rangle\langle 2|_U \otimes \left(\frac{|00\rangle_{RB} + |11\rangle_{RB}}{\sqrt{2}}\right)\left(\frac{\langle 00|_{RB} + \langle 11|_{RB}}{\sqrt{2}}\right) \\ & + |3\rangle\langle 3|_U \otimes \left(\frac{|00\rangle_{RB} - |11\rangle_{RB}}{\sqrt{2}}\right)\left(\frac{\langle 00|_{RB} - \langle 11|_{RB}}{\sqrt{2}}\right)\Bigg),\end{aligned}$$

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

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

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 σ_3^{UV} as

$$\begin{aligned}\sigma_3^{UV} = \frac{1}{16}\Bigg(& \sum_{i=0}^{3}\sum_{j=0}^{3} |i\rangle\langle i|_U \otimes |i\rangle\langle i|_V + \Big(|0\rangle\langle 0|_U \otimes |0\rangle\langle 0|_V + |1\rangle\langle 1|_U \otimes |1\rangle\langle 1|_V\Big) \\ & - \Big(|0\rangle\langle 0|_U \otimes |1\rangle\langle 1|_V + |1\rangle\langle 1|_U \otimes |0\rangle\langle 0|_V\Big)\Bigg),\end{aligned}$$

which gives

$$S(U,V)_{\sigma_3} = 3.75 \quad \text{and} \quad I(U;V)_{\sigma_3} = 0.25.$$

Therefore, we can write the constraints given in Theorem II.6 as

$$R_1 \geq 0.75, \quad R_2 \geq 0.75, \quad R_1 + R_2 \geq 1.75, \quad R_1 + C_1 \geq 1.75,$$
$$R_2 + C_2 \geq 1.75, \quad R_1 + R_2 + C_1 \geq 2.75,$$
$$R_1 + R_2 + C_2 \geq 2.75, \quad \text{and} \quad R_1 + R_2 + C_1 + C_2 \geq 3.75.$$

Consider the case when $C = C_1 + C_2 \geq 2$ 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 (2.5.3.3)), gives an additional gain of a quarter bit, which is characterized by $I(U;V)_{\sigma_3} = 0.25$, thus giving us the achievable sum-rate of 1.75 bits.

2.2.4 Proof of One-Shot Inner Bound (Theorem II.4)

We begin the proof of the theorem by restating the measurement compression theorem (Theorem I.3) in a one-shot quantum information theoretic setting. This restatement allows us to develop a one-shot mutual covering lemma, which is a crucial part of the current proof. The theorem is stated as follows:

Theorem II.9. *(One-shot Point-to-point Faithful Simulation). Consider a density operator $\rho \in \mathcal{D}(\mathcal{H})$ and a sub-POVM $M \triangleq \{\Lambda_x\}_{x\in\mathcal{X}}$ acting on $\mathcal{H}$, and let $\{\lambda_x, \hat{\rho}_x\}_{x\in\mathcal{X}}$ be the canonical ensemble[1] of M with respect to ρ. Suppose there exists a total subspace projector Π_ρ and codeword subspace projectors $\{\Pi_x\}_{x\in\mathcal{X}}$ acting on $\mathcal{H}$ satisfying:*

$$\mathrm{Tr}\{\Pi_\rho \hat{\rho}_x\} \geq 1 - \epsilon \tag{2.7a}$$
$$\mathrm{Tr}\{\Pi_x \hat{\rho}_x\} \geq 1 - \epsilon \tag{2.7b}$$
$$\mathrm{Tr}\{\Pi_\rho\} \leq D \tag{2.7c}$$
$$\Pi_x \hat{\rho}_x \Pi_x \leq \frac{1}{d}\Pi_x \tag{2.7d}$$
$$\Pi_x \hat{\rho}_x \Pi_x \leq \hat{\rho}_x \tag{2.7e}$$
$$\Pi_\rho \rho \Pi_\rho \leq \rho, \tag{2.7f}$$

[1]Note that $\{\lambda_x\}_{x\in\mathcal{X}}$ is a sub-probability vector, i.e., a vector of non-negative real numbers whose sum is not greater than 1.

where $\epsilon \in (0, \frac{1}{2})$, $0 < d < D$. *Then there exists a collection of sub-POVMs* $\tilde{M}^{(\mu)}$ *for* $\mu \in [1, N]$ *each with at most* K *outcomes, with* $K \leq |\mathcal{X}|$, *and acting on* $\mathcal{H}$ *such that*

$$\Xi_\rho(M, \tilde{M}) \leq \frac{2}{(1+\epsilon)\sqrt{NK}} \sum_{x \in \mathcal{X}} \sqrt{\lambda_x} + \frac{2\epsilon}{\epsilon + 1} f(\epsilon, \theta) + 4DN \exp\left[-\frac{K\epsilon^3 d D^{-1}}{4 \ln 2}\right] + 2\theta,$$

where $\tilde{M} \triangleq \frac{1}{N} \sum_\mu \tilde{M}^{(\mu)}$, $\theta \triangleq 1 - \sum_{x \in \mathcal{X}} \lambda_x$, *and*

$$f(\epsilon, \theta) \triangleq \left[4\sqrt{\epsilon} + 4\sqrt{\epsilon + 2\sqrt{\epsilon}} + 4\sqrt{2}(1-\theta)\sqrt{\epsilon + \sqrt{\epsilon}}\right] / (1 - \epsilon).$$

Proof. The proof is provided in Appendix A.1. □

Moving ahead with the proof of the current theorem, assume that the operators of the original sub-POVM $M_{AB} = M_A \otimes M_B$ are denoted by $\{\Lambda_u^A\}_{u \in \mathcal{U}}$ and $\{\Lambda_v^B\}_{v \in \mathcal{V}}$, respectively, where $\mathcal{U}$ and $\mathcal{V}$ are two finite sets. The proof follows by constructing a protocol for faithful simulation of $M_A \otimes M_B$. We start by generating the canonical ensembles[2] corresponding to M_A and M_B. Let Π_{ρ_A} and Π_{ρ_B} denote the total projectors for marginal density operators ρ_A and ρ_B, respectively. Also, for any $u \in \mathcal{U}$ and $v \in \mathcal{V}$, let Π_u^A and Π_v^B denote the codeword projectors. Let the canonical ensembles be $\{\lambda_u^A, \hat{\rho}_u^A\}$ and $\{\lambda_v^B, \hat{\rho}_v^B\}$. For each $u \in \mathcal{U}$ and $v \in \mathcal{V}$ define

$$\tilde{\rho}_u^{A'} \triangleq \Pi_{\rho_A} \Pi_u^A \hat{\rho}_u^A \Pi_u^A \Pi_{\rho_A}, \quad \tilde{\rho}_v^{B'} \triangleq \Pi_{\rho_B} \Pi_v^B \hat{\rho}_v^B \Pi_v^B \Pi_{\rho_B}. \tag{2.8}$$

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

$$\sigma^{A'} \triangleq \frac{1}{(1-\theta_1)} \sum_{u \in \mathcal{U}} \lambda_u^A \tilde{\rho}_u^{A'}, \quad \sigma^{B'} \triangleq \frac{1}{(1-\theta_2)} \sum_{v \in \mathcal{V}} \lambda_v^B \tilde{\rho}_v^{B'}. \tag{2.9}$$

Let $\hat{\Pi}^A$ and $\hat{\Pi}^B$ be the projectors onto the subspaces spanned by the eigenstates of $\sigma^{A'}$ and $\sigma^{B'}$ corresponding to eigenvalues that are larger than ϵ_1 / D_A and ϵ_2 / D_B, respectively. Lastly, define

$$\tilde{\rho}_u^A \triangleq \hat{\Pi}^A \tilde{\rho}_u^{A'} \hat{\Pi}^A, \quad \text{and} \quad \tilde{\rho}_v^B \triangleq \hat{\Pi}^B \tilde{\rho}_v^{B'} \hat{\Pi}^B, \tag{2.10}$$

for all $u \in \mathcal{U}$, and $v \in \mathcal{V}$ and $\sigma^A = \hat{\Pi}^A \sigma^{A'} \hat{\Pi}^A$, $\sigma^B = \hat{\Pi}^B \sigma^{B'} \hat{\Pi}^B$.

[2]Note that $\{\lambda_u^A\}_{u \in \mathcal{U}}$ and $\{\lambda_v^B\}_{v \in \mathcal{V}}$ are sub-probability vectors.

2.2.4.1 Construction of Random POVMs

In what follows, we construct two random POVMs one for each encoder. Fix positive integers K_1, K_2, N_1 and N_2. 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. For each $\mu_1 \in [1, N_1]$ and $\mu_2 \in [1, N_2]$, randomly and independently select $K_1 \times K_2$ pairs denoted by $(U^{(\mu_1)}(l), V^{(\mu_2)}(k))$ from the set $\mathcal{U} \times \mathcal{V}$ according to the distribution:

$$\mathbb{P}\left((U^{(\mu_1)}(l), V^{(\mu_2)}(k)) = (u, v)\right) = \frac{\lambda_u^A \lambda_v^B}{(1-\theta_1)(1-\theta_2)}, \tag{2.11}$$

for $u \in \mathcal{U}, v \in \mathcal{V}$. Let $\mathcal{C}^{(\mu_1,\mu_2)}$ denote the collection $\{U^{(\mu_1)}(l), V^{(\mu_2)}(k)\}_{\{l\in[1,K_1],k\in[1,K_2]\}}$. Construct operators[3]

$$A_u^{(\mu_1)} \triangleq \gamma_u^{(\mu_1)}\left(\sqrt{\rho_A}^{-1}\tilde{\rho}_u^A\sqrt{\rho_A}^{-1}\right) \quad \text{and} \quad B_v^{(\mu_2)} \triangleq \zeta_v^{(\mu_2)}\left(\sqrt{\rho_B}^{-1}\tilde{\rho}_v^B\sqrt{\rho_B}^{-1}\right), \tag{2.12}$$

where

$$\gamma_u^{(\mu_1)} \triangleq \frac{(1-\theta_1)}{(1+\epsilon_1)K_1}|\{l : U^{(\mu_1)}(l) = u\}| \quad \text{and} \quad \zeta_v^{(\mu_2)} \triangleq \frac{(1-\theta_2)}{(1+\epsilon_2)K_2}|\{k : V^{(\mu_2)}(k) = v\}|. \tag{2.13}$$

Let $\mathbb{1}_{\{\text{sP-1}\}}$ denote the indicator random variable corresponding to the event that $\{A_u^{(\mu_1)} : u \in \mathcal{U}\}$ forms a sub-POVM for all $\mu_1 \in [1, N_1]$. Similarly define $\mathbb{1}_{\{\text{sP-2}\}}$ with regard to $\{B_v^{(\mu_2)} : v \in \mathcal{V}\}$. If $\mathbb{1}_{\{\text{sP-1}\}} = 1$, then, for each $\mu_i \in [1, N_i]$ construct $M_i^{(\mu_i)}$, for $i = 1, 2$, as in the following:

$$M_1^{(\mu_1)} \triangleq \{A_u^{(\mu_1)} : u \in \mathcal{U}\}, \quad M_2^{(\mu_2)} \triangleq \{B_v^{(\mu_2)} : v \in \mathcal{V}\}.$$

These collections $M_1^{(\mu_1)}$ and $M_2^{(\mu_2)}$ are completed using the operators $A_{0_U}^{(\mu_1)} \triangleq I - \sum_{u\in\mathcal{U}} A_u^{(\mu_1)}$ and $B_{0_V}^{(\mu_2)} \triangleq I - \sum_{v\in\mathcal{V}} B_v^{(\mu_2)}$, and these operators are associated with symbols 0_U and 0_V. In the case of the complementary event, i.e., $\mathbb{1}_{\{\text{sP-i}\}} = 0$, we define $M_i^{(\mu_i)} \triangleq \{I\}$, for $i = 1, 2$, and denote the output as 0_U or 0_V, respectively. Hence by construction $M_1^{(\mu_1)}$ and $M_2^{(\mu_2)}$ are sub-POVMs for all $\mu_i \in [1, N_i]$, for $i = 1, 2$. For a fixed $\{\mathcal{C}^{(\mu_1,\mu_2)}\}_{\mu_1\in[1,N_1],\mu_2\in[1,N_2]}$, the probability distribution P

[3]The inverse used in $\sqrt{\rho}^{-1}$ refers to the generalized inverse as defined in (*Holevo*, 2012, Section 5.6).

induced on $(\mathcal{U}\cup\{0_U\})\times(\mathcal{V}\cup\{0_V\})$ has the following salient features.

$$P\{(u,v)\} = \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v},$$

if $(u,v)\in\mathcal{U}\times\mathcal{V}$, and

$$\begin{aligned}&P\big((\mathcal{U}\cup\{0_U\})\times(\mathcal{V}\cup\{0_V\})\backslash(\mathcal{U}\times\mathcal{V})\big)\\&\quad=\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\left(1-\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u,v}\gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v}\right)+\left(1-\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\right),\end{aligned}$$

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

$$\Omega_{u,v}\triangleq\operatorname{Tr}\left\{\sqrt{\rho_A\otimes\rho_B}^{-1}(\tilde{\rho}_u^A\otimes\tilde{\rho}_v^B)\sqrt{\rho_A\otimes\rho_B}^{-1}\rho_{AB}\right\}. \tag{2.14}$$

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 positive integers (T_1,T_2) and choose a (μ_1,μ_2) pair. For each symbol $u\in\mathcal{U}$ assign an index from $[1,T_1]$ randomly and uniformly, such that the assignments for different sequences are done independently. Perform a similar random and independent assignment for all $v\in\mathcal{V}$ with indices chosen from $[1,T_2]$. Repeat this assignment for every $\mu_1\in[1,N_1]$ and $\mu_2\in[1,N_2]$. For each $i\in[1,T_1]$ and $j\in[1,T_2]$, let $\mathcal{B}_1^{(\mu_1)}(i)$ and $\mathcal{B}_2^{(\mu_2)}(j)$ denote the i^{th} and the j^{th} bins, respectively. More precisely, $\mathcal{B}_1^{(\mu_1)}(i)$ is the set of all u symbols with assigned index equal to i, and similar is $\mathcal{B}_2^{(\mu_2)}(j)$. Define the following operators:

$$\Gamma_i^{A,(\mu_1)}\triangleq\sum_{u\in\mathcal{B}_1^{(\mu_1)}(i)}A_u^{(\mu_1)},\qquad\Gamma_j^{B,(\mu_2)}\triangleq\sum_{v\in\mathcal{B}_2^{(\mu_2)}(j)}B_v^{(\mu_2)},$$

for all $i\in[1,T_1]$ and $j\in[1,T_2]$. Using these operators, we form the following collection:

$$M_A^{(\mu_1)}\triangleq\{\Gamma_i^{A,(\mu_1)}\}_{i\in[1,T_1]},\quad M_B^{(\mu_2)}\triangleq\{\Gamma_j^{B,(\mu_2)}\}_{j\in[1,T_2]}. \tag{2.15}$$

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

relations

$$\sum_i \Gamma_i^{A,(\mu_1)} = \sum_{u\in\mathcal{U}} A_u^{(\mu_1)}, \quad \text{and} \quad \sum_j \Gamma_j^{B,(\mu_2)} = \sum_{v\in\mathcal{V}} B_v^{(\mu_2)}.$$

To make $M_A^{(\mu_1)}$ and $M_B^{(\mu_2)}$ 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[4]. Now, we intend to use the completions $[M_A^{(n,\mu_1)}]$ and $[M_B^{(n,\mu_2)}]$ as the POVMs for each encoder. In event that $\mathbb{1}_{\text{sP-i}} = 0$, for $i = 1, 2$, then the symbols 0_U and 0_V are mapped to 0. Also, note that the effect of the binning is in reducing the communication rates from $(\log(K_1 + 1), \log(K_2 + 1))$ to $(\log(T_1 + 1), \log(T_2 + 1))$.

Decoder mapping: Note that the operators $\{A_u^{(\mu_1)} \otimes B_v^{(\mu_2)}\}_{u\in\mathcal{U},v\in\mathcal{V}}$ are used to simulate $M_A \otimes M_B$. Binning can be viewed as partitioning of the set of classical outcomes into bins. Suppose an outcome (U, V) occurred in the measurement process. 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 symbols (U, V). More precisely, we define a mapping $F^{(\mu_1,\mu_2)}$, for (μ_1, μ_2), acting on the outputs of $[M_A^{(\mu_1)}] \otimes [M_B^{(\mu_2)}]$ as follows. On observing (μ_1, μ_2) and the classical indices $(i, j) \in [1, T_1] \times [1, T_2]$ communicated by the encoders, the decoder populates

$$D_{i,j}^{(\mu_1,\mu_2)} \triangleq \left\{(u, v) \in \mathcal{C}^{(\mu_1,\mu_2)} : (u, v) \in \mathcal{W} \text{ and } (u, v) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j)\right\},$$

where $\mathcal{W}$ is an arbitrary subset of $\mathcal{U} \times \mathcal{V}$. For every $\mu_l \in [1, N_l]$, for $l = 1, 2$, and $i \in [1, K_1]$ and $j \in [1, K_2]$, define the function $F^{(\mu_1,\mu_2)}(i, j) = (u, v)$ if (u, v) is the only element of $D_{i,j}^{(\mu_1,\mu_2)}$; otherwise $F^{(\mu_1,\mu_2)}(i, j) = (0_U, 0_V)$. Further, $F^{(\mu_1,\mu_2)}(i, j) = (0_U, 0_V)$ for $i = 0$ or $j = 0$. With this mapping, we form the following collection of operators, denoted by $\tilde{M}_{AB}$,

$$\tilde{\Lambda}_{u,v}^{AB} \triangleq \frac{\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}}{N_1 N_2} \sum_{\mu_1=1}^{N_1}\sum_{\mu_2=1}^{N_2} \sum_{(i,j):F^{(\mu_1,\mu_2)}(i,j)=(u,v)} \Gamma_i^{A,(\mu_1)} \otimes \Gamma_j^{B,(\mu_2)}$$
$$+ (1 - \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}})(I \otimes I)\mathbb{1}_{\{(u,v)=(0_U,0_V)\}},$$

$\forall (u, v) \in (\mathcal{U} \bigcup \{0_U\}) \times (\mathcal{V} \bigcup \{0_V\})$. Note that by construction $\tilde{M}_{AB}$ is a sub-POVM.

[4]Note that $\Gamma_0^{A,(\mu_1)} = I - \sum_i \Gamma_i^{A,(\mu_1)} = I - \sum_{u\in\mathcal{U}} A_u^{(\mu_1)}$ and $\Gamma_0^{B,(\mu_2)} = I - \sum_j \Gamma_j^{B,(\mu_2)} = I - \sum_{v\in\mathcal{V}} B_v^{(\mu_2)}$.

2.2.4.2 Analysis of POVM and Trace Distance

We show that $\tilde{M}_{AB}$ is a sub-POVM that is faithful to the sub-POVM $M_A \otimes M_B$, with respect to ρ_{AB}. More precisely, we provide a bound on

$$G_{\rho_{AB}} \triangleq \Xi(M_{AB}, \tilde{M}_{AB}). \tag{2.16}$$

Step 1: $M_1^{\mu_1)}$ and $M_2^{(\mu_2)}$ are sub-POVMs and individually approximating. As a first step, one can show that $M_1^{(\mu_1)}$ and $M_2^{(\mu_2)}$ individually approximate the corresponding POVMs in the expected sense. More precisely the following lemma holds.

Lemma II.10. *For the POVM ensemble described above, we have*

$$\mathbb{E}(\Xi_{\rho_A}(M_A, M_1)) \leq \alpha_A(\epsilon_1, K_1, N_1),$$
$$\mathbb{E}(\Xi_{\rho_B}(M_B, M_2)) \leq \alpha_B(\epsilon_2, K_2, N_2),$$

where $M_1 \triangleq \frac{1}{N_1}\sum_{\mu_1} M_1^{(\mu_1)}$, *and* $M_2 \triangleq \frac{1}{N_2}\sum_{\mu_2} M_2^{(\mu_2)}$.

Proof. Follows from the proof of Theorem II.9, as the assumptions of that theorem (which M_A and M_B have to satisfy) are met as a part of the current theorem statement (see (2.2a-2.2c)). □

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 (2.16). 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

$$S_1 \triangleq \left\| (\mathrm{id} \otimes [M_A] \otimes [M_B])(\Psi^{\rho}_{RAB}) - \frac{1}{N_1 N_2}\sum_{\mu_1,\mu_2} (\mathrm{id} \otimes [M_1^{(\mu_1)}] \otimes [M_2^{(\mu_2)}])(\Psi^{\rho}_{RAB}) \right\|_1,$$
$$S_2 \triangleq \left\| \frac{1}{N_1 N_2}\sum_{\mu_1,\mu_2} (\mathrm{id} \otimes [M_1^{(\mu_1)}] \otimes [M_2^{(\mu_2)}])(\Psi^{\rho}_{RAB}) - (\mathrm{id} \otimes [\tilde{M}_{AB}])(\Psi^{\rho}_{RAB}) \right\|_1, \tag{2.17}$$

where S_1 captures the effect of using approximating sub-POVMs M_1 and M_2 instead of the actual sub-POVMs M_A and M_B, while S_2 captures the error introduced by binning these approximating sub-POVMs. Before we proceed further, we provide the following lemma which will be useful in the rest of the paper.

Lemma II.11. *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 Γ^A acting on $\mathcal{H}_A$, we have*

$$\sum_{y\in\mathcal{Y}} \left\|\sqrt{\rho_{AB}}\left(\Gamma^A \otimes \Lambda_y^B\right)\sqrt{\rho_{AB}}\right\|_1 \leq \left\|\sqrt{\rho_A}\,\Gamma^A\sqrt{\rho_A}\right\|_1, \tag{2.18}$$

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

Proof. The proof is provided in Appendix A.2. □

Next, we provide a bound on S_1 using the following Mutual Covering Lemma.

Lemma II.12. *(Mutual Covering Lemma) Suppose a sub-POVM $\hat{M}_X$ is ϵ_X-faithful to M_X with respect to ρ_X, and a sub-POVM $\hat{M}_Y$ is ϵ_Y-faithful to M_Y with respect to ρ_Y, where $\rho_X = \mathrm{Tr}_Y\{\rho_{XY}\}$ and $\rho_Y = \mathrm{Tr}_X\{\rho_{XY}\}$. Then the sub-POVM $\hat{M}_X \otimes \hat{M}_Y$ is $(\epsilon_X+\epsilon_Y)$-faithful to the POVM $M_X \otimes M_Y$ with respect to ρ_{XY}.*

Proof. The proof is provided in the Appendix A.3. □

Using Lemma II.12 with $\rho_{XY} = \rho_{AB}$, $\hat{M}_X = \frac{1}{N_1}\sum_{\mu_1} M_1^{(\mu_1)}$, $\hat{M}_Y = \frac{1}{N_2}\sum_{\mu_2} M_2^{(\mu_2)}$, $M_X = [M_A]$ and $M_Y = [M_B]$, and Lemma II.10, we have $\mathbb{E}(S_1) \leq \alpha_A(\epsilon_1, N_1, K_1) + \alpha_B(\epsilon_2, N_2, K_2)$. For later convenience, we state the following lemma which will be used in analyzing the binning operation:

Lemma II.13. *We have*

$$\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left|\lambda_{u,v}^{AB} - \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\gamma_v^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v}\mathbb{1}_{\{sP_{-1}\}}\mathbb{1}_{\{sP_{-2}\}}\right| + \mathbb{1}_{\{sP_{-1}\}}\mathbb{1}_{\{sP_{-2}\}}\left(1 - \frac{1}{N_1N_2}\sum_{u,v}\sum_{\mu_1,\mu_2}\gamma_v^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v}\right) + \left(1 - \mathbb{1}_{\{sP_{-1}\}}\mathbb{1}_{\{sP_{-2}\}}\right) \leq S_1, \tag{2.19}$$

where $\Omega_{u,v}$ is defined as in (2.14).

Proof. The proof follows from Lemma 2 in *Wilde et al.* (2012). □

Step 3: Analyzing the effect of Binning. In this step, we provide an upper bound on S_2. For $(u,v) \in \mathcal{B}_1^{(\mu_1)}(i) \times \mathcal{B}_2^{(\mu_2)}(j)$, define $e^{(\mu_1,\mu_2)}(u,v) \triangleq F^{(\mu_1,\mu_2)}(i,j)$. For any $(u,v) \notin \mathcal{C}^{(\mu_!,\mu_2)}$ define $e^{(\mu_1,\mu_2)}(u,v) = (0_U, 0_V)$. Note that $e^{(\mu_1,\mu_2)}$ captures the overall effect of the binning followed by

the decoding function $F^{(\mu_1,\mu_2)}$. For all $u \in \mathcal{U}$ and $v \in \mathcal{V}$, let $\Phi_{u,v} \triangleq |u,v\rangle\langle u,v|$. With this notation, we simplify S_2 using the following proposition.

Proposition II.14. *S_2 can be simplified as*

$$S_2 = S_3 + \frac{2}{N_1}\sum_{\mu_1} \mathrm{Tr}\left(\left(I - \sum_{u\in\mathcal{U}} A_u^{(\mu_1)}\right)\rho_A\right) + \frac{2}{N_2}\sum_{\mu_2} \mathrm{Tr}\left(\left(I - \sum_{v\in\mathcal{V}} B_v^{(\mu_2)}\right)\rho_B\right) + 2\left(2 - \mathbb{1}_{\{sP\text{-}1\}} - \mathbb{1}_{\{sP\text{-}2\}}\right),$$

where

$$S_3 \triangleq \mathbb{1}_{\{sP\text{-}1\}}\mathbb{1}_{\{sP\text{-}2\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left\|\Phi_{u,v} - \Phi_{e^{(\mu_1,\mu_2)}(u,v)}\right\|_1 \gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v}.$$

Proof. The proof is provided in Appendix A.6. □

In the next proposition we provide a bound on the expectation of S_2.

Proposition II.15 (Mutual Packing)**.** *We have*

$$\mathbb{E}[S_2] \leq \alpha_P(\epsilon_1,\epsilon_2,K_1,K_2,N_1,N_2,T_1,T_2,\mathcal{W}).$$

Proof. The proof is provided in Appendix A.7. □

Combining the results from the mutual covering and mutual packing lemmas we obtain

$$G \leq \alpha_A + \alpha_B + \alpha_P.$$

This completes the proof of the theorem.

2.2.5 Proof of Second Inner Bound (Theorem II.6)

We develop a proof as a corollary to Theorem II.4. Assume that the operators of the original POVM M_{AB} are decomposed as

$$\Lambda_z^{AB} \triangleq \sum_{u,v} \mathbb{1}_{\{g(u,v)=z\}}\Lambda_u^A \otimes \Lambda_v^B, \ \forall z \in \mathcal{Z}, \tag{2.20}$$

for some POVMs M_A and M_B with operators denoted by $\{\Lambda_u^A\}_{u\in\mathcal{U}}$ and $\{\Lambda_v^B\}_{v\in\mathcal{V}}$, respectively, and for some function $g:\mathcal{U}\times\mathcal{V}\to\mathcal{Z}$ where $\mathcal{U},\mathcal{V}$ and $\mathcal{Z}$ are three finite sets. In what follows, we show the existence of an (n,T_1,T_2,N_1,N_2) distributed protocol with the associated sub-POVM $\tilde{M}_{AB}^{(n)}$ that is ϵ-faithful to M_{AB} with respect to $\rho_{AB}^{\otimes n}$ (according to Definition I.1), where $\epsilon>0$ can be made arbitrarily small for all sufficiently large n. More precisely, we plan to show that

$$\sum_{z^n}\|\sqrt{\rho_{AB}^{\otimes n}}\left(\Lambda_{z^n}^{AB}-\tilde{\Lambda}_{z^n}^{AB}\right)\sqrt{\rho_{AB}^{\otimes n}}\|_1\leq\epsilon. \tag{2.21}$$

Next we claim that it is sufficient to show that there exists a distributed protocol for the finite set $\mathcal{U}\times\mathcal{V}$ and the Hilbert space $\mathcal{H}_A\otimes\mathcal{H}_B$, with parameters (n,T_1,T_2,N_1,N_2) such that the associated sub-POVM $\tilde{M}_{AB}^{(n)}=\{\tilde{\Lambda}_{u^n,v^n}^{AB}\}_{u^n\in\mathcal{V}^n,v^n\in\mathcal{V}^n}$ satisfies $\Xi_{\rho_{AB}^{\otimes n}}(M_A^{\otimes n}\otimes M_B^{\otimes n},\tilde{M}_{AB}^{(n)})\leq\epsilon$. This is because one can always apply the function $g(\cdot,\cdot)$ componentwise on (u^n,v^n) to yield a sub-POVM with operators

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

that satisfies the constraint (2.21) as

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

Fix three free parameters $\delta>0$, $\epsilon_1>0$, and $\epsilon_2>0$. We make the following identification with regard to Theorem II.4. (a) Let $\rho_{AB}\leftrightarrow\rho_{AB}^{\otimes n}$, $M_A\leftrightarrow M_A^{\otimes n}$, and $M_B\leftrightarrow M_B^{\otimes n}$, which implies that $\lambda_u^A\leftrightarrow\lambda_{u^n}^A$, $\lambda_v^B\leftrightarrow\lambda_{v^n}^B$ and $\lambda_{uv}^{AB}\leftrightarrow\lambda_{u^nv^n}^{AB}$. (b) Let $\mathcal{U}\leftrightarrow\mathcal{T}_\delta^{(n)}(U)$, $\mathcal{V}\leftrightarrow\mathcal{T}_\delta^{(n)}(V)$, and $\mathcal{W}\leftrightarrow\mathcal{T}_\delta^{(n)}(U,V)$, where $\mathcal{T}_\delta^{(n)}(U)$, $\mathcal{T}_\delta^{(n)}(V)$ and $\mathcal{T}_\delta^{(n)}(U,V)$ are the δ-typical sets defined for $\{\lambda_u^A\}$, $\{\lambda_v^B\}$ and $\{\lambda_{uv}^{AB}\}$, respectively. (c) Furthermore, let $\Pi_{\rho_A}\leftrightarrow\Pi_{\rho_A,\delta}$ $\Pi_{\rho_B}\leftrightarrow\Pi_{\rho_B,\delta}$, $\Pi_u^A\leftrightarrow\Pi_{u^n,\delta}^A$, and $\Pi_v^B\leftrightarrow\Pi_{v^n,\delta}^B$, where $\Pi_{\rho_A,\delta}$ and $\Pi_{\rho_B,\delta}$ denote the δ-typical projectors (as in (*Wilde*, 2013a, Def. 15.1.3)) for marginal density operators ρ_A and ρ_B, respectively[5]. Also, for any $u^n\in\mathcal{T}_\delta^{(n)}(U)$ and $v^n\in\mathcal{T}_\delta^{(n)}(V)$, let $\Pi_{u^n,\delta}^A$

[5]Note that $\Pi_{\rho_A,\delta}$ and $\Pi_{\rho_B,\delta}$ also depend on n, however, for ease of notation, we do not make this explicit.

and $\Pi^B_{v^n,\delta}$ denote the strong conditional typical projectors (as in (*Wilde*, 2013a, Def. 15.2.4)) for the canonical ensembles $\{\lambda^A_u, \hat{\rho}^A_u\}$ and $\{\lambda^B_v, \hat{\rho}^B_v\}$, respectively.

With the above identification, and using the property of typical sets and typical projectors, we now find the values of the variables $D_1, D_2, d_1, d_2, F_1, F_2, f_1$ and f_2 that satisfy the hypotheses of Theorem II.4. Firstly, using the properties of strong typical and conditional typical projectors (*Wilde*, 2013a, Properties 15.2.4 and 15.2.7) we have the first four inequalities (hypotheses (2.2a)) satisfied for all $\epsilon_1, \epsilon_2 \in (0,1)$, and for all sufficiently large n. Next, using (*Wilde*, 2013a, Property 15.1.2), there exist functions $\delta_1(\delta), \delta_2(\delta) \searrow 0$ as $\delta \searrow 0$, such that for all sufficiently large n, the first two inequalities of hypotheses (2.2b) are satisfied for $D_1 \triangleq 2^{n(S(RB)_{\sigma_1}+\delta_1(\delta))}$ and $D_2 \triangleq 2^{n(S(RA)_{\sigma_2}+\delta_2(\delta))}$. Further, using (*Wilde*, 2013a, Property 15.2.3), there exist functions $\delta_3(\delta), \delta_4(\delta) \searrow 0$ as $\delta \searrow 0$, such that for all sufficiently large n, the next two inequalities of hypotheses (2.2b) are satisfied for $d_1 \triangleq 2^{n(S(RB|U)_{\sigma_1}-\delta_3(\delta))}$, $d_2 \triangleq 2^{n(S(RA|V)_{\sigma_2}-\delta_4(\delta))}$. The next four inequalities of hypotheses (2.2c) follow from the definition of projectors $\Pi_{\rho_A,\delta}$, $\Pi_{\rho_B,\delta}$, $\Pi^A_{u^n,\delta}$ and $\Pi^B_{v^n,\delta}$. And finally, the four inequalities of hypotheses (2.2d) are satisfied by using (*Wilde*, 2013a, Property 15.1.3) and by defining $F_1 \triangleq 2^{n(S(RB)_{\sigma_1}-\delta_1(\delta))}$, $F_2 \triangleq 2^{n(S(RA)_{\sigma_2}-\delta_2(\delta))}$, and $f_1 \triangleq D_1$ and $f_2 \triangleq D_2$.

This implies the existence of a distributed protocol with parameters (n, T_1, T_2, N_1, N_2) with $\Xi_{\rho_{AB}}(M_{AB}, \tilde{M}_{AB}) \leq \alpha_A + \alpha_B + \alpha_P$. We now evaluate the upper bound. For this we let $T_i = 2^{nR_i}$, $N_i = 2^{nC_i}$, and $K_i = 2^{n\tilde{R}_i}$, for some non-negative real numbers R_i, C_i, and $\tilde{R}_i$ for $i = 1, 2$. Moreover, we assume that $S(U)_{\sigma_3} \geq \tilde{R}_1$ and $S(V)_{\sigma_3} \geq \tilde{R}_2$. If not, then faithful simulation can be achieved in a trivial way.

Using the property of strongly typical sets, note that for all sufficiently large n we have $|\mathcal{U}| \leq 2^{n(S(U)_{\sigma_3}+\delta_5(\delta))}$, $|\mathcal{V}| \leq 2^{n(S(V)_{\sigma_3}+\delta_5(\delta))}$, $\lambda^A_{\mathrm{m}} \leq 2^{-n(S(U)_{\sigma_3}-\delta_5(\delta))}$ $\lambda^B_{\mathrm{m}} \leq 2^{-n(S(V)_{\sigma_3}-\delta_5(\delta))}$. Furthermore, we have the bounds: $|\mathcal{W}| \leq 2^{n(S(U,V)_{\sigma_3}+\delta_5(\delta))}$, $W_A \leq 2^{n(S(U|V)_{\sigma_3}+\delta_5(\delta))}$, and $W_B \leq 2^{n(S(V|U)_{\sigma_3}+\delta_5(\delta))}$, where $\delta_5(\delta) \searrow 0$ as $\delta \searrow 0$. For all sufficiently large n we have $\theta_i \leq \epsilon_i$ for $i = 1, 2$. Hence for $i = 1, 2$, the term $(2\epsilon_i/(1+\epsilon_i)) + f(\epsilon_i, \theta_i)$ can be made arbitrarily small by a suitable choice of ϵ_i and n.

Next we see that

$$\frac{1}{\sqrt{N_1K_1}}\sum_{u\in\mathcal{U}}\sqrt{\lambda_u^A} \le 2^{\left[-\frac{n}{2}(\tilde{R}_1+C_1-S(U)_{\sigma_3}-3\delta_5)\right]}, \text{ and}$$
$$\frac{1}{\sqrt{N_2K_2}}\sum_{v\in\mathcal{V}}\sqrt{\lambda_v^B} \le 2^{\left[-\frac{n}{2}(\tilde{R}_2+C_2-S(V)_{\sigma_3}-3\delta_5)\right]},$$

and hence can be made arbitrarily small for all sufficiently large n if

$$\tilde{R}_1 + C_1 > S(U)_{\sigma_3} + 3\delta_5 \quad \text{and} \quad \tilde{R}_2 + C_2 > S(V)_{\sigma_3} + 3\delta_5.$$

Moving on, we have

$$D_1N_1\exp\left[-\frac{K_1\epsilon_1^3 d_1 D_1^{-1}}{4\ln 2}\right] \le 2^{n(S(RB)_{\sigma_1}+C_1+\delta_1)}\exp\left[-\frac{2^{n(\tilde{R}_1-I(RB;U)_{\sigma_1}-\delta_1-\delta_3)}\epsilon_1^3}{4\ln 2}\right],$$
$$D_2N_2\exp\left[-\frac{K_2\epsilon_2^3 d_2 D_2^{-1}}{4\ln 2}\right] \le 2^{n(S(RA)_{\sigma_2}+C_2+\delta_2)}\exp\left[-\frac{2^{n(\tilde{R}_2-I(RA;V)_{\sigma_2}-\delta_2-\delta_4)}\epsilon_2^3}{4\ln 2}\right],$$

which can be made arbitrarily small for all sufficiently large n if

$$\tilde{R}_1 > I(U;RB)_{\sigma_1} + \delta_1 + \delta_3, \text{ and } \tilde{R}_2 > I(V;RA)_{\sigma_2} + \delta_2 + \delta_4.$$

Next we have

$$\lambda^{AB}(\mathcal{W}^c) = \sum_{(u^n,v^n)\notin T_\delta(U,V)} \lambda_{u^n,v^n}^{AB},$$

which can be made arbitrarily small for sufficiently large n. Finally, we have

$$\begin{aligned}
&\left[\frac{\lambda_{\mathrm{m}}^A\lambda_{\mathrm{m}}^B|\mathcal{W}|K_1K_2}{(1-\theta_1)(1-\theta_2)T_1T_2} + \frac{K_1W_A\lambda_{\mathrm{m}}^A}{(1-\theta_1)T_1}\left(1+\frac{\lambda_{\mathrm{m}}^BK_2}{(1-\theta_2)}\right) + \frac{K_2W_B\lambda_{\mathrm{m}}^B}{(1-\theta_2)T_2}\left(1+\frac{\lambda_{\mathrm{m}}^AK_1}{(1-\theta_1)}\right)\right]\frac{f_1f_2}{F_1F_2} \\
&\le \frac{2^{2n(\delta_1+\delta_2)}}{(1-\theta_1)(1-\theta_2)}\Big[2^{\left[n(\tilde{R}_1+\tilde{R}_2-R_1-R_2-I(U:V)_{\sigma_3}+3\delta_5)\right]} \\
&\qquad + 2^{\left[n(\tilde{R}_1-R_1-I(U:V)_{\sigma_3}+2\delta_5)\right]} + 2^{\left[n(\tilde{R}_1+\tilde{R}_2-R_1-I(U:V)_{\sigma_3}-S(V)_{\sigma_3}+3\delta_5)\right]} \\
&\qquad + 2^{\left[n(\tilde{R}_1+\tilde{R}_2-R_2-I(U:V)_{\sigma_3}-S(U)_{\sigma_3}+3\delta_5)\right]} + 2^{\left[n(\tilde{R}_2-R_2-I(U:V)_{\sigma_3}+2\delta_5)\right]}\Big],
\end{aligned}$$

which again can be made arbitrarily small for all sufficiently large n if

$$\tilde{R}_1 + \tilde{R}_2 - R_1 - R_2 < I(U;V)_{\sigma_3} - 3\delta_5 - 2(\delta_1 + \delta_2).$$

To sum-up, we have showed that the trace distance inequality in (2.21) holds for all sufficiently small δ, ϵ_1, and ϵ_2, and all sufficiently large n, if the following bounds hold:

$$\begin{aligned}
\tilde{R}_1 > I(U;RB)_{\sigma_1}, \quad \tilde{R}_2 > I(V;RA)_{\sigma_2}, \\
C_1 + \tilde{R}_1 > S(U)_{\sigma_3}, \quad C_2 + \tilde{R}_2 > S(V)_{\sigma_3}, \\
(\tilde{R}_1 - R_1) + (\tilde{R}_2 - R_2) < I(U;V)_{\sigma_3}, \\
\tilde{R}_1 \geq R_1 \geq 0, \quad \tilde{R}_2 \geq R_2 \geq 0, \\
C_1 \geq 0, \quad C_2 \geq 0.
\end{aligned} \tag{2.22}$$

Therefore, there exists a distributed protocol with parameters $(n, 2^{nR_1}, 2^{nR_2}, 2^{nC_1}, 2^{nC_2})$ such that its overall POVM $\tilde{M}_{AB}^{(n)}$ is ϵ-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 II.16. *Let $\mathcal{R}_1$ denote the closure of the set of all (R_1, R_2, C_1, C_2) 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 (2.22). Let, $\mathcal{R}_2$ denote the set of all quadruple (R_1, R_2, C_1, C_2) that satisfies the inequalities in (2.6) given in the statement of the theorem. Then, $\mathcal{R}_1 = \mathcal{R}_2$.*

Proof. The proof follows by Fourier-Motzkin elimination *Ziegler* (2012). □

2.3 Q-C Distributed Rate Distortion Theory

As an application of faithful simulation of distributed measurements (Theorem II.6), we consider the distributed extension of QC rate distortion coding *Datta et al.* (2013b). 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 $\rho_A = \mathrm{Tr}_B\{\rho_{AB}\}$ and $\rho_B = \mathrm{Tr}_A\{\rho_{AB}\}$. They both perform a measurement on n copies of

their sub-systems and send the classical bits to Charlie. Upon receiving the classical bits sent by Alice and Bob, a reconstruction state is produced by Charlie. The objective of Charlie is to produce a reconstruction of the source ρ_{AB} within a targeted distortion threshold which is measured by a given distortion observable.

2.3.1 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 II.17 (q-c source coding setup)**.** A QC 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 purification of ρ_{AB}, $\mathcal{H}_{\hat{X}}$ 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 Charlie by the following definition.

Definition II.18 (q-c protocol)**.** An (n, Θ_1, Θ_2) QC protocol for a given input and reconstruction Hilbert spaces $(\mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{H}_{\hat{X}})$ is defined by POVMs $M_A^{(n)}$ and $M_B^{(n)}$ acting on $\mathcal{H}_A^{\otimes n}$ and $\mathcal{H}_B^{\otimes n}$ with Θ_1 and Θ_2 number of outcomes, respectively, and a set of reconstruction states $S_{i,j} \in \mathcal{D}(\mathcal{H}_{\hat{X}}^{\otimes n})$ for all $i \in [1, \Theta_1], j \in [1, \Theta_2]$.

The overall action of Alice, Bob and Charlie, as a QC protocol, on a quantum source ρ_{AB} is given by the following operation

$$\mathcal{N}_{A^n B^n \mapsto \hat{X}^n} : \rho_{AB}^{\otimes n} \mapsto \sum_{i,j} \mathrm{Tr}\{(\Lambda_i^A \otimes \Lambda_j^B)\rho_{AB}^{\otimes n}\}\ S_{i,j}, \tag{2.23}$$

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 II.17, the distortion of a $(n = 1, \Theta_1, \Theta_2)$ QC protocol is measured as

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

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

$$\Delta^{(n)} = \frac{1}{n}\sum_{i=1}^{n} \Delta_{R_i\hat{X}_i} \otimes I_{R\hat{X}}^{\otimes[n]\setminus i}, \tag{2.24}$$

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

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

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

The authors in *Datta et al.* (2013b) studied the point-to-point setup of the above formulation wherein Bob is absent. 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 (please see (*Datta et al.*, 2013b, Sec. 4) for more details). In this paper, we allow Δ to be any non-negative and bounded operator acting on the appropriate Hilbert spaces. Moreover, we allow for the use of any c-q reconstruction mapping as the action of Charlie.

Definition II.19 (Achievability)**.** For a QC 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, Θ_1, Θ_2) QC protocol satisfying

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

where $\mathcal{N}_{A^nB^n\mapsto\hat{X}^n}$ is defined as in (2.23). 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.

2.3.2 An 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 2.2. The proof of the theorem in provided in Section 2.3.3.

Theorem II.20. *For a given QC 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*

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

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* $\sigma_1^{RUBQ} \triangleq \sum_{q\in\mathcal{Q}} P_Q(q)(\mathrm{id}_R \otimes M_A^q \otimes \mathrm{id}_B)(\Psi_{RAB}^{\rho_{AB}}) \otimes |q\rangle\langle q|$*,* $\sigma_2^{RAVQ} \triangleq \sum_{q\in\mathcal{Q}} P_Q(q)(\mathrm{id}_R \otimes \mathrm{id}_A \otimes M_B^q)(\Psi_{RAB}^{\rho_{AB}}) \otimes |q\rangle\langle q|$*, and* $\sigma_3^{RUVQ} \triangleq \sum_{q\in\mathcal{Q}} P_Q(q)(\mathrm{id}_R \otimes M_A^q \otimes M_B^q)(\Psi_{RAB}^{\rho_{AB}}) \otimes |q\rangle\langle q|$*, where* (U, V) *represents the output of* M_{AB}*, and* $\mathcal{N}_{AB\mapsto\hat{X}} : \rho_{AB} \mapsto \sum_{u,v,q} P_Q(q) \operatorname{Tr}\{(\Lambda_u^{A,q} \otimes \Lambda_v^{B,q})\rho_{AB}\}\ S_{u,v,q}$*.*

Remark II.21. Note that for the auxiliary state σ_1, we have

$$\begin{aligned} \sigma_1^{RQ} = \operatorname{Tr}_{UB}\{\sigma_1^{RUBQ}\} &= \sum_q P_Q(q) \operatorname{Tr}_{UAB}\left\{\sum_{u\in\mathcal{U}}\{(I_{RB}\otimes\Lambda_u^q)(\Psi_{RAB}^{\rho_{AB}})\}\otimes|u\rangle\langle u|\right\}\otimes|q\rangle\langle q| \\ &= \sum_q P_Q(q) \operatorname{Tr}_{AB}\left\{\left\{(I_{RB}\otimes\sum_{u\in\mathcal{U}}\Lambda_u^q)(\Psi_{RAB}^{\rho_{AB}})\right\}\right\}\otimes|q\rangle\langle q| \\ &= \rho_R \otimes \sum_q P_Q(q)\,|q\rangle\langle q|\,, \end{aligned}$$

which gives $I(R;Q)_{\sigma_1} = 0$. Similar statements hold for the states σ_2 and σ_3.

One can observe that the rate region in Theorem II.20 matches in form with the classical Berger-Tung region when ρ_{AB} is a mixed state of a collection of orthogonal pure states. Note that the rate region is an inner bound for the set of all achievable rates. The single-letter characterization of the set of achievable rates is still an open problem even in the classical setting. Some progress has been made recently on this problem which provides an improvement over Berger-Tung rate region *Shirani and Pradhan* (2014).

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 II.6. Fix POVMs (M_A, M_B) and reconstruction states $S_{u,v}$ as in the statement of the theorem. Let $\mathcal{N}_{AB\mapsto\hat{X}}$ be the mapping corresponding to these POVMs and the reconstruction states. Then, $d(\rho_{AB}, \mathcal{N}_{AB\mapsto\hat{X}}) \leq D$. According to Theorem II.6, for any $\epsilon > 0$, there exists an $(n, 2^{nR_1}, 2^{nR_2}, N_1, N_2)$ distributed protocol for ϵ-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 (2.6). Let $\tilde{M}_A^{(\mu_1)}, \tilde{M}_B^{(\mu_2)}, \mu_i \in [1, N_i]$, for $i = 1, 2$, and $f^{(\mu_1,\mu_2)}$ 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 QC protocol for distributed quantum source coding.

For each $\mu_i \in [1, N_i]$, for $i = 1, 2$, consider the QC protocol with parameters $\Theta_i = 2^{nR_i}, i = 1, 2$, and POVMs $\tilde{M}_A^{(\mu_1)}, \tilde{M}_B^{(\mu_2)}$. Moreover, we use n-length reconstruction states

$$S_{i,j} \triangleq \sum_{u^n, v^n} \mathbb{1}\{f^{(\mu_1,\mu_2)}(i,j) = (u^n, v^n)\} S_{u^n,v^n},$$

where $S_{u^n,v^n} = \otimes_i S_{u_i,v_i}$. Further, let the corresponding mappings be denoted as $\tilde{\mathcal{N}}^{(\mu_1,\mu_2)}_{A^nB^n\mapsto\hat{X}^n}$. With this notation, for the average of these random protocols, the following bounds hold:

$$\begin{aligned}
&\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2} d(\rho_{AB}^{\otimes n}, \tilde{\mathcal{N}}^{(\mu_1,\mu_2)}_{A^nB^n\mapsto\hat{X}^n}) \\
&= \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2} \mathrm{Tr}\left\{\Delta^{(n)}(\mathrm{id}\otimes\tilde{\mathcal{N}}^{(\mu_1,\mu_2)}_{A^nB^n\mapsto\hat{X}^n})\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\} \\
&= \mathrm{Tr}\left\{\Delta^{(n)}(\mathrm{id}\otimes\mathcal{N}^{\otimes n}_{AB\mapsto\hat{X}})\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\} + \mathrm{Tr}\left\{\Delta^{(n)}(\mathrm{id}\otimes(\tilde{\mathcal{N}}_{A^nB^n\mapsto\hat{X}^n} - \mathcal{N}^{\otimes n}_{AB\mapsto\hat{X}}))\Psi^{\rho_{AB}}_{R^nA^nB^n}\right\}
\end{aligned}$$

$$
\begin{aligned}
&\overset{(a)}{\leq} \operatorname{Tr}\left\{\Delta\left((\mathrm{id}_R \otimes \mathcal{N}_{AB\mapsto\hat{X}})(\Psi_{RAB}^{\rho_{AB}})\right)\right\} + \|\Delta^{(n)}(\mathrm{id} \otimes (\mathcal{N}_{AB\mapsto\hat{X}}^{\otimes n} - \tilde{\mathcal{N}}_{A^nB^n\mapsto\hat{X}^n}))\Psi_{R^nA^nB^n}^{\rho_{AB}}\|_1 \\
&\overset{(b)}{\leq} D + \|\Delta^{(n)}\|_\infty \|(\mathrm{id} \otimes (\mathcal{N}_{AB\mapsto\hat{X}}^{\otimes n} - \tilde{\mathcal{N}}_{A^nB^n\mapsto\hat{X}^n}))\Psi_{R^nA^nB^n}^{\rho_{AB}}\|_1 \\
&\overset{(c)}{\leq} D + \|\Delta^{(n)}\|_\infty \|(\mathrm{id} \otimes (M_A^{\otimes n} \otimes M_B^{\otimes n} - \tilde{M}_{AB}))\Psi_{R^nA^nB^n}^{\rho_{AB}}\|_1 \\
&\overset{(d)}{\leq} D + \epsilon\|\Delta\|_\infty,
\end{aligned}
$$

where $\tilde{\mathcal{N}}_{AB\mapsto\hat{X}}$ is the average of $\tilde{\mathcal{N}}_{AB\mapsto\hat{X}}^{(\mu_1,\mu_2)}$, and $\tilde{M}_{AB}$ is the overall POVM of the underlying distributed protocol as given in (2.1). The inequality (a) holds by the fact that $|\operatorname{Tr}\{A\}| \leq \|A\|_1$. (b) follows from the fact that for any two operators A and B acting on a Hilbert space $\mathcal{H}$ the following inequalities hold.

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

(see in (*Wilde*, 2013a, Exercise 12.2.1) for a proof). (c) is due to the monotonicity of the trace-distance *Wilde* (2013a) with respect to the quantum channel given by $\mathrm{id} \otimes \mathcal{L}_{UV\mapsto\hat{X}}^{\otimes n}$, where

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

And (d) follows by Theorem II.6, and the fact that $\|\Delta^{(n)}\|_\infty \leq \|\Delta\|_\infty$. Hence using the collection of codebooks $\{\mathcal{C}^{(\mu_1,\mu_2)}\}_{\mu_1\in[1,N_1],\mu_2\in[1,N_2]}$, constructed in Theorem II.6, and averaged over the common randomness, the distortion constraint $\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2} d(\rho_{AB}^{\otimes n}, \tilde{\mathcal{N}}_{A^nB^n\mapsto\hat{X}^n}^{(\mu_1,\mu_2)}) \leq D + \epsilon\|\Delta\|_\infty$ is met. Hence there must exist a realization of the common randomness (μ_1, μ_2), and the corresponding codebook $\mathcal{C}^{(\mu_1,\mu_2)}$ that achieves this distortion. This completes the proof of the theorem, since Δ is a bounded operator. □

2.3.3 An Outer Bound

In this section, we provide an outer bound for the achievable rate-distortion region. The proof of this theorem is provided in Section 2.3.3.

Theorem II.22. *Given a QC 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

$$R_1 \geq I(W_1; R|W_2, Q)_\sigma, \tag{2.25a}$$

$$R_2 \geq I(W_2; R|W_1, Q)_\sigma, \tag{2.25b}$$

$$R_1 + R_2 \geq I(W_1, W_2; R|Q)_\sigma, \tag{2.25c}$$

$$D \geq \mathrm{Tr}\left\{\Delta \sigma^{R\hat{X}}\right\}, \tag{2.25d}$$

for some state $\sigma^{W_1W_2RQ\hat{X}}$ *which can be written as*

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

where W_1, W_2 *and* Q *represent auxiliary quantum states, and* $\mathcal{N}_{AB\mapsto W_1W_2Q\hat{X}}$ *is a quantum test channel with* $I(R;Q)_\sigma = 0$.

Remark II.23. One may question the computability of the outer bound provided in Theorem II.22. 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 *Berger* (1977) and a recent substantial improvement was made by authors in *Wagner and Anantharam* (2008). 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.

Proof. Suppose the triplet (R_1, R_2, D) is achievable. Then, from Definition II.19, for all $\epsilon > 0$, there exists an (n, Θ_1, Θ_2) q-c protocol satisfying the inequalities in the definition. Let $M_A \triangleq \{\Lambda_{l_1}^A\}_{l_1\in[1,\Theta_1]}$, $M_B \triangleq \{\Lambda_{l_2}^B\}_{l_2\in[1,\Theta_2]}$, and $S_{l_1,l_2} \in \mathcal{D}(\mathcal{H}_{\hat{X}}^{\otimes n})$ be the corresponding POVMs and reconstruction states. Let L_1, L_2 denote the outcomes of the measurements. Then, for Alice's rate, we obtain

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

where the state τ is defined as $\tau^{L_1L_2R^n\hat{X}^n} \triangleq$

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

and the inequalities follow from L_1 and L_2 being classical. 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 definition of direct-sum of Hilbert spaces *Conway* (1985). With this definition, consider the following single-letterization:

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

where the state σ is defined as:

$$\sigma^{L_1L_2RQ\hat{X}} \triangleq \sum_{l_1,l_2}\frac{|l_1,l_2\rangle\langle l_1,l_2|}{n}\otimes\Big(\sum_{j=1}^{n}\Big(\mathrm{Tr}_{R^n_{j+1}A^nB^n}\Big\{(\mathrm{id}\otimes\Lambda^A_{l_1}\otimes\Lambda^B_{l_2})\Psi^{\rho_{AB}}_{R^nA^nB^n}\Big\}\otimes|j\rangle\langle j|\otimes\mathrm{Tr}_{\hat{X}^n\sim j}\{S_{l_1,l_2}\}\Big)\Big), \tag{2.26}$$

and $\mathrm{Tr}_{\hat{X}^n\sim j}$ denotes tracing over $(\hat{X}^{\otimes j-1}\otimes\hat{X}^{\otimes n}_{j+1})$, and $Q \triangleq (R^{J-1},J)$, and J is an averaging random variable which is uniformly distributed over $[1,n]$. We have attached a quantum register for this classical random variable yielding the state σ. The equality (a) follows from the following lemma.

Lemma II.24. *Consider the classical-quantum state*

$$\sigma_{JABC} \triangleq \sum_{j=1}^{n} P_J(j)\,|j\rangle\,\langle j|\otimes\rho^j_{ABC},$$

where $\{|j\rangle\}_{j\in[1,n]}$ *is an orthonormal set in some Hilbert space* $\mathcal{H}_J$, $\rho^j_{ABC}\in\mathcal{D}(\mathcal{H}^j_A\otimes\mathcal{H}^j_B\otimes\mathcal{H}^j_C)$, *where* $\{\mathcal{H}^j_A\otimes\mathcal{H}^j_B\otimes\mathcal{H}^j_C\}_{j\in[1,n]}$ *is a collection of finite-dimensional Hilbert spaces. Note that* $\sigma_{ABC} = \mathrm{Tr}_J(\sigma_{XABC})$ *is a state on* $\bigoplus_{j=1}^n(\mathcal{H}^j_A\otimes\mathcal{H}^j_B\otimes\mathcal{H}^j_C)$. *Then* $I(A;B|C,J)_\sigma = \sum_{j=1}^n P_J(j)I(A;B|C)_{\rho^j}$.

Proof. The proof is provided in Appendix A.4. □

We elaborate on the Hilbert space associated with Q as follows. Suppose $\{|\phi_i\rangle\}_{i\in\mathcal{I}}$ is an orthonormal basis for $\mathcal{H}_R$. Then, a basis for $\mathcal{H}_R^{\otimes k}$ is given by

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

for all $\mathbf{i}^k \in \mathcal{I}^k$. Consider the direct-sum of the Hilbert spaces $\bigoplus_{k=1}^n \mathcal{H}_R^{\otimes k}$ and the Hilbert space $\mathcal{H}_J \otimes \mathcal{H}_R^{\otimes k}$. 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 *Meyer* (2006).

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

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

where the inequalities follow from L_1 and L_2 being classical. In addition, the distortion of this q-c protocol satisfies $d(\rho_{AB}^{\otimes n}, \mathcal{N}_{A^nB^n\mapsto\hat{X}^n}) \leq D+\epsilon$, where $\mathcal{N}_{A^nB^n\mapsto\hat{X}^n}$ is the quantum channel associated with the protocol. Therefore, as the distortion observable is symbol-wise additive, we obtain

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

$$\stackrel{(a)}{=} \mathrm{Tr}\{(\Delta \otimes I_Q)\sigma^{RQ\hat{X}}\},$$

where (a) holds because of the following argument. From (2.26), one can show by partially tracing over (L_1, L_2), that

$$\begin{aligned}\sigma^{RQ\hat{X}} &= \mathrm{Tr}_{L_1,L_2}\{\sigma^{L_1L_2RQ\hat{X}}\} \\ &= \sum_{j=1}^{n} \frac{1}{n} |j\rangle\langle j| \otimes \mathrm{Tr}_{R^n_{j+1}\hat{X}^n\sim j}\{(\mathrm{id} \otimes \mathcal{N}_{A^nB^n\mapsto\hat{X}^n})(\Psi^{\rho_{AB}}_{R^nA^nB^n})\},\end{aligned}$$

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

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

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

$$\begin{aligned}\sigma^{L_1L_2RQ\hat{X}} = \Big(\sum_{j=1}^{n} \frac{1}{n} \sum_{l_1,l_2} |l_1,l_2\rangle\langle l_1,l_2| \otimes \big(\mathrm{Tr}_{R^n_{j+1}A^nB^n}\Big\{(\mathrm{id} \\ \otimes \Lambda^A_{l_1} \otimes \Lambda^B_{l_2})\Psi^{\rho_{AB}}_{R^nA^nB^n}\Big\} \otimes |j\rangle\langle j| \otimes \mathrm{Tr}_{\hat{X}^n\sim j}\{S_{l_1,l_2}\}\big)\Big).\end{aligned}$$

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

$$\begin{aligned}&\mathcal{N}^{(j)}_{AB\mapsto L_1L_2R^{(j-1)}\hat{X}}(\omega_{AB}) \\ &\triangleq \sum_{l_1,l_2} |l_1,l_2\rangle\langle l_1,l_2| \otimes \big(\mathrm{Tr}_{R^n_{j+1}A^nB^n}\Big\{(\mathrm{id}_{R^n\sim j} \otimes \Lambda^A_{l_1} \otimes \Lambda^B_{l_2})(\omega_{AB} \otimes E_j)\Big\} \otimes \mathrm{Tr}_{\hat{X}^n\sim j}\{S_{l_1,l_2}\}\big),\end{aligned}$$

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

these definitions, let

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

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

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

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

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

where the last equality is due to the distributive property of tensor product over direct sum operation. Hence, σ^{RQ} is in a tensor product of the form $\sigma^R\otimes\sigma^Q$, and therefore, $I(R;Q)_\sigma = 0$. The proof completes by identifying W_1 and W_2 with L_1 and L_2, respectively.

□

2.4 Simulation of P2P POVMs with Stochastic Processing

Before we detail into the distributed setup, in this section, we discuss an extension of the Winter's point-to-point measurement compression scheme *Winter* (2004), incorporating additional stochastic processing at the receiver. This problem was first discussed in *Wilde et al.* (2012), where a theorem characterizing the achievable rate region was provided. The results were also rederived

in (*Anshu et al.*, 2019, Corollary 4) and *Berta et al.* (2014) for the same problem. Since this problem provides us with some of the tools required for the proof of the final result of this chapter (Theorem II.33), developed in sequel (Section 2.5.2), we rederive the achievability of its rate region using the approximating POVMs developed in *Winter* (2004). This will serve as a building block toward proving the main result. In this problem, 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 *Wilde et al.* (2012) 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.

2.4.1 Problem Formulation

Definition II.25 (Protocol)**.** For a given finite set $\mathcal{X}$, and a Hilbert space $\mathcal{H}_A$, a measurement simulation protocol with stochastic processing with parameters (n, Θ, 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 $[1, \Theta]$, and
2) a collection of Bob's classical stochastic maps $P^{(\mu)}(x^n|l)$ for all $l \in [1, \Theta]$, $x^n \in \mathcal{X}^n$ and $\mu \in [1, N]$.
The overall sub-POVM of this protocol, given by $\tilde{M}$, is characterized by the following operators:

$$\tilde{\Lambda}_{x^n} \triangleq \frac{1}{N} \sum_{\mu,l} P^{(\mu)}(x^n|l)\ \Lambda_l^{(\mu)}, \quad \forall x^n \in \mathcal{X}^n, \tag{2.27}$$

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

In the above definition, Θ characterizes the amount of classical bits communicated from Alice to Bob, and the amount of common randomness is determined by N, with μ 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 II.26 (Achievability)**.** 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, Θ, N) such that its overall sub-POVM $\tilde{M}$ is ϵ-faithful to $M^{\otimes n}$ with respect to $\rho^{\otimes n}$ (see Definition I.1), 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.

2.4.2 Achievable Rate Region

The following theorem characterizes the achievable rate region.

Theorem II.27. *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 $M_A \triangleq \{\Lambda_w^A\}_{w\in\mathcal{W}}$, with $\mathcal{W}$ being a finite set, and a stochastic map $P_{X|W} : \mathcal{W} \to \mathcal{X}$ such that*

$$R \geq I(R;W)_\sigma \quad \textit{and} \quad R + C \geq I(RX;W)_\sigma,$$
$$\Lambda_x \triangleq \sum_{w\in\mathcal{W}} P_{X|W}(x|w)\Lambda_w^A, \quad \forall x \in \mathcal{X}.$$

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

Remark II.28. 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

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

for $\left\{\lambda_w^A, \hat{\rho}_w^A\right\}$ being the canonical ensemble associated with the POVM M and the state ρ as defined in (2.37). With this ensemble, we have

$$\begin{aligned} I(R;W)_\sigma &= \chi\left(\left\{\lambda_w^A, \hat{\rho}_w^A\right\}\right) \quad \text{and} \\ I(RX;W)_\sigma &= I(X;W)_\sigma + I(R;XW)_\sigma - I(R;X)_\sigma \\ &= I(X;W)_\sigma + \chi\left(\left\{\lambda_w^A, \hat{\rho}_w^A\right\}\right) - \chi\left(\{\lambda_x, \hat{\rho}_x\}\right), \end{aligned}$$

where we have used the Markov Chain $R - W - X$ which is evident from the structure of σ^{RWX}.

Remark II.29. As was pointed out in Section 1.5.1, a proof of achievability and converse for Theorem II.27 was provided by Wilde et al. in (*Wilde et al.*, 2012, Section III). With regards to the proof of achievability, the authors assume (*Wilde et al.*, 2012, Eqns. 53 and 54) to be true, but do not provide a proof for it. Due to the presence of the cut-off operator, which is constructed for the ensemble and not for the individual operators, these equations may not always be true. Since the proof hinges on these two equations and we do not see a straightforward way to prove the two assumptions made (also confirmed in *Wilde* (Aug 2019)), we provide an alternate proof for achievability below. For the proof of converse, we refer the readers to (*Wilde et al.*, 2012, Section III.3).

Proof. Suppose there exist a POVM M_A and a stochastic map $P_{X|W} : \mathcal{W} \to \mathcal{X}$, such that M can be decomposed as

$$\Lambda_x \triangleq \sum_w P_{X|W}(x|w)\Lambda_w^A, \ \forall x \in \mathcal{X}. \tag{2.28}$$

We begin by defining a canonical ensemble corresponding to M_A as $\{\lambda_w^A, \hat{\rho}_w^A\}_{w\in\mathcal{W}}$. Similarly. for each $w^n \in \mathcal{W}^n$, we also define

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

where $\hat{\rho}_{w^n}^A \triangleq \bigotimes_i \hat{\rho}_{w_i}^A$, Π_ρ denotes the δ-typical projector (as in (*Wilde*, 2013a, Def. 15.1.3)) corresponding to the density operator ρ, Π_{w^n} denotes the strong conditional typical projector (as in (*Wilde*, 2013a, Def. 15.2.4)) corresponding to the canonical ensemble $\{\lambda_w^A, \hat{\rho}_w^A\}_{w\in\mathcal{W}}$, and $\hat{\Pi}$ denotes the projector onto the subspace spanned by the eigenstates of

$$\sum_{w^n\in\mathcal{T}_\delta^{(n)}(W)} \frac{\lambda_{w^n}^A}{(1-\varepsilon)}\Pi_\rho\Pi_{w^n}\hat{\rho}_{w^n}^A\Pi_{w^n}\Pi_\rho$$

corresponding to eigenvalues larger than $\varepsilon 2^{-n(S(\rho)+\delta_1)}$, where $\delta_1(\delta)$ is such that $\mathrm{Tr}(\Pi_\rho) \leq 2^{n(S(\rho)+\delta_1)}$, and $\varepsilon \triangleq \sum_{w^n\notin\mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n}^A$[6], and $\delta_1 \searrow 0$ as $\delta \searrow 0$.

Using the above definitions, we now construct the approximating POVM.

[6]Note that Π_ρ, Π_{w^n} and $\hat{\Pi}$ depend on n, and δ however, for ease of notation, we do not make this explicit.

2.4.2.1 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 $W^{n,(\mu)}(l)$ from the set $\mathcal{W}^n$, according to the pruned distributions, i.e.,

$$\mathbb{P}\left(W^{n,(\mu)}(l) = w^n\right) \triangleq \begin{cases} \dfrac{\lambda^A_{w^n}}{(1-\varepsilon)} & \text{for} \quad w^n \in \mathcal{T}^{(n)}_\delta(W) \\ 0 & \text{otherwise} \end{cases}. \tag{2.29}$$

Let the collection of operators $\tilde{M}^{(n,\mu)}_A$ be defined as $\{A^{(\mu)}_{w^n} : w^n \in \mathcal{T}^{(n)}_\delta(W)\}$ for each $\mu \in [1, 2^{nC}]$, where $A^{(\mu)}_{w^n}$ is defined as

$$A^{(\mu)}_{w^n} \triangleq \gamma^{(\mu)}_{w^n}\left(\sqrt{\rho}^{-1}\tilde{\rho}^A_{w^n}\sqrt{\rho}^{-1}\right) \quad \text{and} \quad \gamma^{(\mu)}_{w^n} \triangleq \frac{1}{2^{nR}}\sum_{l=1}^{2^{nR}} \frac{(1-\varepsilon)}{(1+\eta)}\mathbb{1}_{\{W^{n,(\mu)}(l)=w^n\}}, \tag{2.30}$$

with $\eta \in (0,1)$ determining the probability that $\tilde{M}^{(n,\mu)}_A$ does not form a sub-POVM, for all $\mu \in [1, 2^{nC}]$. Since the construction is very similar to the one used in Section 2.2.4 and 2.2.5, we make a claim similar to the one in Lemma II.10 (also see Proposition A.1). 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 with high probability. Let $\mathbb{1}_{\{\text{sP}\}}$ denote the indicator random variable corresponding to this event. The claim is as follows. For any $\epsilon \in (0,1)$, $\eta \in (0,1)$, any $\delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[\mathbb{1}_{\{\text{sP}\}}] \geq (1-\epsilon)$ if $R > I(R;W)_\sigma$, where the definition of σ_{RWX} follows from the statement of theorem. From this, let $[\tilde{M}^{(n,\mu)}_A]$ denote the completion of the corresponding sub-POVM $\tilde{M}^{(n,\mu)}_A$ for $\mu \in [1, 2^{nC}]$. Let the operators completing these POVMs, given by $I - \sum_{w^n \in \mathcal{T}^{(n)}_\delta(W)} A^{(\mu)}_{w^n}$, be denoted by $A^{(\mu)}_{w^n_0}$ for some $w^n_0 \notin \mathcal{T}^{(n)}_\delta(W)$, for all $\mu \in [1, 2^{nC}]$, and $A^{(\mu)}_{w^n} = 0$ for $w^n \notin \mathcal{T}^{(n)}_\delta(W) \bigcup \{w^n_0\}$. We use the trivial POVM $\{I\}$ in the case of the complementary event that the operators do not form sub-POVMs for all μ, and associate it with the sequence $\{w^n_0\}$. The POVM is given by $\{\mathbb{1}_{\{\text{sP}\}}A^{(\mu)}_{w^n} + (1 - \mathbb{1}_{\{\text{sP}\}})\mathbb{1}_{\{w^n = w^n_0\}}I\}_{w^n \in \mathcal{W}^n}$. Using this construction, we define the intermediate

approximating 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

$$\tilde{\Lambda}_{w^n}^A \triangleq \left(\frac{1}{2^{nC}} \sum_\mu A_{w^n}^{(\mu)} \right) \mathbb{1}_{\{\text{sP}\}} + (1 - \mathbb{1}_{\{\text{sP}\}}) \mathbb{1}_{\{w^n = w_0^n\}} I.$$

Now, we define Bob's stochastic map as $P_{X|W}^n$, yielding the operators of the final approximating POVM as

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

2.4.2.2 Trace Distance

Fix an arbitrary $\epsilon \in (0,1)$. Now, we compare the action of this approximating POVM on the input state $\rho^{\otimes n}$ with that of the given POVM $M^{\otimes n}$, using the characterization provided in Definition I.1. Specifically, we show using the expressions for canonical ensemble that, under certain conditions on (R, C), for all sufficiently large n we have $\mathbb{E}[G] \leq \epsilon$, where

$$G \triangleq \sum_{x^n \in \mathcal{X}^n} \left\| \sum_{w^n \in \mathcal{W}^n} P_{X|W}^n(x^n|w^n) \sqrt{\rho^{\otimes n}} (\Lambda_{w^n}^A - \tilde{\Lambda}_{w^n}^A) \sqrt{\rho^{\otimes n}} \right\|_1. \tag{2.31}$$

As a first step, we split and bound G as $G \leq S_1 + S_2 + 2(1 - \mathbb{1}_{\{\text{sP}\}})$, where

$$\begin{aligned} S_1 &\triangleq \sum_{x^n} \left\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) - \frac{1}{2^{nC}} \sum_{w^n \neq w_0^n} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \tilde{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) \right\|_1, \\ S_2 &\triangleq \sum_{x^n} \left\| P_{X|W}^n(x^n|w_0^n) \frac{1}{2^{nC}} \times \sum_{\mu=1}^{2^{nC}} \left[\sqrt{\rho^{\otimes n}} (I - \sum_{w^n \neq w_0^n} A_{w^n}^{(\mu)}) \sqrt{\rho^{\otimes n}} \right] \right\|_1. \end{aligned} \tag{2.32}$$

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

$$\begin{aligned} S_{11} &\triangleq \left\| \sum_{x^n} \left[\sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) \otimes |x^n\rangle\langle x^n| - \frac{1}{2^{nC}} \sum_{w^n \neq w_0^n} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) \otimes |x^n\rangle\langle x^n| \right] \right\|_1, \\ S_{12} &\triangleq \left\| \sum_{x^n} \sum_{w^n \neq w_0^n} \left[\frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \hat{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) - \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \gamma_{w^n}^{(\mu)} \tilde{\rho}_{w^n}^A P_{X|W}^n(x^n|w^n) \right] \otimes |x^n\rangle\langle x^n| \right\|_1. \end{aligned}$$

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 $\mathbb{E}[S_{11}]$ is small. To simplify the notation, we define $\sigma_{w^n} = \sum_{x^n} P^n_{X|W}(x^n|w^n)\,|x^n\rangle\langle x^n|$ which gives S_{11} as

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

We develop the following lemma to bound this term.

Lemma II.30. *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 (2.29) and $\mathcal{T}_{w^n}$ is any tensor product state of the form $\mathcal{T}_{w^n} = \bigotimes_{i=1}^n \mathcal{T}_{w_i}$. Then, for any $\epsilon_2 \in (0,1)$, and for all $\eta, \delta \in (0,1)$ sufficiently small, and n sufficiently large, we have*

$$\mathbb{E}\left\| \sum_{w^n} \lambda^A_{w^n} \mathcal{T}_{w^n} - \frac{1}{2^{n(R+C)}} \frac{(1-\varepsilon)}{(1+\eta)} \sum_{l,\mu} \mathcal{T}_{W^{n,(\mu)}(l)} \right\|_1 \leq \epsilon_2, \tag{2.33}$$

if $R+C > S(\sum_w \lambda^A_w \mathcal{T}_w) - \sum_w \lambda^A_w S(\mathcal{T}_w) = \chi\left(\{\lambda^A_w, \mathcal{T}_w\}\right)$, where $\{W^{n,(\mu)}(l) : l \in [1, 2^{nR}], \mu \in [1, 2^{nC}]\}$ are independent random vectors generated from $\mathcal{W}^n$ according to the pruned distribution given in (2.29).

Proof. The proof of the lemma is provided in Appendix A.5 □

Therefore, using the lemma above with $\mathcal{T}_{w^n} \triangleq \hat{\rho}^A_{w^n} \otimes \sigma_{w^n}$, for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[S_{11}] \leq \epsilon$ if $R + C > S(\sum_w \lambda^A_w \hat{\rho}^A_w \otimes \sigma_w) - \sum_w \lambda^A_w S(\hat{\rho}^A_w \otimes \sigma_w) = \chi\left(\{\lambda^A_w\}, \{\hat{\rho}^A_w \otimes \sigma_w\}\right) = I(RX;W)_\sigma$, where σ is as defined in the statement of the theorem. Secondly, we bound S_{12} by applying expectation with respect to the codebook generation, and using Gentle Measurement Lemma *Wilde* (2013a) as follows,

$$\mathbb{E}\left[S_{12}\right] \overset{(a)}{\leq} \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{x^n} \sum_{w^n \neq w_0^n} P^n_{X|W}(x^n|w^n) \mathbb{E}\left[\gamma^{(\mu)}_{w^n} \left\|(\hat{\rho}^A_{w^n} - \tilde{\rho}^A_{w^n})\right\|_1\right]$$

$$\overset{(b)}{=} \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \frac{\lambda_{w^n}^A}{(1+\eta)} \left\| \hat{\rho}_{w^n}^A - \tilde{\rho}_{w^n}^A \right\|_1$$

$$= \frac{1}{(1+\eta)} \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n}^A \left\| \hat{\rho}_{w^n}^A - \hat{\Pi}\Pi_\rho \Pi_{w^n} \hat{\rho}_{w^n}^A \Pi_{w^n} \Pi_\rho \hat{\Pi} \right\|_1$$

$$\overset{(c)}{\leq} \frac{(1-\varepsilon)}{(1+\eta)} (2\sqrt{\varepsilon'} + 2\sqrt{\varepsilon''}) \triangleq \varepsilon_3, \tag{2.34}$$

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 $\mathbb{E}[\gamma_{w^n}^{(\mu)}] = \frac{\lambda_{w^n}^A}{(1+\eta)}$, and finally (c) uses repeated application of the average gentle measurement lemma, by setting $\varepsilon_3 = \frac{(1-\varepsilon)}{(1+\eta)}(2\sqrt{\varepsilon'} + 2\sqrt{\varepsilon''})$ with $\varepsilon_3 \searrow 0$ as $n \to \infty$ for all sufficiently small $\delta > 0$, and, $\varepsilon' \triangleq \varepsilon_p + 2\sqrt{\varepsilon_p}$ and $\varepsilon'' \triangleq 2\varepsilon_p + 2\sqrt{\varepsilon_p}$ for $\varepsilon_p \triangleq 1 - \min\left\{ \mathrm{Tr}\left\{ \Pi_\rho \hat{\rho}_{w^n}^A \right\}, \mathrm{Tr}\left\{ \Pi_{w^n} \hat{\rho}_{w^n}^A \right\}, 1-\varepsilon \right\}$ (see (35) in *Wilde et al.* (2012) 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

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

where

$$\begin{aligned} S_{21} &\triangleq \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \left\| \sum_{w^n} \lambda_{w^n}^A \hat{\rho}_{w^n}^A - \frac{(1-\varepsilon)}{(1+\eta)} \frac{1}{2^{nR}} \sum_{l=1}^{2^{nR}} \hat{\rho}_{W^{n,(\mu)}(l)}^A \right\|_1, \\ S_{22} &\triangleq \frac{1}{2^{nC}} \sum_{\mu=1}^{2^{nC}} \sum_{w^n \neq w_0^n} \gamma_{w^n}^{(\mu)} \left\| \hat{\rho}_{w^n}^A - \tilde{\rho}_{w^n}^A \right\|_1. \end{aligned} \tag{2.35}$$

Now, for the first term in (2.35) we use Lemma II.30 and claim that for any $\epsilon \in (0,1)$, any

$\eta, \delta \in (0,1)$, sufficiently small, any n sufficiently large, we have $\mathbb{E}[S_{21}] \leq \epsilon$, if

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

where σ is as defined in the statement of the theorem. Note that the requirement we obtain on R was already imposed when claiming the collection of operators $A_{w^n}^{(\mu)}$ forms a sub-POVM. As for the second term in (2.35) we again use the gentle measurement Lemma and bound its expected value as

$$\mathbb{E}\left[\frac{1}{2^{nC}}\sum_{\mu=1}^{2^{nC}}\sum_{w^n\neq w_0^n}\gamma_{w^n}^{(\mu)}\left\|\hat{\rho}_{w^n}^A - \tilde{\rho}_{w^n}^A\right\|_1\right] = \sum_{w^n\in\mathcal{T}_\delta^{(n)}(W)}\frac{\lambda_{w^n}}{(1+\eta)}\left\|\hat{\rho}_{w^n}^A - \tilde{\rho}_{w^n}^A\right\|_1 \leq \varepsilon_3,$$

where ε_3 is defined in (2.34).

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 I.1. As a first step in establishing this bound, we showed that $G \leq S_1+S_2+2(1-\mathbb{1}_{\{\mathrm{sP}\}})$. Firstly, we have shown that $\mathbb{E}[\mathbb{1}_{\{\mathrm{sP}\}}] \geq (1-\epsilon)$ if $R > I(R;W)_\sigma$. Then considering S_1, we used the triangle inequality and divided it into two terms: S_{11} and S_{12}. Then, using Lemma II.30, we showed that for any given $\epsilon \in (0,1)$, $\mathbb{E}[S_{11}]$ can be made smaller than ϵ, if $R + C > I(RX;W)_\sigma$. As for S_{12}, we showed that it goes to zero in the expected sense using (2.34). 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 (2.35). We showed that they can be made arbitrarily small in the expected sense if $R > I(R;W)_\sigma$ for all sufficiently large n.

Hence for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large we have $\mathbb{E}[G] \leq 6\epsilon$ if

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

Therefore, using random coding arguments, there exists at least one collection of sub-POVMs with the above construction satisfying the statement of Theorem II.27.

□

2.5 Simulation of Distributed POVMs with Stochastic Processing

This brings us to the final result of this chapter. We begin by considering the simulation of distributed POVMs with stochastic processing. Consider a bipartite composite quantum system (A, B) represented by a Hilbert Space $\mathcal{H}_A \otimes \mathcal{H}_B$. Let ρ_{AB} be a density operator on $\mathcal{H}_A \otimes \mathcal{H}_B$. Consider a joint measurement M_{AB} on the system. Imagine that three parties, named Alice, Bob and Charlie, are trying to collectively simulate the joint measurement, using two measurements, one applied on each sub-system. The resources available to these parties are: some amount of classical common randomness pairwise shared among them, and classical communication links of specified rates between Alice and Charlie, and Bob and Charlie. Alice and Bob perform measurements $\tilde{M}_A^{(n)} \triangleq \{\Lambda_{l_1}^A\}$ and $\tilde{M}_B^{(n)} \triangleq \{\Lambda_{l_2}^B\}$ on n copies of sub-systems A and B, respectively. The measurements are performed in a distributed fashion with no communication taking place between Alice and Bob. Based on their respective measurements and the common randomness, Alice and Bob send some classical bits to Charlie. Upon receiving these classical bits, Charlie applies a stochastic processing operation on them, given by $P(\cdot|l_1, l_2)$, 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 $M_{AB}^{\otimes n}$. Further, the operators of the given measurement M_{AB} admit a decomposition of the form given in Definition II.1. We formally define the problem as follows.

2.5.1 Problem Formulation

The problem is defined in the following.

Definition II.31 (Distributed Protocol)**.** 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 $(n, \Theta_1, \Theta_2, N_1, N_2)$ is characterized by

1) a collection of Alice's sub-POVMs $\tilde{M}_A^{(\mu_1)}, \mu_1 \in [1, N_1]$ each acting on $\mathcal{H}_A^{\otimes n}$ and with outcomes in $[1, \Theta_1]$.

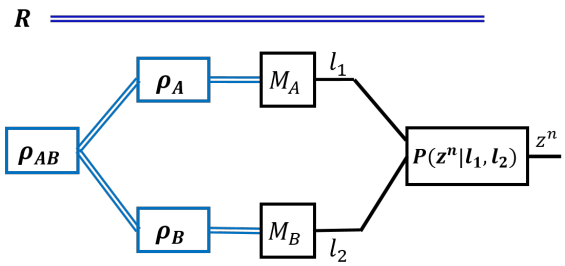

Figure 2.3: The diagram depicting the distributed POVM simulation problem with stochastic processing. In this setting, Charlie additionally has access to unlimited private randomness.

2) a collection of Bob's sub-POVMs $\tilde{M}_B^{(\mu_2)}, \mu_2 \in [1, N_2]$ each acting on $\mathcal{H}_B^{\otimes n}$ and with outcomes in $[1, \Theta_2]$.

3) a collection of Charlie's classical stochastic maps $P^{(\mu_1,\mu_2)}(z^n|l_1, l_2)$ for all $l_1 \in [1, \Theta_1], l_2 \in [1, \Theta_2], z^n \in \mathcal{Z}^n$, $\mu_1 \in [1, N_1]$, and $\mu_2 \in [1, N_2]$.

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

$$\tilde{\Lambda}_{z^n} \triangleq \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{l_1 \in [1,\Theta_1], l_2 \in [1,\Theta_2]} P^{(\mu_1,\mu_2)}(z^n|l_1, l_2) \Lambda_{l_1}^{A,(\mu_1)} \otimes \Lambda_{l_2}^{B,(\mu_2)}, \quad \forall z^n \in \mathcal{Z}^n,$$

where $\Lambda_{l_1}^{A,(\mu_1)}$ and $\Lambda_{l_2}^{B,(\mu_2)}$ are the operators corresponding to the sub-POVMs $\tilde{M}_A^{(\mu_1)}$ and $\tilde{M}_B^{(\mu_2)}$, respectively.

In the above definition, (Θ_1, Θ_2) determines the amount of classical bits communicated from Alice and Bob to Charlie, respectively. N_1 and N_2 denote the amount of pairwise common randomness. The classical stochastic maps $P^{(\mu_1,\mu_2)}(z^n|l_1, l_2)$ represent the action of Charlie on the received classical bits.

Definition II.32 (Achievability)**.** 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 quadruple (R_1, R_2, C_1, C_2) 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 $(n, \Theta_1, \Theta_2, N_1, N_2)$ such that its overall sub-POVM $\tilde{M}_{AB}$ is ϵ-faithful to $M_{AB}^{\otimes n}$ with respect to $\rho_{AB}^{\otimes n}$

(see Definition I.1), and

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

The set of all achievable quadruples (R_1,R_2,C_1,C_2) is called the achievable rate region.

2.5.2 An Inner Bound

The following theorem provides an inner bound to the achievable rate region. The proof of the theorem is provided below, while some of the tools required for the proof are borrowed from Section 2.4.

Theorem II.33. *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 II.1), a quadruple (R_1,R_2,C_1,C_2) is achievable if the following inequalities are satisfied:*

$$\begin{aligned}
R_1 &\geq I(U;RB)_{\sigma_1}-I(U;V)_{\sigma_3},\\
R_2 &\geq I(V;RA)_{\sigma_2}-I(U;V)_{\sigma_3},\\
R_1+R_2 &\geq I(U;RB)_{\sigma_1}+I(V;RA)_{\sigma_2}-I(U;V)_{\sigma_3},\\
R_1+C_1 &\geq I(U;RZ)_{\sigma_3}-I(U;V)_{\sigma_3},\\
R_2+C_2 &\geq I(V;RZ)_{\sigma_3}-I(U;V)_{\sigma_3},\\
R_1+R_2+C_1 &\geq I(U;RZ)_{\sigma_3}+I(V;RA)_{\sigma_2}-I(U;V)_{\sigma_3},\\
R_1+R_2+C_2 &\geq I(V;RZ)_{\sigma_3}+I(U;RB)_{\sigma_1}-I(U;V)_{\sigma_3},\\
R_1+R_2+C_1+C_2 &\geq I(UV;RZ)_{\sigma_3},
\end{aligned} \tag{2.36}$$

for some decomposition with POVMs $M_A=\{\Lambda_u^A\}_{u\in\mathcal{U}}$ and $M_B=\{\Lambda_v^B\}_{v\in\mathcal{V}}$ and a stochastic map $P_{Z|U,V}:\mathcal{U}\times\mathcal{V}\to\mathcal{Z}$, where the above information quantities are computed for the auxiliary states $\sigma_1^{RUB}\triangleq(\mathrm{id}_R\otimes M_A\otimes\mathrm{id}_B)(\Psi_{RAB}^{\rho_{AB}})$, $\sigma_2^{RAV}\triangleq(\mathrm{id}_R\otimes\mathrm{id}_A\otimes M_B)(\Psi_{RAB}^{\rho_{AB}})$, and $\sigma_3^{RUVZ}\triangleq\sum_{u,v,z}[\sqrt{\rho_{AB}}\left(\Lambda_u^A\otimes\Lambda_v^B\right)\sqrt{\rho_{AB}}]^R\otimes P_{Z|U,V}(z|u,v)\,|u\rangle\langle u|\otimes|v\rangle\langle v|\otimes|z\rangle\langle z|$, and $\Psi_{RAB}^{\rho_{AB}}$ is a purification[7] of ρ_{AB}.

[7]The information theoretic quantities remain independent of the purification used in their definitions.

Remark II.34. 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

$$\lambda_u^A \triangleq \mathrm{Tr}\{\Lambda_u^A \rho_A\}, \quad \lambda_v^B \triangleq \mathrm{Tr}\{\Lambda_v^B \rho_B\},$$
$$\lambda_{uv}^{AB} \triangleq \mathrm{Tr}\{(\Lambda_u^A \otimes \Lambda_v^B)\rho_{AB}\}, \quad \text{and}$$

$$\hat{\rho}_u^A \triangleq \frac{1}{\lambda_u^A}\sqrt{\rho_A}\Lambda_u^A\sqrt{\rho_A}, \quad \hat{\rho}_v^B \triangleq \frac{1}{\lambda_v^B}\sqrt{\rho_B}\Lambda_v^B\sqrt{\rho_B},$$
$$\hat{\rho}_{uv}^{AB} \triangleq \frac{1}{\lambda_{uv}^{AB}}\sqrt{\rho_{AB}}(\Lambda_u^A \otimes \Lambda_v^B)\sqrt{\rho_{AB}}. \tag{2.37}$$

Note that the post-measurement states corresponding to the outcomes u and v are given by $(\hat{\rho}_u^A)^T, (\hat{\rho}_v^B)^T$ and $(\hat{\rho}_{uv}^{AB})^T$, where transposes are defined with respect to the eigenbasis of the corresponding density operators. This entails that the states $\hat{\rho}_u^A, \hat{\rho}_v^B$ and $\hat{\rho}_{uv}^{AB}$ defined above have the same spectrum as the states induced on the purifying reference R after the measurement. However, these canonical states are not on the same "operational level" as the latter. Further, we define the following ensemble $\{\lambda_z, \hat{\rho}_z\}$ as

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

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

2.5.3 Proof of the Inner Bound

2.5.3.1 Construction of An Ensemble of POVMs

Suppose there exist POVMs $M_A \triangleq \{\Lambda_u^A\}_{u\in\mathcal{U}}$ and $M_B \triangleq \{\Lambda_v^B\}_{v\in\mathcal{V}}$ and a stochastic map $P_{Z|UV} : \mathcal{U}\times\mathcal{V}\to\mathcal{Z}$, such that M_{AB} can be decomposed as

$$\Lambda_z^{AB} = \sum_{u,v} P_{Z|UV}(z|u,v)\Lambda_u^A \otimes \Lambda_v^B,\ \forall z\in\mathcal{Z}. \tag{2.38}$$

Note that the proof technique here is very different to the one used in Section 2.2.5 for proving Theorem II.6. Recall that in Theorem II.6 we initiated the proof by constructing a protocol to faithfully simulate $M_A^{\otimes n}\otimes M_B^{\otimes n}$. However, here we are not interested in faithfully simulating $M_A^{\otimes n}\otimes M_B^{\otimes n}$. Instead, by carefully exploiting the private randomness Charlie possesses, manifested in terms of the stochastic processing applied by him 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 (2.6e) of Theorem II.6. This requires a considerably different methodology. More specifically, Lemma I.2 was employed in Theorem II.6, 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 does not allow a straightforward segmentation into two point-to-point problems. The problem becomes analytically tractable using an asymmetric partitioning.

2.5.3.2 Random Coding

We start by generating the canonical ensembles corresponding to M_A and M_B, as given in (2.37). With this notation, corresponding to each of the probability distributions, we can associate a δ-typical set. Let us denote $\mathcal{T}_\delta^{(n)}(U)$, $\mathcal{T}_\delta^{(n)}(V)$ and $\mathcal{T}_\delta^{(n)}(UV)$ as the δ-typical sets defined for $\{\lambda_u^A\}$, $\{\lambda_v^B\}$ and $\{\lambda_{uv}^{AB}\}$, respectively. Let Π_{ρ_A} and Π_{ρ_B} denote the δ-typical projectors (as in (*Wilde*, 2013a, Def. 15.1.3)) for marginal density operators ρ_A and ρ_B, respectively. Also, for any

$u^n \in \mathcal{U}^n$ and $v^n \in \mathcal{V}^n$, let $\Pi^A_{u^n}$ and $\Pi^B_{v^n}$ denote the strong conditional typical projectors (as in (*Wilde*, 2013a, Def. 15.2.4)) for the canonical ensembles $\{\lambda^A_u, \hat{\rho}^A_u\}$ and $\{\lambda^B_v, \hat{\rho}^B_v\}$, respectively. For each $u^n \in \mathcal{U}^n$ and $v^n \in \mathcal{V}^n$ define

$$\tilde{\rho}^{A'}_{u^n} \triangleq \Pi_{\rho_A} \Pi^A_{u^n} \hat{\rho}^A_{u^n} \Pi^A_{u^n} \Pi_{\rho_A}, \quad \tilde{\rho}^{B'}_{v^n} \triangleq \Pi_{\rho_B} \Pi^B_{v^n} \hat{\rho}^B_{v^n} \Pi^B_{v^n} \Pi_{\rho_B}, \tag{2.39}$$

where $\hat{\rho}^A_{u^n} \triangleq \bigotimes_i \hat{\rho}^A_{u_i}$ and $\hat{\rho}^B_{v^n} \triangleq \bigotimes_i \hat{\rho}^B_{v_i}$ [8].

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

$$\sigma^{A'} \triangleq \sum_{u^n \in \mathcal{T}^{(n)}_\delta(U)} \frac{\lambda^A_{u^n}}{(1-\varepsilon)} \tilde{\rho}^A_{u^n}, \quad \sigma^{B'} \triangleq \sum_{v^n \in \mathcal{T}^{(n)}_\delta(V)} \frac{\lambda^B_{v^n}}{(1-\varepsilon')} \tilde{\rho}^B_{v^n}, \tag{2.40}$$

where $\varepsilon = \sum_{u^n \in \mathcal{T}^{(n)}_\delta(U)} \lambda^A_{u^n}$ and $\varepsilon' = \sum_{v^n \in \mathcal{T}^{(n)}_\delta(V)} \lambda^B_{v^n}$. Note that $\sigma^{A'}$ and $\sigma^{B'}$ defined above are expectations with respect to the pruned distribution *Wilde* (2013a). Let $\hat{\Pi}^A$ and $\hat{\Pi}^B$ be the projectors onto the subspaces spanned by the eigenstates of $\sigma^{A'}$ and $\sigma^{B'}$ corresponding to eigenvalues that are larger than $\varepsilon 2^{-n(S(\rho_A)+\delta_1)}$ and $\varepsilon' 2^{-n(S(\rho_B)+\delta_1)}$, where $\delta_1 > 0$ is such that $\mathrm{Tr}(\Pi_{\rho_A}) \leq 2^{n(S(\rho_A)+\delta_1)}$, and $\mathrm{Tr}(\Pi_{\rho_B}) \leq 2^{n(S(\rho_B)+\delta_1)}$, and $\delta_1 \searrow 0$ as $\delta \searrow 0$. Lastly, define

$$\Lambda^A_{u^n} \triangleq \hat{\Pi}^A \tilde{\rho}^{A'}_{u^n} \hat{\Pi}^A, \quad \text{and} \quad \Lambda^B_{v^n} \triangleq \hat{\Pi}^B \tilde{\rho}^{B'}_{v^n} \hat{\Pi}^B. \tag{2.41}$$

In what follows, we construct two random POVMs one for each encoder. Fix a positive integer N and positive real numbers $\tilde{R}_1$ and $\tilde{R}_2$ satisfying $\tilde{R}_1 < S(U)_{\sigma_3}$ and $\tilde{R}_2 < S(V)_{\sigma_3}$, where σ_3 is defined as

$$\sigma_3^{RUV} \triangleq (\mathrm{id}_R \otimes M_A \otimes M_B)(\Psi^{\rho_{AB}}_{RAB}),$$

with $\Psi^{\rho_{AB}}_{RAB}$ being any purification of ρ_{AB}. 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. Let $\tilde{\mu}_1 \in [1, \tilde{N}_1]$ and $\tilde{\mu}_2 \in [1, \tilde{N}_2]$ denote additional pairwise shared randomness used for random coding purposes. This randomness is only

[8] Note that $\tilde{\rho}^A_{u^n}$ and $\tilde{\rho}^B_{v^n}$ are not tensor products operators.

used to show the existence of a desired distributed protocol (as defined in Definition II.31), and is used only for bounding purposes. We denote $\bar{\mu}_i \triangleq (\mu_i, \tilde{\mu}_i)$, and $\bar{N}_i \triangleq N_i \cdot \tilde{N}_i$ for $i = 1, 2$. For each $\bar{\mu}_1 \in [1, \bar{N}_1]$ and $\bar{\mu}_2 \in [1, \bar{N}_2]$, randomly and independently select $2^{n\tilde{R}_1}$ and $2^{n\tilde{R}_2}$ sequences $(U^{n,(\bar{\mu}_1)}(l), V^{n,(\bar{\mu}_2)}(k))$ according to the pruned distributions, i.e.,

$$\mathbb{P}\left((U^{n,(\bar{\mu}_1)}(l), V^{n,(\bar{\mu}_2)}(k)) = (u^n, v^n)\right) = \begin{cases} \dfrac{\lambda^A_{u^n}}{(1-\varepsilon)} \dfrac{\lambda^B_{v^n}}{(1-\varepsilon')} & \text{for} \quad u^n \in \mathcal{T}^{(n)}_\delta(U), v^n \in \mathcal{T}^{(n)}_\delta(V) \\ 0 & \text{otherwise} \end{cases}. \tag{2.42}$$

Let $\mathcal{C}^{(\bar{\mu}_1,\bar{\mu}_2)}$ denote the codebook containing all pairs of codewords $(U^{n,(\bar{\mu}_1)}(l), V^{n,(\bar{\mu}_2)}(k))$. Construct operators

$$A^{(\bar{\mu}_1)}_{u^n} \triangleq \gamma^{(\bar{\mu}_1)}_{u^n}\left(\sqrt{\rho_A}^{-1}\Lambda^A_{u^n}\sqrt{\rho_A}^{-1}\right) \quad \text{and} \quad B^{(\bar{\mu}_2)}_{v^n} \triangleq \zeta^{(\bar{\mu}_2)}_{v^n}\left(\sqrt{\rho_B}^{-1}\Lambda^B_{v^n}\sqrt{\rho_B}^{-1}\right), \tag{2.43}$$

where

$$\gamma^{(\bar{\mu}_1)}_{u^n} \triangleq \frac{1-\varepsilon}{1+\eta}2^{-n\tilde{R}_1}|\{l : U^{n,(\bar{\mu}_1)}(l) = u^n\}| \quad \text{and} \quad \zeta^{(\bar{\mu}_2)}_{v^n} \triangleq \frac{1-\varepsilon'}{1+\eta}2^{-n\tilde{R}_2}|\{k : V^{n,(\bar{\mu}_2)}(k) = v^n\}|, \tag{2.44}$$

where $\eta \in (0, 1)$ is a parameter that determines the probability of not obtaining sub-POVMs. Then, for each $\bar{\mu}_1 \in [1, \bar{N}_1]$ and $\bar{\mu}_2 \in [1, \bar{N}_2]$, construct $M^{(n,\bar{\mu}_1)}_1$ and $M^{(n,\bar{\mu}_2)}_2$ as in the following

$$M^{(n,\bar{\mu}_1)}_1 \triangleq \{A^{(\bar{\mu}_1)}_{u^n} : u^n \in \mathcal{T}^{(n)}_\delta(U)\}, \quad \text{and} \quad M^{(n,\bar{\mu}_2)}_2 \triangleq \{B^{(\bar{\mu}_2)}_{v^n} : v^n \in \mathcal{T}^{(n)}_\delta(V)\}. \tag{2.45}$$

We show later that $M^{(n,\bar{\mu}_1)}_1$ and $M^{(n,\bar{\mu}_2)}_2$ form sub-POVMs, with high probability, for all $\bar{\mu} \in [1, \bar{N}_1]$ and $\bar{\mu}_2 \in [1, \bar{N}_2]$, respectively. These collections $\tilde{M}^{(n,\bar{\mu}_1)}_1$ and $\tilde{M}^{(n,\bar{\mu}_2)}_2$ are completed using the operators $I - \sum_{u^n \in \mathcal{T}^{(n)}_\delta(U)} A^{(\bar{\mu}_1)}_{u^n}$ and $I - \sum_{v^n \in \mathcal{T}^{(n)}_\delta(V)} B^{(\bar{\mu}_2)}_{v^n}$, and these operators are associated with sequences u^n_0 and v^n_0, which are chosen arbitrarily from $\mathcal{U}^n \backslash \mathcal{T}^{(n)}_\delta(U)$ and $\mathcal{V}^n \backslash \mathcal{T}^{(n)}_\delta(V)$, respectively. For $(\tilde{\mu}_1, \tilde{\mu}_2) \in [1, \tilde{N}_1] \times [1, \tilde{N}_2]$, let $\mathbb{1}_{\{\text{sP-i}\}}(\tilde{\mu}_1, \tilde{\mu}_2)$ denote the indicator random variable corresponding to the event that $M^{(n,\mu_i,\tilde{\mu}_i)}_i$ form sub-POVM for all $\mu_i \in [1, N_i]$ for $i = 1, 2$. We use the trivial POVM $\{I\}$ in the case of the complementary event and associate it with u^n_0 and v^n_0 as the case

maybe. In summary, the POVMs are given by $\{\mathbb{1}_{\{\text{sP-1}\}}A_{u^n}^{(\bar{\mu}_1)} + (1-\mathbb{1}_{\{\text{sP-1}\}})\mathbb{1}_{\{u^n=u_0^n\}}I\}_{u^n\in\mathcal{U}^n}$, and $\{\mathbb{1}_{\{\text{sP-2}\}}B_{v^n}^{(\bar{\mu}_2)} + (1-\mathbb{1}_{\{\text{sP-2}\}})\mathbb{1}_{\{v^n=v_0^n\}}I\}_{v^n\in\mathcal{V}^n}$.

2.5.3.3 Binning of POVMs

Fix binning rates (R_1, R_2) and choose a $(\bar{\mu}_1, \bar{\mu}_2)$ 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 $\bar{\mu}_1 \in [1, \bar{N}_1]$ and $\bar{\mu}_2 \in [1, \bar{N}_2]$. For each $i \in [1, 2^{nR_1}]$ and $j \in [1, 2^{nR_2}]$, let $\mathcal{B}_1^{(\bar{\mu}_1)}(i)$ and $\mathcal{B}_2^{(\bar{\mu}_2)}(j)$ denote the i^{th} and the j^{th} bins, respectively. More precisely, $\mathcal{B}_1^{(\bar{\mu}_1)}(i)$ is the set of all u^n sequences with assigned index equal to i, and similar is $\mathcal{B}_2^{(\bar{\mu}_2)}(j)$. Moreover let $\iota_1^{(\bar{\mu}_1)} : \mathcal{T}_\delta^{(n)}(U) \to [1, 2^{nR_1}]$, and $\iota_2^{(\bar{\mu}_2)} : \mathcal{T}_\delta^{(n)}(V) \to [1, 2^{nR_2}]$, denote the corresponding random binning functions. Define the following operators:

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

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

$$M_A^{(n,\bar{\mu}_1)} \triangleq \{\Gamma_i^{A,(\bar{\mu}_1)}\}_{i\in[1,2^{nR_1}]},\ M_B^{(n,\bar{\mu}_2)} \triangleq \{\Gamma_j^{B,(\bar{\mu}_2)}\}_{j\in[1,2^{nR_2}]}.$$

Note that if $M_1^{(n,\bar{\mu}_1)}$ and $M_2^{(n,\bar{\mu}_2)}$ are sub-POVMs, then so are $M_A^{(n,\bar{\mu}_1)}$ and $M_B^{(n,\bar{\mu}_2)}$. This is due to the relations

$$\sum_i \Gamma_i^{A,(\bar{\mu}_1)} = \sum_{u^n\in\mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\bar{\mu}_1)}, \quad \sum_j \Gamma_j^{B,(\bar{\mu}_2)} = \sum_{v^n\in\mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\bar{\mu}_2)}.$$

To make $M_A^{(n,\bar{\mu}_1)}$ and $M_B^{(n,\bar{\mu}_2)}$ complete, we define $\Gamma_0^{A,(\bar{\mu}_1)}$ and $\Gamma_0^{B,(\bar{\mu}_2)}$ as $\Gamma_0^{A,(\bar{\mu}_1)} = I - \sum_i \Gamma_i^{A,(\bar{\mu}_1)}$ and $\Gamma_0^{B,(\bar{\mu}_2)} = I - \sum_j \Gamma_j^{B,(\bar{\mu}_2)}$, respectively[9]. In the event that the operators do not form sub-POVM, the sequence u_0^n and v_0^n are mapped to 0. Now, we intend to use the completions $[M_A^{(n,\bar{\mu}_1)}]$ and $[M_B^{(n,\bar{\mu}_2)}]$ 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).

[9] Note that $\Gamma_0^{A,(\bar{\mu}_1)} = I - \sum_i \Gamma_i^{A,(\bar{\mu}_1)} = I - \sum_{u^n\in T_\delta^{(n)}(U)} A_{u^n}^{(\bar{\mu}_1)}$ and $\Gamma_0^{B,(\bar{\mu}_2)} = I - \sum_j \Gamma_j^{B,(\bar{\mu}_2)} = I - \sum_{v^n\in T_\delta^{(n)}(V)} B_{v^n}^{(\bar{\mu}_2)}$.

2.5.3.4 Decoder mapping

We define a mapping $F^{(\bar{\mu}_1,\bar{\mu}_2)}$ acting on the outputs of $[M_A^{(n,\bar{\mu}_1)}] \otimes [M_B^{(n,\bar{\mu}_2)}]$ as follows. On observing $(\bar{\mu}_1, \bar{\mu}_2)$, and the classical indices $(i,j) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ communicated by the encoders, the decoder creates a set as follows:

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

For every $\bar{\mu}_i \in [1 : \bar{N}_i]$, $i \in [1 : 2^{nR_1}]$ and $j \in [1, 2^{nR_2}]$ define the function $F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j) = (u^n, v^n)$ if (u^n, v^n) is the only element of $D_{i,j}^{(\bar{\mu}_1,\bar{\mu}_2)}$; otherwise $F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j) = (u_0^n, v_0^n)$ Further, $F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j) = (u_0^n, v_0^n)$ for $i = 0$ or $j = 0$. Finally, the decoder produces $z^n \in \mathcal{Z}^n$ according to the stochastic map $P_{Z|UV}^n(z^n|F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j))$. With this mapping, we form the following collections of operators, for every $(\tilde{\mu}_1, \tilde{\mu}_2)$,

$$\tilde{\Lambda}_{u^n,v^n}^{AB}(\tilde{\mu}_1,\tilde{\mu}_2) \triangleq \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}} \frac{1}{N_1 N_2} \sum_{\mu_1=1}^{N_1}\sum_{\mu_2=1}^{N_2} \sum_{(i,j):F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j)=(u^n,v^n)} \Big(\Gamma_i^{A,(\bar{\mu}_1)} \otimes \Gamma_j^{B,(\bar{\mu}_2)} \Big) + (1 - \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}})(I \otimes I)\mathbb{1}_{\{(u^n,v^n)=(u_0^n,v_0^n)\}},$$

for all $(u^n, v^n) \in \mathcal{U}^n \times \mathcal{V}^n$. Note that for

$$\tilde{\Lambda}_{u^n,v^n}^{AB}(\tilde{\mu}_1,\tilde{\mu}_2) = 0 \text{ for } (u^n, v^n) \notin (\mathcal{T}_\delta^{(n)}(U) \times \mathcal{T}_\delta^{(n)}(V)) \bigcup \{(u_0^n, v_0^n)\}.$$

We use the stochastic mapping to define the approximating sub-POVM $\tilde{M}_{AB}^{(n)}(\tilde{\mu}_1,\tilde{\mu}_2) \triangleq \{\hat{\Lambda}_{z^n}(\tilde{\mu}_1,\tilde{\mu}_2)\}$ as

$$\hat{\Lambda}_{z^n}^{AB}(\tilde{\mu}_1,\tilde{\mu}_2) \triangleq \sum_{u^n,v^n} \tilde{\Lambda}_{u^n,v^n}^{AB}(\tilde{\mu}_1,\tilde{\mu}_2) P_{Z|U,V}^n(z^n|u^n,v^n),$$

$\forall z^n \in \mathcal{Z}^n$. The performance of the above ensemble is bounded from above as

$$\frac{1}{\tilde{N}_1\tilde{N}_2}\sum_{\tilde{\mu}_1,\tilde{\mu}_2} \Xi_{\rho_{AB}^{\otimes n}}(M_{AB}^{\otimes n}, \tilde{M}_{AB}^{(n)}(\tilde{\mu}_1,\tilde{\mu}_2)) \leq \frac{1}{\tilde{N}_1\tilde{N}_2}\sum_{\tilde{\mu}_1,\tilde{\mu}_2} \Big[G(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1,\tilde{\mu}_2))\mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1,\tilde{\mu}_2) + 2(1 - \mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1,\tilde{\mu}_2)) \Big],$$

where

$$G(\tilde{\mu}_1, \tilde{\mu}_2) \triangleq \sum_{z^n} \left\| \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B P_{Z|U,V}^n(z^n|u^n, v^n) - \tilde{\Lambda}_{u^n, v^n}^{AB}(\tilde{\mu}_1, \tilde{\mu}_2) P_{Z|U,V}^n(u^n, v^n) \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 . \tag{2.46}$$

In what follows, under the conditions on the rates given in the theorem, we show the existence of a pair $(\tilde{\mu}_1, \tilde{\mu}_2)$, and codebooks $\mathcal{C}^{(\bar{\mu}_1, \bar{\mu}_2)}$ and binning functions $\iota_i^{(\bar{\mu}_i)}$, for $\mu_i \in [1, N_i]$, $i = 1, 2$, such that the ϵ-faithfulness is satisfied for an arbitrary $\epsilon > 0$ for all sufficiently large n.

2.5.3.5 Performance Analysis

Step 0: Operators form sub-POVM: Fix an arbitrary $\epsilon > 0$. To start with, for all $(\tilde{\mu}_1, \tilde{\mu}_2) \in [1, \tilde{N}_1] \times [1, \tilde{N}_2]$, one can show using a result similar to Lemma II.10 the following proposition.

Proposition II.35 (sub-POVM)**.** *For any $\epsilon \in (0, 1)$, any $\eta \in (0, 1)$, any $\delta \in (0, 1)$ sufficiently small, and any n sufficiently large, we have*

$$\frac{1}{\tilde{N}_1 \tilde{N}_2} \sum_{\tilde{\mu}_1, \tilde{\mu}_2} 2 \left(1 - \mathbb{E} \left[\mathbb{1}_{\{sP\text{-}1\}}(\tilde{\mu}_1, \tilde{\mu}_2) \mathbb{1}_{\{sP\text{-}2\}}(\tilde{\mu}_1, \tilde{\mu}_2) \right] \right) < 2\epsilon,$$

if $\tilde{R}_1 > I(U; RB)_{\sigma_1}$ and $\tilde{R}_2 > I(V; RA)_{\sigma_2}$, where σ_1, σ_2 are defined as in the statement of the theorem.

Proof. We skip the proof for brevity. □

Next we focus on G.

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

Consider the second term within $G(\tilde{\mu}_1, \tilde{\mu}_2)$, which, under the event $\mathbb{1}_{\{\text{sP-1}\}} = 1$ and $\mathbb{1}_{\{\text{sP-2}\}} = 1$, can be written as

$$\sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \tilde{\Lambda}_{u^n, v^n}^{AB}(\tilde{\mu}_1, \tilde{\mu}_2)) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(u^n, v^n)$$
$$= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{i,j} \sqrt{\rho_{AB}^{\otimes n}} \left(\Gamma_i^{A,(\bar{\mu}_1)} \otimes \Gamma_j^{B,(\bar{\mu}_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|F^{(\bar{\mu}_1, \bar{\mu}_2)}(i, j)) \underbrace{\sum_{u^n, v^n} \mathbb{1}_{\{F^{(\bar{\mu}_1, \bar{\mu}_2)}(i,j) = (u^n, v^n)\}}}_{=1}$$

$$= T(\tilde{\mu}_1, \tilde{\mu}_2) + \widetilde{T}(\tilde{\mu}_1, \tilde{\mu}_2),$$

where

$$T(\tilde{\mu}_1, \tilde{\mu}_2) \triangleq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{i,j>0} \sqrt{\rho_{AB}^{\otimes n}} \left(\Gamma_i^{A,(\bar{\mu}_1)} \otimes \Gamma_j^{B,(\bar{\mu}_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | F^{(\bar{\mu}_1, \bar{\mu}_2)}(i,j)),$$

$$\widetilde{T}(\tilde{\mu}_1, \tilde{\mu}_2) \triangleq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{i=0 \text{ or } j=0} \sqrt{\rho_{AB}^{\otimes n}} \left(\Gamma_i^{A,(\bar{\mu}_1)} \otimes \Gamma_j^{B,(\bar{\mu}_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | u_0^n, v_0^n).$$

Hence, we have

$$G(\tilde{\mu}_1, \tilde{\mu}_2) \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \leq [S(\tilde{\mu}_1, \tilde{\mu}_2) + \widetilde{S}(\tilde{\mu}_1, \tilde{\mu}_2)] \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}}, \tag{2.47}$$

where $S(\tilde{\mu}_1, \tilde{\mu}_2) \triangleq$

$$\sum_{z^n} \left\| \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B P_{Z|U,V}^n(z^n | u^n, v^n) \right) \sqrt{\rho_{AB}^{\otimes n}} - T(\tilde{\mu}_1, \tilde{\mu}_2) \right\|_1, \tag{2.48}$$

and $\widetilde{S}(\tilde{\mu}_1, \tilde{\mu}_2) \triangleq \sum_{z^n} \|\widetilde{T}(\tilde{\mu}_1, \tilde{\mu}_2)\|_1$. Note that $\widetilde{S}$ captures the error induced by not covering the state $\rho_{AB}^{\otimes n}$.

Remark II.36. The terms corresponding to the operators that complete the sub-POVMs $M_A^{(n,\bar{\mu}_1)}$ and $M_B^{(n,\bar{\mu}_2)}$, i.e., $I - \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}^{(\bar{\mu}_1)}$ and $I - \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}^{(\bar{\mu}_2)}$ are taken care of in $\widetilde{T}$. The expression T excludes the completing operators. Therefore, in the analysis of the term S, we use $A_{u^n}^{(\bar{\mu}_1)}$ and $B_{v^n}^{(\bar{\mu}_2)}$ 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

Noting that $e^{(\bar{\mu}_1, \bar{\mu}_2)}(u^n, v^n) = F^{(\bar{\mu}_1, \bar{\mu}_2)}(i,j)$, for each $(u^n, v^n) \in \mathcal{B}_1^{(\bar{\mu}_1)}(i) \times \mathcal{B}_2^{(\bar{\mu}_2)}(j)$ and $(u^n, v^n) \in \mathcal{C}^{(\bar{\mu}_1, \bar{\mu}_2)}$. For any $(u^n, v^n) \notin \mathcal{C}^{(\bar{\mu}_1, \bar{\mu}_2)}$ let $e^{(\bar{\mu}_1, \bar{\mu}_2)}(u^n, v^n) = (u_0^n, v_0^n)$. This simplifies T as

$$T(\tilde{\mu}_1, \tilde{\mu}_2)$$
$$= \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{\substack{i>0, \\ j>0}} \sqrt{\rho_{AB}^{\otimes n}} \Bigg(\sum_{u^n \in B_1^{(\bar{\mu}_1)}(i)} A_{u^n}^{(\bar{\mu}_1)} \otimes \sum_{v^n \in B_2^{(\bar{\mu}_2)}(j)} B_{v^n}^{(\bar{\mu}_2)} \Bigg) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | F^{(\bar{\mu}_1, \bar{\mu}_2)}(i,j))$$

$$= \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(A_{u^n}^{(\bar{\mu}_1)} \otimes B_{v^n}^{(\bar{\mu}_2)} \right) \sqrt{\rho_{AB}^{\otimes n}}$$
$$\times \sum_{\substack{i>0,\\ j>0}} \mathbb{1}_{\left\{u^n \in B_1^{(\bar{\mu}_1)}(i), v^n \in B_2^{(\bar{\mu}_2)}(j)\right\}} P_{Z|U,V}^n(z^n | e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n, v^n))$$
$$= \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(A_{u^n}^{(\bar{\mu}_1)} \otimes B_{v^n}^{(\bar{\mu}_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n, v^n)),$$

where we use the fact that $\sum_{u^n \in B_1^{(\bar{\mu}_1)}(i)} A_{u^n}^{(\bar{\mu}_1)} = \sum_{u^n} A_{u^n}^{(\bar{\mu}_1)} \mathbb{1}_{\left\{u^n \in B_1^{(\bar{\mu}_1)}(i)\right\}}$ and $\sum_{i>0} \mathbb{1}_{\left\{u^n \in B_1^{(\bar{\mu}_1)}(i)\right\}} = 1$ for all $u^n \in \mathcal{T}_\delta^{(n)}(U)$, and a similar argument holds for the sub-POVM $\{B_{v^n}^{(\bar{\mu}_2)}\}$. 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 (2.48) to obtain

$$S(\tilde{\mu}_1, \tilde{\mu}_2) = \sum_{z^n} \Bigg\| \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | u^n, v^n)$$
$$- \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sqrt{\rho_{AB}^{\otimes n}} \left(A_{u^n}^{(\bar{\mu}_1)} \otimes B_{v^n}^{(\bar{\mu}_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n, v^n)) \Bigg\|_1 .$$

Recall that $\bar{\mu}_i = (\mu_i, \tilde{\mu}_i)$ for $i = 1, 2$. 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

$$S_1(\tilde{\mu}_1, \tilde{\mu}_2) \triangleq \sum_{z^n} \Bigg\| \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \Bigg(\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B - \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} A_{u^n}^{(\bar{\mu}_1)} \otimes B_{v^n}^{(\bar{\mu}_2)} \Bigg) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n | u^n, v^n) \Bigg\|_1,$$
$$S_2(\tilde{\mu}_1, \tilde{\mu}_2) \triangleq \sum_{z^n} \Bigg\| \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(A_{u^n}^{(\bar{\mu}_1)} \otimes B_{v^n}^{(\bar{\mu}_2)} \right) \sqrt{\rho_{AB}^{\otimes n}}$$
$$\times \left(P_{Z|U,V}^n(z^n | u^n, v^n) - P_{Z|U,V}^n \left(z^n | e^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n, v^n) \right) \right) \Bigg\|_1 . \tag{2.49}$$

This gives

$$G \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \leq [S_1 + S_2 + \widetilde{S}] \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}}.$$

Note that the term S_1 characterizes the error introduced by approximation of the original POVM with the collection of approximating sub-POVMs $M_1^{(n,\bar{\mu}_1)}$ and $M_2^{(n,\bar{\mu}_2)}$, and the term S_2 characterizes the error caused by binning of these approximating sub-POVMs. Next, we analyze S_2 and prove the following proposition.

Proposition II.37 (Mutual Packing)**.** *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large n, we have*

$$\frac{1}{\tilde{N}_1 \tilde{N}_2} \sum_{\tilde{\mu}_1, \tilde{\mu}_2} \mathbb{E}\left[S_2(\tilde{\mu}_1, \tilde{\mu}_2) \mathbb{1}_{\{sP\text{-}1\}}(\tilde{\mu}_1, \tilde{\mu}_2) \mathbb{1}_{\{sP\text{-}2\}}(\tilde{\mu}_1, \tilde{\mu}_2) \right] < 5\epsilon,$$

if $\tilde{R}_1 > I(U;RB)_{\sigma_1}$, $\tilde{R}_2 > I(V;RA)_{\sigma_2}$, $\tilde{R}_1 + \frac{1}{n}\log(\bar{N}_1) > S(U)_{\sigma_3}$, $\tilde{R}_2 + \frac{1}{n}\log(\bar{N}_2) > S(V)_{\sigma_3}$, $\tilde{R}_1 + \tilde{R}_2 - R_1 - R_2 < I(U;V)_{\sigma_3}$, where σ_i for $i = 1,2,3$, is the auxiliary state defined in the theorem.

Proof. The proof is provided in Appendix A.8 □

Hence there must exist a pair $(\tilde{\mu}_1, \tilde{\mu}_2)$ such that

$$\begin{aligned} &2\left(1 - \mathbb{E}\left[\mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1, \tilde{\mu}_2) \mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1, \tilde{\mu}_2))\right]\right) \\ &\quad + \mathbb{E}\left[S_2(\tilde{\mu}_1, \tilde{\mu}_2) \mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1, \tilde{\mu}_2) \mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1, \tilde{\mu}_2) \right] < 7\epsilon, \end{aligned}$$

for the rates satisfying the constraints in Propositions II.35 and II.37. For the rest of the proof, we fix $(\tilde{\mu}_1, \tilde{\mu}_2)$ to be this pair. The dependence of functions defined in the sequel on this pair is not made explicit for ease of notation.

Remark II.38. Since the shared randomness given by $(\tilde{\mu}_1, \tilde{\mu}_2)$ is only used for random coding purposes, two of the constraints in Proposition II.37, given by $\tilde{R}_1 + \frac{1}{n}\log(\bar{N}_1) > S(U)_{\sigma_3}$, $\tilde{R}_2 + \frac{1}{n}\log(\bar{N}_2) > S(V)_{\sigma_3}$, are superfluous.

For the term corresponding to $\widetilde{S}$, we prove the following result.

Proposition II.39. *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have*

$$\mathbb{E}\left[\widetilde{S} \mathbb{1}_{\{sP\text{-}1\}} \mathbb{1}_{\{sP\text{-}2\}} \right] < 8\epsilon,$$

if $\tilde{R}_1 > I(U;RB)_{\sigma_1}$ and $\tilde{R}_2 > I(V;RA)_{\sigma_2}$, where σ_1 and σ_2 are auxiliary states defined in the theorem.

Proof. The proof is provided in Appendix A.9. □

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

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

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

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

$$J \triangleq \sum_{z^n, v^n} \left\| \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B - \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} A_{u^n}^{(\mu_1)} \otimes \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n, v^n) \right\|_1. \tag{2.50}$$

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 σ_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 all 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.

Proposition II.40. *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[Q_1] \leq \mathbb{E}[J] < 2\epsilon$, if $\tilde{R}_1 + C_1 > I(U;RZV)_{\sigma_3}$, where the auxiliary state σ_3 is defined in the theorem.*

Proof. The proof is provided in Appendix A.10. □

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 II.41 (Non-product Covering Lemma)**.** *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have*

$$\mathbb{E}\left[Q_2 \mathbb{1}_{\{sP\text{-}1\}} \mathbb{1}_{\{sP\text{-}2\}}\right] < 4\epsilon,$$

if $\tilde{R}_1 + C_1 > I(U;RZV)_{\sigma_3}$, and $\tilde{R}_2 + C_2 > I(V;RZ)_{\sigma_3}$, where the auxiliary state σ_3 is defined in the theorem.

Proof. The proof is provided in Appendix A.11. □

2.5.3.6 Rate Constraints

To sum-up, we showed that the trace distance satisfies:

$$\Xi_{\rho_{AB}^{\otimes n}}(M_{AB}^{\otimes n}, \tilde{M}_{AB}^{(n)}(\tilde{\mu}_1, \tilde{\mu}_2)) \leq 21\epsilon,$$

if the following bounds hold:

$$\begin{aligned}
\tilde{R}_1 > I(U;RB)_{\sigma_1}, \quad \tilde{R}_2 > I(V;RA)_{\sigma_2}, \\
\tilde{R}_1 + C_1 > I(U;RZV)_{\sigma_3}, \quad \tilde{R}_2 + C_2 > I(V;RZ)_{\sigma_3}, \\
(\tilde{R}_1 - R_1) + (\tilde{R}_2 - R_2) < I(U;V)_{\sigma_3}, \\
\tilde{R}_1 \geq R_1 \geq 0, \quad \tilde{R}_2 \geq R_2 \geq 0, \quad C_1 \geq 0, \quad C_2 \geq 0.
\end{aligned} \tag{2.51}$$

Let us denote the above achievable rate-region by $\mathcal{R}_1$. By doing an exact symmetric analysis, but by replacing the first encoder by a product distribution instead of the second encoder in S_1 (as defined in (2.49)), all the constraints remain the same, except that the constraints on $\tilde{R}_1 + C_1$ and

$\tilde{R}_2 + C_2$ change as follows

$$\tilde{R}_1 + C_1 \geq I(U; RZ)_{\sigma_3}, \quad \tilde{R}_2 + C_2 \geq I(V; RZU)_{\sigma_3}. \tag{2.52}$$

Let us denote the above region by $\mathcal{R}_2$. By time sharing between the any two points of $\mathcal{R}_1$ and $\mathcal{R}_2$ one can achieve any point in the convex closure of $(\mathcal{R}_1 \bigcup \mathcal{R}_2)$. The following lemma gives a symmetric characterization of the closure of convex hull of the union of the above achievable rate-regions.

Lemma II.42. *For the above defined rate regions $\mathcal{R}_1$ and $\mathcal{R}_2$, we have*

$$\mathcal{R}_3 = \textit{Convex Closure}(\mathcal{R}_1 \bigcup \mathcal{R}_2),$$

where $\mathcal{R}_3$ is given by the set of all the sextuples $(\tilde{R}_1, \tilde{R}_2, R_1, R_2, C_1, C_2)$ satisfying the following constraints:

$$\begin{aligned}
\tilde{R}_1 \geq I(U; RB)_{\sigma_1}, \quad \tilde{R}_2 \geq I(V; RA)_{\sigma_2}, \\
\tilde{R}_1 + C_1 \geq I(U; RZ)_{\sigma_3}, \quad \tilde{R}_2 + C_2 \geq I(V; RZ)_{\sigma_3}, \\
\tilde{R}_1 + \tilde{R}_2 + C_1 + C_2 \geq I(U; RZ)_{\sigma_3} + I(V; RZ)_{\sigma_3} + I(U; V|RZ)_{\sigma_3}, \\
\tilde{R}_1 + \tilde{R}_2 - (R_1 + R_2) \leq I(U; V)_{\sigma_3} \\
0 \leq R_1 \leq \tilde{R}_1 \quad 0 \leq R_2 \leq \tilde{R}_2 \quad C_1 \geq 0, C_2 \geq 0.
\end{aligned} \tag{2.53}$$

Proof. The proof follows from elementary convex analysis. □

Lemma II.43. *Let $\bar{\mathcal{R}}_3$ denote the set of all quadruples (R_1, R_2, C_1, C_2) 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 (2.53). Let $\mathcal{R}_F$ denote the set of all quadruples (R_1, R_2, C_1, C_2) that satisfy the inequalities in (2.36) given in the statement of the theorem. Then, $\bar{\mathcal{R}}_3 = \mathcal{R}_F$.*

Proof. This follows by Fourier-Motzkin elimination *Ziegler* (2012). □

2.6 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. In the next chapter, we aim to improve the achievable rate region by using structured POVMs based on algebraic codes.

CHAPTER III

Algebraic Structured Distributed Measurement Compression

3.1 Introduction

We begin this chapter by asking the following question: If one is interested in solely reconstructing a given *function* of the distributively stored measurement outcomes, can the rate of communication be further reduced? This chapter answers this in affirmative. For this, we employ structured coding techniques and impose further structure on these approximating POVMs. This ensures that the joint decoder (Charlie) is able to reconstruct a lower dimensional quantum state with minimal use of the classical communication resource. In particular, the structure of the POVM is aligned to match with the structure of the *function* being computed.

As been highlighted earlier, there are two conventional approaches for deriving performance limits of communication problems in the stationary memoryless setting. One is based on random coding involving IID code ensembles, and characterizing the performance in terms of single-letter information quantities. And the other is based on random coding in the one-shot setting, and characterizing the performance using smooth entropic quantities. Since there is no global structure in the codes in these ensemble, we refer to them as random coding techniques based on unstructured code ensembles. The work by Korner-Marton *Korner and Marton* (1979) demonstrated that unstructured code ensembles may not always achieve optimality for distributed multi-terminal settings. In particular, this work showed sub-optimality for the problem of classical distributed lossless compression with the objective of computing the sum of the sources. The standard approach that conventional codes take is to recover the individual messages and then compute the function. However, it is known that this technique cannot characterize the performance limit. Korner and Marton

considered a special binary symmetric case, and using random linear codes (based on finite fields) derived the performance limits in terms of single-letter information quantities. Their idea was to perform coding in a way that allows computing the function directly without the need to recover each individual messages.

Traditionally, algebraic-structured codes have been used extensively in information coding problems toward achieving computationally efficient (polynomial-time) encoding and decoding algorithms. However, in multi-terminal communication problems, even if computational complexity is a non-issue (which is generally the assumption in an information theoretic setting), there are many instances were random algebraic structured codes can be preferred. The works in *Krithivasan and Pradhan* (2011); *Nazer and Gastpar* (2007); *Philosof and Zamir* (2009); *Jafarian and Vishwanath* (2012) have considered multi-terminal setups and showed that random algebraic structured codes outperform random unstructured codes in terms of achieving improved asymptotic rate regions.

Motivated by this, in this chapter, we consider the quantum distributed faithful measurement simulation problem in the memoryless setting using algebraic structured coding techniques. There are two main challenges in using these algebraic structured codes toward an asymptotic analysis in quantum information theory. The first challenge is to be able to induce arbitrary empirical single-letter distributions. For example, if we were to send codewords from a linear code with uniform probability, then the induced empirical distribution of codeword symbols (single-letter distribution on the symbols of the codewords) is uniform. To address this challenge, we use a collection of cosets of a linear code called Unionized Coset Codes (UCCs) *Pradhan et al.* (2021). The second challenge is that unlike the random unstructured codes, the codewords generated from a random linear code are only pairwise-independent *Gallager* (1968). This renders the some of techniques that are developed for standard measurement compression problem unusable. These mainly include the technique of operator Chernoff bound and the covering lemma. Similarly, the mutual covering and the mutual packing lemmas developed in Chapter II also require a new analysis. Since our approach relies on the use of UCCs for generating the approximating POVMs, the binning of these POVM elements is performed in a correlated fashion as governed by these structured codes. This is in contrast to the common technique of independent binning. Due to the correlated binning, the pairwise-independence issue gets exacerbated.

We address these challenges using three main ideas summarized as follows:

- **Random structured generation of pruned POVMs** - We generate a collection of algebraic structured approximating POVMs randomly using the above described UCC technique, and then prune them. This pruning ensures that these POVMs form a positive resolution of identity, and thus eliminates any need for the operator Chernoff inequality. However, such pruning comes at the cost of additional approximating error. To bound the approximating error caused by pruning the POVMs, we develop a new operator inequality which provides a handle to convert the pruning error in the form of covering error expression which is dealt within the next idea.

- **Covering lemma with change of measure for pairwise-independent ensemble** - Since the traditional covering lemma is based on the Chernoff inequality, we develop an alternative form of the aforementioned covering lemma (*Wilde*, 2013a, Lemma 17.2.1). This alternative form is based on a second-order analysis using the operator trace inequalities and hence requires the operators to be only pairwise-independent.

- **Multi-partite Packing Lemma** - We develop a binning technique for performing computation on the fly so as to achieve a low dimensional reconstruction of a function at the location of Charlie. In an effort towards analysing this binning technique, we develop a multi-partite packing Lemma for the structured POVMs.

Combining these techniques, we provide a multi-party distributed faithful simulation and function computation protocol in a quantum information theoretic setting. We provide a characterization of the asymptotic performance limit of this problem in terms of an inner bound expressed using computable single-letter quantum information quantities, which is the main result of the paper (see Theorem III.22). This new inner bound subsumes the inner bound developed in Theorem II.33 of Chapter II, derived using random unstructured coding techniques. Further, we identify examples where the current inner bound strictly improves upon the former.

The organization of the paper is as follows. In Section 3.2, we set the notation, state requisite definitions, and also provide related results. In Section 3.3, for pedagogical reasons, we consider the point-to-point setup, and provide a theorem characterizing the rate-region using algebraic structured codes, while developing a new Covering lemma with change of measure for pairwise-independent ensembles in Section 3.3.2. In Section 3.4, we state our main result of this chapter on

the distributed measurement compression and provide the theorem (Theorem III.22) characterizing the rate-region. We prove the main result (Theorem III.22) in Section 3.5 using the point-to-point result as a building block. Finally, we conclude the chapter in Section 3.6.

3.2 Preliminaries

Notation: We use the notation developed in Chapters 1.5 and II. In addition, let $\mathbb{F}_p$ denote the prime finite field of size p with addition operation given by $+$.

We recall Lemma II.11 below for convenience.

Lemma III.1. *Given a density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, 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 Γ^A acting on $\mathcal{H}_A$, we have*

$$\sum_{y\in\mathcal{Y}} \left\| \sqrt{\rho_{AB}} \left(\Gamma^A \otimes \Lambda_y^B\right) \sqrt{\rho_{AB}} \right\|_1 \leq \left\| \sqrt{\rho_A}\Gamma^A\sqrt{\rho_A} \right\|_1, \tag{3.1}$$

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

Below we present a lemma which will be used extensively in the sequel.

Definition III.2 (Pruning Operator)**.** Consider an operator $A \geq 0$ acting on Hilbert space $\mathcal{H}_A$. We say that a projector P prunes A with respect to Identity I_A on $\mathcal{H}_A$, if P is a projector on to the non-negative eigenspace of $I_A - A$. Further, observe that $\mathrm{Tr}\{I_A - P\} \leq \mathrm{Tr}\{A\}$.

Lemma III.3. *(Pruning Trace Inequality) Consider the above random operator $X \geq 0$ acting on a Hilbert space $\mathcal{H}_A$. Further, suppose $\mathbb{E}[X] \leq (1-\eta)I_A$ for $\eta \in (0,1)$. Let P be a pruning operator for X with respect to I_A, as in Definition III.2. Then, we have*

$$\mathbb{E}[\mathrm{Tr}\{I_A - P\}] \leq \frac{1}{\eta}\mathbb{E}\left[\|X - \mathbb{E}[X]\|_1\right].$$

Proof. Note that $X \geq 0$, implies $\mathbb{E}[X] \geq 0$. Therefore, if P prunes X, then $PXP \leq P$, and $P(X - \mathbb{E}[X])P \leq P$, and thus P also prunes $X - \mathbb{E}[X]$. This implies

$$\mathrm{Tr}\{I_A - P\} \leq \mathrm{Tr}\{X - \mathbb{E}[X]\} \leq \mathrm{Tr}\{|X - \mathbb{E}[X]|\} = \|X - \mathbb{E}[X]\|. \tag{3.2}$$

□

3.3 Point-to-point Measurement Compression using Structured Random POVMs

Before presenting the main result of this chapter, i.e., a distributed measurement compression theorem using algebraic structured code ensembles, as a pedagogical first step, we consider the non-feedback measurement compression problem in the point-to-point setup. This problem was addressed in *Wilde et al.* (2012), where the performance limits were derived using unstructured random POVM ensembles. An alternate approach for this problem was also developed in Chapter II (Section 2.4) using unstructured coding ensembles. Here, we redrive the performance limit using random algebraic structured POVM ensembles. Since the algebraic structured codes can only induce a uniform distribution, we consider a collection of cosets of a random linear code for this task. The problem setup is described as follows. An agent (Alice) performs a measurement M on a quantum state ρ, and sends a set of classical bits to a receiver (Bob). 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. 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. This problem serves as a building block toward the proof of the main result of this chapter (Theorem III.22). Formally, the problem is stated as:

3.3.1 Problem Formulation and Main Result

Definition III.4. For a given finite set $\mathcal{Z}$, and a Hilbert space $\mathcal{H}$, a measurement simulation protocol with parameters (n, Θ, N) is characterized by

1) a collection of codes $\mathcal{C}^{(\mu)} \subseteq \mathcal{W}^n$, for $\mu \in [1, N]$, such that $|\mathcal{C}^{(\mu)}| \leq \Theta$, and $\mathcal{W}$, a finite set, is called the code alphabet,

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

3) a collection of Bob's classical stochastic maps $P^{(\mu)}(z^n|w^n)$ for all $w^n \in \mathcal{C}^{(\mu)}$, $z^n \in \mathcal{Z}^n$ and $\mu \in [1, N]$.

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

$$\tilde{\Lambda}_{z^n} \triangleq \frac{1}{N} \sum_{\mu=1}^{N} \sum_{w^n \in \mathcal{C}^{(\mu)}} P^{(\mu)}(z^n|w^n)\, \Lambda_{w^n}^{(\mu)}, \quad \forall z^n \in \mathcal{Z}^n, \tag{3.3}$$

where $\{\Lambda_{w^n}^{(\mu)} : w^n \in \mathcal{C}^{(\mu)}\}$ is the set of operators corresponding to the sub-POVM $\tilde{M}^{(\mu)}$. Let $\mathcal{C}^{(\mu)}(i)$ denote the ith codeword of $\mathcal{C}^{(\mu)}$.

In the above definition, Θ characterizes the amount of classical bits communicated from Alice to Bob, and the amount of common randomness is determined by N, with μ 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. In building the code, we use the Unionized Coset Code (UCC) *Pradhan et al.* (2021) defined below. These codes involve two layers of codes (i) a coarse code and (ii) a fine code. The coarse code is a coset of the linear code and the fine code is the union of several cosets of the linear code.

For a fixed $k \times n$ matrix $G \in \mathbb{F}_p^{k\times n}$ with $k \leq n$, and p being a prime number, and a $1 \times n$ vector $B \in \mathbb{F}_p^n$, define the coset code as

$$\mathbb{C}(G, B) \triangleq \{x^n : x^n = a^k G + B, \text{ for some } a^k \in \mathbb{F}_p^k\}. \tag{3.4}$$

In other words, $\mathbb{C}(G, B)$ is a shift of the row space of the matrix G. The row space of G is a linear code. If the rank of G is k, then there are p^k codewords in the coset code.

Definition III.5. An (n, k, l, p) UCC is characterized by a pair (G, h) consisting of a $k \times n$ matrix $G \in \mathbb{F}_p^{k\times n}$, and a mapping $h : \mathbb{F}_p^l \to \mathbb{F}_p^n$, and the code is the following union: $\bigcup_{m\in\mathbb{F}_p^l} \mathbb{C}(G, h(m))$, where $\mathbb{C}(\cdot, \cdot)$ is defined in (3.4).

Definition III.6. Given a finite set $\mathcal{Z}$, and a Hilbert space $\mathcal{H}$, an $(n, \Theta, \kappa, N, p)$ UCC-based measurement simulation protocol is a pair of (n, Θ, N) measurement simulation protocol and a collection of N UCCs with parameters (n, k, l, p) characterized by $\{(G, h^{(\mu)})\}_{\mu\in[1,N]}$ such that (i) the code alphabet of the protocol $\mathcal{W} \subseteq \mathbb{F}_p$ (with suitable relabeling), (ii) $\kappa = p^k$, $\Theta = p^l$, and (iii) for all $m \in \mathbb{F}_p^l$, we have $\mathcal{C}^{(\mu)}(m) \in \{a^k G + h^{(\mu)}(m) : a^k \in \mathbb{F}_p^k\}$.

Definition III.7. The UCC grand ensemble is the ensemble of N UCCs where G, and $\{h^{(\mu)}\}_{\mu\in[1,N]}$

are chosen randomly, independently and uniformly, where the latter is chosen from the set of all mappings with replacement.

Definition III.8. Given a POVM M acting on $\mathcal{H}$, and a density operator $\rho \in \mathcal{D}(\mathcal{H})$, a tuple (R, R_1, C, p) is said to be achievable using the grand UCC ensemble, if for all $\epsilon > 0$ and for all sufficiently large n, there exists an ensemble of UCC-based measurement simulation protocols with parameters $(n, \Theta, \kappa, N, p)$ (based on the UCC grand ensemble) such that their overall sub-POVM $\tilde{M}$ is ϵ-faithful to $M^{\otimes n}$ with respect to $\rho^{\otimes n}$ in the expected sense:

$$\mathbb{E}\left[\sum_{z^n}\left\|\sqrt{\rho^{\otimes n}}(\Lambda_{z^n}-\tilde{\Lambda}_{z^n})\sqrt{\rho^{\otimes n}}\right\|+\mathrm{Tr}\{I-\sum_{z^n}\tilde{\Lambda}_{z^n}\}\right] \leq \epsilon,$$

where the expectation is with respect to the ensemble, and

$$\frac{1}{n}\log_2\Theta \leq R+\epsilon, \quad \left|\frac{1}{n}\log\kappa - R_1\right| \leq \epsilon, \quad \frac{1}{n}\log_2 N \leq C+\epsilon.$$

Define $\mathscr{R}_{\mathrm{UCC}}$ as $\mathscr{R}_{\mathrm{UCC}} \triangleq \{(R, R_1, C, p) : (R, R_1, C, p)$ is achievable using the UCC grand ensemble$\}$.

Remark III.9. The appearance of the modulus in the second constraint needs justification. Note that R is the rate of transmission of information from Alice to Bob and C is the rate of the common information shared between them. So if (R, R_1, C, p) is achievable, then it is clear that any $(\tilde{R}, \tilde{C})$ is also achievable if $\tilde{R} \geq R$ and $\tilde{C} \geq C$. However R_1 is a parameter of the UCC grand ensemble, and there is no natural order on R_1, i.e., it does not naturally follows that $(R, \tilde{R}_1, C, p)$ is achievable for all $\tilde{R}_1 \geq R_1$.

The following theorem characterizes the achievable rate region which characterizes the asymptotic performance of the UCC grand ensemble.

Theorem III.10. *For any density operator $\rho \in \mathcal{D}(\mathcal{H})$ and any POVM $M \triangleq \{\Lambda_z\}_{z\in\mathcal{Z}}$ acting on the Hilbert space $\mathcal{H}$, a tuple (R, R_1, C, p) is achievable using the UCC grand ensemble, i.e., $(R, R_1, C, p) \in \mathscr{R}_{UCC}$ if there exist a POVM $\bar{M} \triangleq \{\bar{\Lambda}_w\}_{w\in\mathcal{W}}$, with $|\mathcal{W}| \leq p$, and a stochastic map $P_{Z|W} : \mathcal{W} \to \mathcal{Z}$ such that $\Lambda_z = \sum_{w\in\mathcal{W}} P_{Z|W}(z|w)\bar{\Lambda}_w$, $\forall z \in \mathcal{Z}$, and the following holds:*

$$R_1 + R \geq I(W;R)_\sigma - S(W)_\sigma + \log p,$$

$$R_1 + R + C \geq I(W;RZ)_\sigma - S(W)_\sigma + \log p,$$
$$R_1 \leq \log p - S(W)_\sigma, \tag{3.5}$$

and $R_1, R, C \geq 0$*, where* $\sigma^{RWZ} \triangleq \sum_{w,z} \sqrt{\rho}\bar{\Lambda}_w\sqrt{\rho} \otimes P_{Z|W}(z|w)\,|w\rangle\langle w| \otimes |z\rangle\langle z|$*, for some orthogonal sets* $\{|w\rangle\}_{w\in\mathcal{W}}$ *and* $\{|z\rangle\}_{z\in\mathcal{Z}}$*.*

Remark III.11. Choosing $R_1 = \log p - S(W)_\sigma$, we recover the rate region of Wilde et. al (*Wilde et al.*, 2012, Theorem 9).

Proof. The proof is provided in Section 3.3.3. □

3.3.2 Covering Lemma with Change of Measure for Pairwise-Independent Ensemble

The proof of the theorem is based on a construction of algebraic-structured POVM ensemble where the elements are only pairwise independent and not mutually independent. To analyze these POVMs we retreat back to first principles and develop a new one-shot Covering Lemma based on a change of measure technique and a second order analysis. This lemma, which can be of independent interest, is one of the main contributions of this work. Although we have developed this lemma in a one-shot setting, this forms a minor component in the proof of the faithful simulation result, which is realized in the n-letter setting using asymptotically good coset codes.

Lemma III.12 (Covering Lemma)**.** *Let* $\{\lambda_x, \sigma_x\}_{x\in\mathcal{X}}$ *be an ensemble, with* $\sigma_x \in \mathcal{D}(\mathcal{H})$ *for all* $x \in \mathcal{X}$*,* $\mathcal{X}$ *being a finite set, and* $\sigma = \sum_{x\in\mathcal{X}} \lambda_x\sigma_x$*. Further, suppose we are given a total subspace projector* Π *and a collection of codeword subspace projectors* $\{\Pi_x\}_{x\in\mathcal{X}}$ *which satisfy the following hypotheses*

$$\mathrm{Tr}\{\Pi\sigma_x\} \geq 1-\epsilon, \tag{3.6a}$$
$$\mathrm{Tr}\{\Pi_x\sigma_x\} \geq 1-\epsilon, \tag{3.6b}$$
$$\|\Pi\sqrt{\sigma}\|_1^2 \leq D, \tag{3.6c}$$
$$\Pi_x\sigma_x\Pi_x \leq \frac{1}{d}\Pi_x, \quad \textit{and} \tag{3.6d}$$
$$\Pi_x\sigma_x\Pi_x \leq \sigma_x. \tag{3.6e}$$

for some $\epsilon \in (0,1)$ *and* $d < D$*. Let* M *be a finite non-negative integer. Additionally, assume*

that there exists some set $\bar{\mathcal{X}}$ *containing* $\mathcal{X}$*, with* $\sigma_x \triangleq 0$ *(null operator) and* $\lambda_x \triangleq 0$ *for* $x \in \bar{\mathcal{X}} \backslash \mathcal{X}$*. Suppose* $\{\mu_{\bar{x}}\}_{\bar{x}\in\bar{\mathcal{X}}}$ *be any distribution on the set* $\bar{\mathcal{X}}$ *such that the distribution is* $\{\lambda_x\}_{x\in\mathcal{X}}$ *is absolutely continuous with respect to the distribution* $\{\mu_{\bar{x}}\}_{\bar{x}\in\bar{\mathcal{X}}}$*. Further, assume that* $\lambda_x/\mu_x \leq \kappa$ *for all* $x \in \mathcal{X}$*. Let a random covering code* $\mathbb{C} \triangleq \{C_m\}_{m\in[1,M]}$ *be defined as a collection of codewords* C_m *that are chosen pairwise independently according to the distribution* $\{\mu_{\bar{x}}\}_{\bar{x}\in\bar{\mathcal{X}}}$*. Then we have*

$$\mathbb{E}_{\mathbb{C}}\left[\Big\| \sum_{x\in\bar{\mathcal{X}}} \lambda_x \sigma_x - \frac{1}{M}\sum_{m=1}^{M} \frac{\lambda_{C_m}}{\mu_{C_m}} \sigma_{C_m} \Big\|_1\right] \leq \sqrt{\frac{\kappa D}{Md}} + 2\delta(\epsilon), \tag{3.7}$$

where $\delta(\epsilon) = 4\sqrt{\epsilon}$*. Futhermore, for* $\tilde{\sigma}_x$ *defined as* $\tilde{\sigma}_x \triangleq \Pi\Pi_x \sigma_x \Pi_x \Pi$*, we have*

$$\mathbb{E}_{\mathbb{C}}\left[\Big\| \sum_{x\in\bar{\mathcal{X}}} \lambda_x \tilde{\sigma}_x - \frac{1}{M}\sum_{m=1}^{M} \frac{\lambda_{C_m}}{\mu_{C_m}} \tilde{\sigma}_{C_m} \Big\|_1\right] \leq \sqrt{\frac{\kappa D}{Md}}. \tag{3.8}$$

Proof. The proof is provided in Appendix B.1 □

3.3.3 Proof of the Main Result (Theorem III.10) Using UCC Code Ensemble

As stated earlier, the main objective of proving this theorem is to build a framework for the main theorem of the paper (Theorem III.22). In doing so, we observe that the structured POVMs constructed below are only pairwise independent. Since the results in *Atif et al.* (2021a) are based on the assumption that approximating POVMs are all mutually independent, the proof below becomes significantly different from *Atif et al.* (2021a).

Suppose there exist a POVM $\bar{M} \triangleq \{\bar{\Lambda}_w\}_{w\in\mathcal{W}}$ and a stochastic map $P_{Z|W} : \mathcal{W} \to \mathcal{Z}$, such that $M \triangleq \{\Lambda_z\}_{z\in\mathcal{Z}}$ can be decomposed as

$$\Lambda_z \triangleq \sum_{w\in\mathcal{W}} P_{Z|W}(z|w)\bar{\Lambda}_w, \quad \forall z \in \mathcal{Z}. \tag{3.9}$$

We generate the canonical ensemble corresponding to $\bar{M}$ as

$$\lambda_w \triangleq \mathrm{Tr}\{\bar{\Lambda}_w \rho\}, \quad \hat{\rho}_w \triangleq \frac{1}{\lambda_w}\sqrt{\rho}\bar{\Lambda}_w\sqrt{\rho}. \tag{3.10}$$

Let $\mathcal{T}_\delta^{(n)}(W)$ denote a δ-typical set associated with the probability distribution induced by $\{\lambda_w\}_{w\in\mathcal{W}}$,

corresponding to a random variable W. Let Π_ρ denote the δ-typical projector (as in (*Wilde*, 2013a, Def. 15.1.3)) corresponding to the density operator $\rho \triangleq \sum_{w\in\mathcal{W}} \lambda_w \hat{\rho}_w$, and Π_{w^n} denote the strong conditional typical projector (as in (*Wilde*, 2013a, Def. 15.2.4)) corresponding to the canonical ensemble $\{\lambda_w, \hat{\rho}_w\}_{w\in\mathcal{W}}$. For each $w^n \in \mathcal{T}_\delta^{(n)}(W)$, define

$$\tilde{\rho}_{w^n} \triangleq \Pi_\rho \Pi_{w^n} \hat{\rho}_{w^n} \Pi_{w^n} \Pi_\rho,$$

and $\tilde{\rho}_{w^n} = 0$, for $w^n \notin \mathcal{T}_\delta^{(n)}(W)$, with $\hat{\rho}_{w^n} \triangleq \bigotimes_i \hat{\rho}_{w_i}$.

3.3.3.1 Construction of Structured POVMs

We now construct random structured POVM elements. Fix a block length $n > 0$, a positive integer N, and a finite field $\mathbb{F}_p$ with $p \geq |\mathcal{W}|$. Without loss of generality, we assume $\mathcal{W} \triangleq \{0, 1, \cdots, |\mathcal{W}| - 1\}$. Furthermore, we assume $\lambda_w = 0$ for all $|\mathcal{W}| - 1 < w < p$. From now on, we assume that W takes values in $\mathbb{F}_p$ with this distribution. Let $\mu \in [1, N]$ denote the common randomness shared between the encoder and decoder. In building the code, we use the UCCs *Pradhan et al.* (2021) as defined in Definition III.5 .

For every $\mu \in [1, N]$, consider a UCC $(G, h^{(\mu)})$ with parameters (n, k, l, p). For each μ, the generator matrix G along with the function $h^{(\mu)}$ generates p^{k+l} codewords. Each of these codewords are characterized by a triple (a, i, μ), where $a \in \mathbb{F}_p^k$ and $i \in \mathbb{F}_p^l$ correspond to the coarse code and the coset indices, respectively. Let $W^{n,(\mu)}(a, i)$ denote the codewords associated with the encoder (Alice), generated using the above procedure, where

$$W^{n,(\mu)}(a, i) = aG + h^{(\mu)}(i). \tag{3.11}$$

Now, construct the operators

$$\bar{A}_{w^n}^{(\mu)} \triangleq \alpha_{w^n} \left(\sqrt{\rho^{\otimes n}}^{-1} \tilde{\rho}_{w^n} \sqrt{\rho^{\otimes n}}^{-1} \right), \qquad \alpha_{w^n} \triangleq \frac{1}{(1+\eta)} \frac{p^n \lambda_{w^n}}{p^{k+l}}, \tag{3.12}$$

with $\eta \in (0, 1)$ being a parameter to be determined. Note that, following the definition of $\tilde{\rho}_{w^n}$, we have $\bar{A}_{w^n}^{(\mu)} = 0$ for $w^n \notin \mathcal{T}_\delta^{(n)}(W)$. Having constructed the operators $\bar{A}_{w^n}^{(\mu)}$, we normalize these

operators, so that they constitute a valid sub-POVM. To do so, we define

$$\Sigma^{(\mu)} \triangleq \sum_{w^n} \gamma_{w^n}^{(\mu)} \bar{A}_{w^n}^{(\mu)}, \qquad \gamma_{w^n}^{(\mu)} \triangleq |\{(a,i) : W^{n,(\mu)}(a,i) = w^n\}|.$$

Now, we define Π^μ as the pruning operator for $\Sigma^{(\mu)}$ with respect to Π_ρ using Definition III.2. Note that, the pruning operator Π^μ depends on the pair $(G, h^{(\mu)})$. For ease of analysis, the subspace of Π^μ is restricted to Π_ρ and hence Π^μ is a projector onto a subspace of Π_ρ. Using these pruning operators, for each $\mu \in [1, N]$, construct the sub-POVM $\tilde{M}^{(n,\mu)}$ as

$$\tilde{M}^{(n,\mu)} \triangleq \{\gamma_{w^n}^{(\mu)} A_{w^n}^{(\mu)}\}_{w^n \in \mathcal{W}^n}, \tag{3.13}$$

where $A_{w^n}^{(\mu)} \triangleq \Pi^\mu \bar{A}_{w^n}^{(\mu)} \Pi^\mu$. Further, using Π^μ we have $\sum_{w^n} \gamma_{w^n}^{(\mu)} A_{w^n}^{(\mu)} = \Pi^\mu \Sigma^{(\mu)} \Pi^\mu \leq \Pi_\rho \leq I$, and thus $\tilde{M}^{(n,\mu)}$ is a valid sub-POVM for all $\mu \in [1, N]$. Moreover, the collection $\tilde{M}^{(n,\mu)}$ is completed using the operators $I - \sum_{w^n \in \mathcal{W}^n} \gamma_{w^n}^{(\mu)} A_{w^n}^{(\mu)}$.

3.3.3.2 Binning of POVMs

The next step is to bin the above constructed sub-POVMs. Since, UCC is a union of several cosets, we associate a bin to each coset, and hence place all the codewords of a coset in the same bin. For each $i \in \mathbb{F}_p^l$, let $\mathcal{B}^{(\mu)}(i) \triangleq \mathbb{C}(G, h^{(\mu)}(i))$ denote the ith bin. Further, for all $i \in \mathbb{F}_p^l$, we define

$$\Gamma_i^{A,(\mu)} \triangleq \sum_{w^n \in \mathcal{W}^n} \sum_{a \in \mathbb{F}_p^k} A_{w^n}^{(\mu)} \mathbb{1}_{\{aG + h^{(\mu)}(i) = w^n\}}.$$

Using these operators, we form the collection $M^{(n,\mu)} \triangleq \{\Gamma_i^{A,(\mu)}\}_{i \in \mathbb{F}_p^l}$. Note that if the collection $\tilde{M}^{(n,\mu)}$ is a sub-POVM for each $\mu \in [1, N]$, then so is the collection $M^{(n,\mu)}$, which is due to the relation $\sum_{i \in \mathbb{F}_p^l} \Gamma_i^{A,(\mu)} = \sum_{w^n \in \mathcal{W}^n} \gamma_{w^n}^{(\mu)} A_{w^n}^{(\mu)} \leq I$. To complete $M^{(n,\mu)}$, we define $\Gamma_0^{A,(\mu)}$ as $\Gamma_0^{A,(\mu)} = I - \sum_i \Gamma_i^{A,(\mu)}$ [1]. Now, we intend to use the completions $[M^{(n,\mu)}]$ as the POVM for the encoder.

[1] Note that $\Gamma_0^{A,(\mu)} = I - \sum_i \Gamma_i^{A,(\mu)} = I - \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \gamma_{w^n}^{(\mu)} A_{w^n}^{(\mu)}$.

3.3.3.3 Decoder mapping

We create a decoder which, on receiving the classical bits from the encoder, generates a sequence $W^n \in \mathbb{F}_p^n$ as follows. The decoder first creates a set $D_i^{(\mu)}$ and a function $F^{(\mu)}$ defined as

$$D_i^{(\mu)} \triangleq \left\{\tilde{a} \in \mathbb{F}_p^k : \tilde{a}G + h^{(\mu)}(i) \in \mathcal{T}_\delta^{(n)}(W)\right\} \quad \text{and}$$
$$F^{(\mu)}(i) \triangleq \begin{cases} \tilde{a}G + h^{(\mu)}(i) & \text{if } D_i^{(\mu)} \equiv \{\tilde{a}\} \\ w_0^n & \text{otherwise ,} \end{cases} \tag{3.14}$$

where w_0^n is an arbitrary sequence in $\mathbb{F}_p^n \backslash \mathcal{T}_\delta^{(n)}(W)$. Further, $F^{(\mu)}(i) = w_0^n$ for $i = 0$. Given this and the stochastic processing $P_{Z|W}$, we obtain the approximating sub-POVM $\hat{M}^{(n)}$ with the following operators.

$$\hat{\Lambda}_{z^n} \triangleq \frac{1}{N}\sum_{\mu=1}^{N} \sum_{w^n \in \mathbb{F}_p^n} \sum_{i: F^{(\mu)}(i) = w^n} \Gamma_i^{A,(\mu)} P_{Z|W}^n(z^n|w^n), \ \forall z^n \in \mathcal{Z}^n.$$

The generator matrix G and the function $h^{(\mu)}$ are chosen randomly uniformly and independently.

3.3.3.4 Trace Distance

In what follows, we show that $\hat{M}^{(n)}$ is ϵ-faithful to $M^{\otimes n}$ with respect to $\rho^{\otimes n}$ (according to Definition I.1), where $\epsilon > 0$ can be made arbitrarily small. More precisely, using (3.9), we show that, $\mathbb{E}[K] \leq \epsilon$, where

$$K \triangleq \sum_{z^n} \left\| \sum_{w^n} \sqrt{\rho^{\otimes n}} \bar{\Lambda}_{w^n} \sqrt{\rho^{\otimes n}} P_{Z|W}^n(z^n|w^n) - \sqrt{\rho^{\otimes n}} \hat{\Lambda}_{z^n} \sqrt{\rho^{\otimes n}} \right\|_1, \tag{3.15}$$

where the expectation is with respect to the codebook generation.

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

Consider the second term within K, which can be written as

$$\sqrt{\rho^{\otimes n}} \hat{\Lambda}_{z^n} \sqrt{\rho^{\otimes n}} = \frac{1}{N} \sum_{\mu} \sum_{i} \sqrt{\rho^{\otimes n}} \Gamma_i^{A,(\mu)} \sqrt{\rho^{\otimes n}} P_{Z|W}^n(z^n|F^{(\mu)}(i)) \underbrace{\sum_{w^n} \mathbb{1}_{\{F^{(\mu)}(i) = w^n\}}}_{=1} = T + \widetilde{T},$$

where

$$T \triangleq \frac{1}{N}\sum_{\mu}\sum_{i>0}\sqrt{\rho^{\otimes n}}\Gamma_i^{A,(\mu)}\sqrt{\rho^{\otimes n}}P_{Z|W}^n(z^n|F^{(\mu)}(i)), \quad \widetilde{T} \triangleq \frac{1}{N}\sum_{\mu}\sqrt{\rho^{\otimes n}}\Gamma_0^{A,(\mu)}\sqrt{\rho^{\otimes n}}P_{Z|W}^n(z^n|w_0^n).$$

Hence, we have $K \leq S + \widetilde{S}$, where

$$S \triangleq \sum_{z^n}\left\|\sum_{w^n}\sqrt{\rho^{\otimes n}}\bar{\Lambda}_{w^n}\sqrt{\rho^{\otimes n}}P_{Z|W}^n(z^n|w^n) - T\right\|_1, \tag{3.16}$$

and $\widetilde{S} \triangleq \sum_{z^n}\|\widetilde{T}\|_1$. Note that $\widetilde{S}$ captures the error induced by not covering the state $\rho^{\otimes n}$. We further bound $\widetilde{S}$ as

$$\widetilde{S} \leq \frac{1}{N}\sum_{\mu}\sum_{z^n}P_{Z|W}^n(z^n|w_0^n)\left\|\sqrt{\rho^{\otimes n}}\Gamma_0^{A,(\mu)}\sqrt{\rho^{\otimes n}}\right\|_1 \leq \frac{1}{N}\sum_{\mu}\left\|\sqrt{\rho^{\otimes n}}(I - \sum_{w^n}\gamma_{w^n}^{(\mu)}A_{w^n}^{(\mu)})\sqrt{\rho^{\otimes n}}\right\|_1,$$

which gives $\widetilde{S} \leq \widetilde{S}_1 + \widetilde{S}_2$ where

$$\widetilde{S}_1 \triangleq \frac{1}{N}\sum_{\mu}\left\|\sum_{w^n}\lambda_{w^n}\hat{\rho}_{w^n} - \sum_{w^n}\sqrt{\rho^{\otimes n}}\gamma_{w^n}^{(\mu)}\bar{A}_{w^n}^{(\mu)}\sqrt{\rho^{\otimes n}}\right\|_1,$$
$$\widetilde{S}_2 \triangleq \frac{1}{N}\sum_{\mu}\sum_{w^n}\left\|\sqrt{\rho^{\otimes n}}\gamma_{w^n}^{(\mu)}\left(\bar{A}_{w^n}^{(\mu)} - A_{w^n}^{(\mu)}\right)\sqrt{\rho^{\otimes n}}\right\|_1.$$

To provide a bound for the term $\widetilde{S}_1$, we (i) develop a n-letter version of Lemma III.12 and (ii) provide a proposition bounding the term corresponding to $\widetilde{S}_1$, using this n-letter lemma.

Lemma III.13. *Let $\{\lambda_w, \theta_w\}_{w\in\mathcal{W}}$ be an ensemble, with $\theta_w \in \mathcal{D}(\mathcal{H})$ for all $w \in \mathcal{W}$, $\mathcal{W} \subseteq \mathbb{F}_p$ for some finite prime p. Then, for any $\epsilon_c \in (0,1)$, and for any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have*

$$\mathbb{E}\left[\left\|\sum_{w^n}\lambda_{w^n}\theta_{w^n} - \frac{p^n}{p^{k+l}}\frac{1}{N'}\sum_{\mu=1}^{N'}\sum_{w^n}\sum_{a,m}\frac{\lambda_{w^n}}{(1+\eta)}\theta_{w^n}\mathbb{1}_{\{W^{n,(\mu)}(a,m)=w^n\}}\right\|_1\right] \leq \epsilon_c, \tag{3.17}$$

if $\left(\frac{k+l}{n}\right)\log p + \frac{1}{n}\log N' > I(W;R)_{\sigma_\theta} - S(W)_{\sigma_\theta} + \log p$, where $\theta_{w^n} \triangleq \bigotimes_{i=1}^n \theta_{w_i}$ and $\lambda_{w^n} \triangleq \Pi_{i=1}^n\lambda_{w_i}$, $\sigma_\theta^{RW} \triangleq \sum_{w\in\mathcal{W}}\lambda_w\theta_w \otimes |w\rangle\langle w|$, for some orthogonal set $\{|w\rangle\}_{w\in\mathcal{W}}$, and $\{W^{n,(\mu)}(a,m) : a \in \mathbb{F}_p^k, m \in \mathbb{F}_p^l, \mu \in [2^{nC}]\}$ are as defined in (3.11), with G and $h^{(\mu)}$ generated randomly uniformly and independently.

Proof. The proof of the lemma is provided in Appendix B.2 □

Now we provide the following proposition.

Proposition III.14. *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have we have $\mathbb{E}[\widetilde{S}_1] \leq \epsilon$, if $\frac{k+l}{n} \log p > I(W;R)_\sigma - S(W)_\sigma + \log p$, where σ is the auxiliary state defined in the theorem.*

Proof. The proof is provided in Appendix B.3. □

Now we provide a bound for $\widetilde{S}_2$. For that, we first develop another n-letter lemma as follows.

Lemma III.15. *For $\gamma_{w^n}^{(\mu)}, \bar{A}_{w^n}^{(\mu)}$,and $A_{w^n}^{(\mu)}$ as defined above, we have*

$$\sum_{w^n} \gamma_{w^n}^{(\mu)} \left\| \sqrt{\rho^{\otimes n}} \left(\bar{A}_{w^n}^{(\mu)} - A_{w^n}^{(\mu)} \right) \sqrt{\rho^{\otimes n}} \right\|_1 \leq 2\, 2^{3n\delta_\rho} \left(H_0 + \frac{\sqrt{(1-\varepsilon)}}{(1+\eta)} \sqrt{H_1 + H_2 + H_3} \right),$$

where

$$\begin{aligned} H_0 &\triangleq \left| \Delta^{(\mu)} - \mathbb{E}[\Delta^{(\mu)}] \right|, \quad H_1 \triangleq \mathrm{Tr}\left\{ (\Pi_\rho - \Pi^\mu) \sum_{w^n} \lambda_{w^n} \tilde{\rho}_{w^n} \right\}, \\ H_2 &\triangleq \left\| \sum_{w^n} \lambda_{w^n} \tilde{\rho}_{w^n} - (1-\varepsilon) \sum_{w^n} \frac{\alpha_{w^n} \gamma_{w^n}^{(\mu)}}{\mathbb{E}[\Delta^{(\mu)}]} \tilde{\rho}_{w^n} \right\|_1, \\ H_3 &\triangleq (1-\varepsilon) \left\| \sum_{w^n} \frac{\alpha_{w^n} \gamma_{w^n}^{(\mu)}}{\Delta^{(\mu)}} \tilde{\rho}_{w^n} - \sum_{w^n} \frac{\alpha_{w^n} \gamma_{w^n}^{(\mu)}}{\mathbb{E}[\Delta^{(\mu)}]} \tilde{\rho}_{w^n} \right\|_1, \end{aligned} \tag{3.18}$$

$\Delta^{(\mu)} = \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \alpha_{w^n} \gamma_{w^n}^{(\mu)}, \varepsilon \triangleq \sum_{w^n \notin \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n}$ *and* $\delta_\rho(\delta) \searrow 0$ *as* $\delta \searrow 0$.

Proof. The proof is provided in Appendix B.4 □

Using the above lemma on $\widetilde{S}_2$ gives

$$\widetilde{S}_2 \leq \frac{2}{N} \sum_{\mu=1}^{N} 2^{3n\delta_\rho} \left(H_0 + \frac{\sqrt{(1-\varepsilon)}}{(1+\eta)} \sqrt{H_1 + H_2 + H_3} \right).$$

Let us first consider H_1. By observing $\sum_{w^n} \lambda_{w^n} \tilde{\rho}_{w^n} \leq \Pi_\rho \rho^{\otimes n} \Pi_\rho \leq 2^{-n(S(\rho)-\delta_\rho)} \Pi_\rho$, we bound H_1 as $H_1 \leq 2^{-n(S(\rho)-\delta_\rho)} \mathrm{Tr}\{(\Pi_\rho - \Pi^\mu)\}$. Note that

$$\mathbb{E}[\Sigma^{(\mu)}] = \mathbb{E}\left[\sum_{w^n} \alpha_{w^n} \gamma_{w^n}^{(\mu)} \sqrt{\rho^{\otimes n}}^{-1} \tilde{\rho}_{w^n} \sqrt{\rho^{\otimes n}}^{-1} \right] = \frac{1}{(1+\eta)} \sum_{w^n} \lambda_{w^n} \sqrt{\rho^{\otimes n}}^{-1} \tilde{\rho}_{w^n} \sqrt{\rho^{\otimes n}}^{-1} \leq \frac{\Pi_\rho}{(1+\eta)}.$$

Now, we use the Pruning Trace Inequality developed in Lemma III.3 on $\Sigma^{(\mu)}$, with $\eta \in (0,1)$ to obtain

$$\begin{aligned}
\mathbb{E}[H_1] &\leq 2^{-n(S(\rho)-\delta_\rho)} \frac{(1+\eta)}{\eta} \mathbb{E}\left[\|\Sigma^{(\mu)} - \mathbb{E}[\Sigma^{(\mu)}]\|_1\right] \\
&\leq 2^{-n(S(\rho)-\delta_\rho)} \frac{(1+\eta)}{\eta} \left\|\Pi_\rho \sqrt{\rho^{\otimes n}}^{-1}\right\|_\infty \mathbb{E}\left[\left\|\sum_{w^n} \alpha_{w^n} \gamma_{w^n}^{(\mu)} \tilde{\rho}_{w^n} - \mathbb{E}\big[\sum_{w^n} \alpha_{w^n} \gamma_{w^n}^{(\mu)} \tilde{\rho}_{w^n}\big]\right\|_1\right] \left\|\Pi_\rho \sqrt{\rho^{\otimes n}}^{-1}\right\|_\infty \\
&\leq 2^{2n\delta_\rho} \frac{(1+\eta)}{\eta} \mathbb{E}\left[\left\|\sum_{w^n} \frac{\lambda_{w^n} \tilde{\rho}_{w^n}}{(1+\eta)} - \frac{1}{(1+\eta)} \frac{p^n}{p^{k+l}} \sum_{w^n} \sum_{a,i} \lambda_{w^n} \tilde{\rho}_{w^n} \mathbb{1}_{\{W^{n,(\mu)}(a,i)=w^n\}}\right\|_1\right] \\
&= 2^{2n\delta_\rho} \frac{(1-\varepsilon)}{\eta} \mathbb{E}[\widetilde{H}], \qquad (3.19)
\end{aligned}$$

where the second inequality follows from Hólders inequality, and the equality follows by defining $\widetilde{H}$ as

$$\widetilde{H} \triangleq \left\|\sum_{w^n} \frac{\lambda_{w^n}}{(1-\varepsilon)} \tilde{\rho}_{w^n} - \frac{p^n}{p^{k+l}} \sum_{w^n} \sum_{a,i} \frac{\lambda_{w^n}}{(1-\varepsilon)} \tilde{\rho}_{w^n} \mathbb{1}_{\{W^{n,(\mu)}(a,i)=w^n\}}\right\|_1 . \qquad (3.20)$$

Similarly, using $\mathbb{E}[\Delta^{(\mu)}] = \frac{(1-\varepsilon)}{(1+\eta)}$, H_2 can be simplified as

$$H_2 = \left\|\sum_{w^n} \lambda_{w^n} \tilde{\rho}_{w^n} - \frac{p^n}{p^{k+l}} \sum_{w^n} \sum_{a,i} \lambda_{w^n} \tilde{\rho}_{w^n} \mathbb{1}_{\{W^{n,(\mu)}(a,i)=w^n\}}\right\|_1 = (1-\varepsilon)\tilde{H}. \qquad (3.21)$$

Now we consider H_3 and convert it into a similar expression as H_0.

$$H_3 \leq (1-\varepsilon) \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \alpha_{w^n} \gamma_{w^n}^{(\mu)} \left|\frac{1}{\Delta^{(\mu)}} - \frac{1}{\mathbb{E}[\Delta^{(\mu)}]}\right| = (1+\eta)\left|\Delta^{(\mu)} - \mathbb{E}[\Delta^{(\mu)}]\right| = (1+\eta) H_0. \qquad (3.22)$$

Using the above simplification and the concavity of square-root function we obtain:

$$\begin{aligned}
\mathbb{E}[\widetilde{S}_2] &\leq \frac{2}{N} 2^{3n\delta_\rho} \sum_{\mu=1}^{N} \left(\mathbb{E}[H_0] + \frac{\sqrt{(1-\varepsilon)}}{(1+\eta)} \sqrt{(1-\varepsilon)\left(\frac{2^{2n\delta_\rho}}{\eta} + 1\right)\mathbb{E}[\widetilde{H}] + (1+\eta)\mathbb{E}[H_0]}\right) \\
&\leq \frac{2}{N} 2^{3n\delta_\rho} \sum_{\mu=1}^{N} \left(\mathbb{E}[H_0] + \frac{(1-\varepsilon)}{(1+\eta)} \sqrt{\left(\frac{2^{2n\delta_\rho}}{\eta} + 1\right)\mathbb{E}[\widetilde{H}]} + \sqrt{\frac{(1-\varepsilon)}{(1+\eta)}} \sqrt{\mathbb{E}[H_0]}\right).
\end{aligned}$$

The following proposition provides a bound on the above term.

Proposition III.16. *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}\left[\widetilde{S}_2\right] \leq \epsilon$, if $\frac{k+l}{n}\log p > I(W;R)_\sigma - S(W)_\sigma + \log p$, where σ is the auxiliary state defined in the theorem.*

Proof. The proof is provided in Appendix B.5 □

Remark III.17. The term corresponding to the operators that complete the sub-POVMs $M^{(n,\mu)}$, i.e., $I - \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \gamma_{w^n}^{(\mu)} A_{w^n}^{(\mu)}$ is taken care in $\widetilde{T}$. The expression T excludes these completing operators.

Step 2: Isolating the effect of error induced by binning

For this, we simplify T as

$$T = \frac{1}{N}\sum_{\mu}\sum_{w^n}\sum_{i>0}\sum_{a\in\mathbb{F}_p^k} \sqrt{\rho^{\otimes n}} A_{w^n}^{(\mu)} \sqrt{\rho^{\otimes n}} P_{Z|W}^n(z^n|F^{(\mu)}(i)) \mathbb{1}_{\{aG+h^{(\mu)}(i)=w^n\}}.$$

We substitute the above expression into S defined in (3.16), and isolate the effect of binning by adding and subtracting an appropriate term within S and applying triangle inequality to obtain $S \leq S_1 + S_2$, where

$$S_1 \triangleq \sum_{z^n} \left\| \sum_{w^n} \sqrt{\rho^{\otimes n}} \left(\bar{\Lambda}_{w^n} - \frac{1}{N}\sum_{\mu} \gamma_{w^n}^{(\mu)} A_{w^n}^{(\mu)} \right) \sqrt{\rho^{\otimes n}} P_{Z|W}^n(z^n|w^n) \right\|_1,$$

$$S_2 \triangleq \sum_{z^n} \left\| \frac{1}{N}\sum_{\mu}\sum_{a,i>0}\sum_{w^n} \sqrt{\rho^{\otimes n}} A_{w^n}^{(\mu)} \sqrt{\rho^{\otimes n}} \mathbb{1}_{\{aG+h^{(\mu)}(i)=w^n\}} \left(P_{Z|W}^n(z^n|w^n) - P_{Z|W}^n\left(z^n|F^{(\mu)}(i)\right)\right) \right\|_1,$$

where $F^{(\mu)}(\cdot)$ is as defined in (3.14). Note that the term S_1 characterizes the error introduced by approximation of the original POVM with the collection of approximating sub-POVM $\tilde{M}^{(n,\mu)}$, and the term S_2 characterizes the error caused by binning this approximating sub-POVM. In this step, we analyze S_2 and prove the following proposition.

Proposition III.18. *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[S_2] \leq \epsilon$, if $\frac{k+l}{n}\log p - R < \log p - S(W)_\sigma$, where σ is the auxiliary state defined in the statement of the theorem.*

Proof. The proof is provided in Appendix B.6 □

Step 3: Isolating the effect of approximating measurement

In this step, we finally analyze the error induced from employing the approximating measurement, given by the term S_1. We add and subtract appropriate terms within S_1 and use triangle inequality to obtain $S_1 \leq S_{11} + S_{12} + S_{13}$, where

$$S_{11} \triangleq \sum_{z^n} \left\| \sum_{w^n} \sqrt{\rho^{\otimes n}} \left(\bar{\Lambda}_{w^n} - \frac{1}{N} \sum_{\mu=1}^{N} \frac{\alpha_{w^n} \gamma_{w^n}^{(\mu)}}{\lambda_{w^n}} \bar{\Lambda}_{w^n} \right) \sqrt{\rho^{\otimes n}} P_{Z|W}^n(z^n|w^n) \right\|_1 ,$$

$$S_{12} \triangleq \sum_{z^n} \left\| \frac{1}{N} \sum_{\mu=1}^{N} \sum_{w^n} \sqrt{\rho^{\otimes n}} \left(\frac{\alpha_{w^n} \gamma_{w^n}^{(\mu)}}{\lambda_{w^n}} \bar{\Lambda}_{w^n} - \gamma_{w^n}^{(\mu)} \bar{A}_{w^n}^{(\mu)} \right) \sqrt{\rho^{\otimes n}} P_{Z|W}^n(z^n|w^n) \right\|_1 ,$$

$$S_{13} \triangleq \sum_{z^n} \left\| \frac{1}{N} \sum_{\mu=1}^{N} \sum_{w^n} \sqrt{\rho^{\otimes n}} \left(\gamma_{w^n}^{(\mu)} \bar{A}_{w^n}^{(\mu)} - \gamma_{w^n}^{(\mu)} A_{w^n}^{(\mu)} \right) \sqrt{\rho^{\otimes n}} P_{Z|W}^n(z^n|w^n) \right\|_1 .$$

Now with the intention of employing Lemma III.13, we express S_{11} as

$$S_{11} = \left\| \sum_{w^n} \lambda_{w^n} \hat{\rho}_{w^n} \otimes \phi_{w^n} - \frac{1}{N} \frac{1}{(1+\eta)} \frac{p^n}{p^{k+l}} \sum_{\mu} \sum_{w^n} \sum_{a,i \neq 0} \mathbb{1}_{\{W^{n,(\mu)}(a,i) = w^n\}} \hat{\rho}_{w^n} \otimes \phi_{w^n} \right\|_1 ,$$

where the equality above is obtained by defining $\phi_{w^n} = \sum_{z^n} P_{Z|W}^n(z^n|w^n) \otimes |z^n\rangle\langle z^n|$ and using the definitions of $\alpha_{w^n}, \gamma_{w^n}^{(\mu)}$ and $\hat{\rho}_{w^n}$, followed by using the triangle inequality for the block diagonal operators, Note that the triangle inequality becomes an equality for such block diagonal operators. By identifying θ_w with $\hat{\rho}_w \otimes \phi_w$ in Lemma III.13 we obtain the following: for all $\epsilon > 0$ and $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, $\mathbb{E}[S_{11}] \leq \epsilon$, if $\frac{k+l}{n} \log p + \frac{1}{n} \log N > I(W;R,Z)_\sigma + \log p - S(W)_\sigma$, where σ is the auxiliary state defined in the theorem.

Now we consider the term corresponding to S_{12}, and prove that its expectation is small. Recalling S_{12}, we get

$$\begin{aligned} S_{12} &\leq \frac{1}{N} \sum_{\mu=1}^{N} \sum_{w^n} \sum_{z^n} P_{Z|W}^n(z^n|w^n) \left\| \sqrt{\rho^{\otimes n}} \left(\frac{\alpha_{w^n} \gamma_{w^n}^{(\mu)}}{\lambda_{w^n}} \bar{\Lambda}_{w^n} - \gamma_{w^n}^{(\mu)} \bar{A}_{w^n}^{(\mu)} \right) \sqrt{\rho^{\otimes n}} \right\|_1 , \\ &= \frac{1}{N} \sum_{\mu=1}^{N} \sum_{w^n} \alpha_{w^n} \gamma_{w^n}^{(\mu)} \left\| \sqrt{\rho^{\otimes n}} \left(\frac{1}{\lambda_{w^n}} \bar{\Lambda}_{w^n} - \sqrt{\rho^{\otimes n}}^{-1} \tilde{\rho}_{w^n} \sqrt{\rho^{\otimes n}}^{-1} \right) \sqrt{\rho^{\otimes n}} \right\|_1 , \end{aligned}$$

where the inequality above is obtained by using triangle inequality. Applying the expectation, we

get

$$\begin{aligned}
\mathbb{E}[S_{12}] &\leq \frac{1}{(1+\eta)} \sum_{w^n} \lambda_{w^n} \left\| \sqrt{\rho^{\otimes n}} \left(\frac{1}{\lambda_{w^n}} \bar{\Lambda}_{w^n} - \sqrt{\rho^{\otimes n}}^{-1} \tilde{\rho}_{w^n} \sqrt{\rho^{\otimes n}}^{-1} \right) \sqrt{\rho^{\otimes n}} \right\|_1, \\
&\leq \frac{1}{(1+\eta)} \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \left\| (\hat{\rho}_{w^n} - \tilde{\rho}_{w^n}) \right\|_1 + \frac{1}{(1+\eta)} \sum_{w^n \notin \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \left\| \hat{\rho}_{w^n} \right\|_1 \\
&\leq \frac{(2\sqrt{\varepsilon'} + 2\sqrt{\varepsilon''}) + \varepsilon}{(1+\eta)} = \epsilon_{S_{12}},
\end{aligned}$$

where we have used the fact that $\mathbb{E}[\alpha_{w^n} \gamma_{w^n}^{(\mu)}] = \frac{\lambda_{w^n}}{(1+\eta)}$, and the last inequality is obtained by the repeated usage of the Average Gentle Measurement Lemma *Wilde* (2013a) and setting $\epsilon_{S_{12}} = \frac{1}{(1+\eta)}(2\sqrt{\varepsilon'} + 2\sqrt{\varepsilon''} + \varepsilon)$ with $\epsilon_{S_{12}} \searrow 0$ as $n \to \infty$ and $\varepsilon' \triangleq \varepsilon'_p + 2\sqrt{\varepsilon'_p}$ and $\varepsilon'' \triangleq 2\varepsilon'_p + 2\sqrt{\varepsilon'_p}$ for $\varepsilon'_p \triangleq 1 - \min\{\mathrm{Tr}\{\Pi_\rho \hat{\rho}_{w^n}\}, \mathrm{Tr}\{\Pi_{w^n} \hat{\rho}_{w^n}\}, 1-\varepsilon\}$ (see (35) in *Wilde et al.* (2012) for details). Now, we move on to bounding the last term within S_1, i.e., S_{13}. We start by applying triangle inequality to obtain

$$\begin{aligned}
S_{13} &\leq \sum_{z^n} \sum_{w^n} P_{Z|W}^n(z^n|w^n) \left\| \frac{1}{N} \sum_{\mu=1}^{N} \sqrt{\rho^{\otimes n}} \left(\gamma_{w^n}^{(\mu)} \bar{A}_{w^n}^{(\mu)} - \gamma_{w^n}^{(\mu)} A_{w^n}^{(\mu)} \right) \sqrt{\rho^{\otimes n}} \right\|_1 \\
&\leq \frac{1}{N} \sum_{\mu=1}^{N} \sum_{w^n} \gamma_{w^n}^{(\mu)} \left\| \sqrt{\rho^{\otimes n}} \left(\bar{A}_{w^n}^{(\mu)} - A_{w^n}^{(\mu)} \right) \sqrt{\rho^{\otimes n}} \right\|_1 = \widetilde{S}_2. \qquad (3.23)
\end{aligned}$$

Since the above term is exactly same as $\widetilde{S}_2$, we obtain the same rate constraints as in $\widetilde{S}_2$ to bound S_{13}, i.e., for all $\epsilon > 0$ and $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, $\mathbb{E}[S_{13}] \leq \epsilon$ if $\frac{k+l}{n} \log p > I(W;R)_\sigma + \log p - S(W)_\sigma$.

Since $S_1 \leq S_{11} + S_{12} + S_{13}$, S_1 can be made arbitrarily small for sufficiently large n, if $\frac{k+l}{n} \log p + \frac{1}{n} \log N > I(W;RZ)_\sigma - S(W)_\sigma + \log p$ and $\frac{k+l}{n} \log p > I(W;R)_\sigma - S(W)_\sigma + \log p$.

3.3.3.5 Rate Constraints

To sum-up, we showed $\mathbb{E}[K] \leq \epsilon$ holds for sufficiently large n if the inequalities in (3.5) provided in the statement of the theorem are satisfied, where $R_1 \triangleq \frac{k}{n} \log p$ and $C \triangleq \frac{1}{n} \log_2 N$, and $R = \frac{l}{n} \log p$. Therefore, there exists a distributed protocol with parameters $(n, 2^{nR}, 2^{nC})$ such that its overall POVM $\hat{M}$ is ϵ-faithful to $M^{\otimes n}$ with respect to $\rho^{\otimes n}$. This completes the proof of the theorem.

3.4 Distributed Measurement Compression using Structured Random POVMs

Let ρ_{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 Charlie, that are trying to collectively simulate the action of a given measurement M_{AB} performed on the state ρ_{AB}, as shown in Fig. 3.1. Charlie additionally has access to unlimited private randomness. The problem is defined in the following.

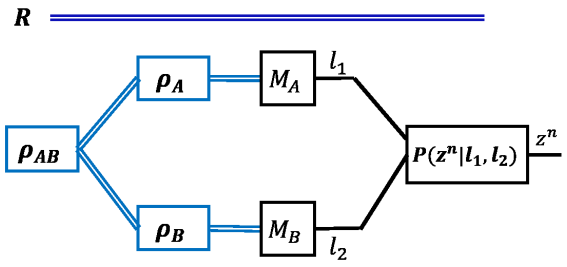

Figure 3.1: The diagram depicting the distributed POVM simulation problem with stochastic processing. In this setting, Charlie additionally has access to unlimited private randomness.

3.4.1 Problem Formulation

Definition III.19. 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 $(n, \Theta_1, \Theta_2, N_1, N_2)$ is characterized by

1) a collection of Alice's sub-POVMs $\tilde{M}_A^{(\mu_1)}, \mu_1 \in [1, N_1]$ 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 collection of Bob's sub-POVMs $\tilde{M}_B^{(\mu_2)}, \mu_2 \in [1, N_2]$ 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 Charlie's classical stochastic maps $P^{(\mu_1,\mu_2)}(z^n|l_1, l_2)$ for all $l_1 \in \mathcal{L}_1, l_2 \in \mathcal{L}_2, z^n \in \mathcal{Z}^n$, $\mu_1 \in [1, N_1]$ and $\mu_2 \in [1, N_2]$.

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

$$\tilde{\Lambda}_{z^n} \triangleq \frac{1}{N_1}\frac{1}{N_2}\sum_{\mu_1,\mu_2}\sum_{l_1,l_2} P^{(\mu_1,\mu_2)}(z^n|l_1,l_2)\Lambda_{l_1}^{A,(\mu_1)} \otimes \Lambda_{l_2}^{B,(\mu_2)}, \quad \forall z^n \in \mathcal{Z}^n,$$

where $\Lambda_{l_1}^{A,(\mu_1)}$ and $\Lambda_{l_2}^{B,(\mu_2)}$ are the operators corresponding to the sub-POVMs $\tilde{M}_A^{(\mu_1)}$ and $\tilde{M}_B^{(\mu_2)}$, respectively.

In the above definition, (Θ_1, Θ_2) determines the amount of classical bits communicated from Alice and Bob to Charlie. The amount of pairwise shared randomness is determined by N_1 and N_2. The classical stochastic maps $P^{(\mu_1,\mu_2)}(z^n|l_1, l_2)$ represent the action of Charlie on the received classical bits.

Definition III.20. 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 quadruple (R_1, R_2, C_1, C_2) 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 $(n, \Theta_1, \Theta_2, N_1, N_2)$ such that its overall sub-POVM $\tilde{M}_{AB}$ is ϵ-faithful to $M_{AB}^{\otimes n}$ with respect to $\rho_{AB}^{\otimes n}$ (see Definition I.1), and

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

The set of all achievable quadruples (R_1, R_2, C_1, C_2) is called the achievable rate region.

Definition III.21 (Joint Measurements)**.** A POVM $M_{AB} = \{\Lambda_z^{AB}\}_{z\in\mathcal{Z}}$, acting on a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, is said to have a separable decomposition with stochastic integration given by $(\bar{M}_A, \bar{M}_B, P_{Z|S,T})$ if there exist POVMs $\bar{M}_A = \{\bar{\Lambda}_s^A\}_{s\in\mathcal{S}}$ and $\bar{M}_B = \{\bar{\Lambda}_t^B\}_{t\in\mathcal{T}}$ and a stochastic mapping $P_{Z|S,T} : \mathcal{S} \times \mathcal{T} \to \mathcal{Z}$ such that

$$\Lambda_z^{AB} = \sum_{s,t} P_{Z|S,T}(z|s,t)\bar{\Lambda}_s^A \otimes \bar{\Lambda}_t^B, \quad \forall z \in \mathcal{Z},$$

where $\mathcal{S}, \mathcal{T}$, and $\mathcal{Z}$ are finite sets.

3.4.2 An Inner Bound

The following theorem provides an inner bound to the achievable rate region, which is proved in Section 3.5. This is one of the main results of this paper.

Theorem III.22. *Consider 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 III.21), yielding POVMs $\bar{M}_A = \{\bar{\Lambda}_s^A\}_{s\in\mathcal{S}}$ *and* $\bar{M}_B = \{\bar{\Lambda}_t^B\}_{t\in\mathcal{T}}$ *and a stochastic map* $P_{Z|S,T} : \mathcal{S}\times\mathcal{T}\to\mathcal{Z}$. *Define the auxiliary states* $\sigma_1^{RSB} \triangleq (\mathrm{id}_R\otimes\bar{M}_A\otimes\mathrm{id}_B)(\Psi_{RAB}^{\rho_{AB}})$, $\sigma_2^{RTV} \triangleq (\mathrm{id}_R\otimes\mathrm{id}_A\otimes\bar{M}_B)(\Psi_{RAB}^{\rho_{AB}})$, *and* $\sigma_3^{RSTZ} \triangleq \sum_{s,t,z}\sqrt{\rho_{AB}}\left(\bar{\Lambda}_s^A\otimes\bar{\Lambda}_t^B\right)\sqrt{\rho_{AB}}\otimes P_{Z|S,T}(z|s,t)\,|s\rangle\langle s|\otimes|t\rangle\langle t|\otimes|z\rangle\langle z|$, *for some orthonormal sets* $\{|s\rangle\}_{s\in\mathcal{S}}, \{|t\rangle\}_{t\in\mathcal{T}}$, *and* $\{|z\rangle\}_{z\in\mathcal{Z}}$, *where* $\Psi_{RAB}^{\rho_{AB}}$ *is a purification of* ρ_{AB}. *A quadruple* (R_1, R_2, C_1, C_2) *is achievable if there exists, a pair of finite sets* $\mathcal{U}$ *and* $\mathcal{V}$, *a pair of mappings* $f_S : \mathcal{S}\to\mathcal{U}$ *and* $f_T : \mathcal{T}\to\mathcal{V}$, *and a stochastic mapping* $P_{Z|W} : \mathcal{W}\to\mathcal{Z}$ *yielding* $U = f_S(S)$, $V = f_T(T)$, *and either* $W = U + V$, *for some finite field* $\mathbb{F}_p$ *and a prime* p, *such that*

$$P_{Z|S,T}(z|s,t) = P_{Z|W}(z|f_S(s) + f_T(t)), \quad \forall s\in\mathcal{S}, t\in\mathcal{T}, z\in\mathcal{Z},$$

or $W = (U, V)$ *such that*

$$P_{Z|S,T}(z|s,t) = P_{Z|W}(z|f_S(s), f_T(t)), \quad \forall s\in\mathcal{S}, t\in\mathcal{T}, z\in\mathcal{Z},$$

and the following inequalities are satisfied:

$$\begin{aligned}
R_1 &\geq I(U;R,B)_{\sigma_1} + \bar{I}(W;V)_{\sigma_3} - I(U;V)_{\sigma_3},\\
R_2 &\geq I(V;R,A)_{\sigma_2} + \bar{I}(W;U)_{\sigma_3} - I(U;V)_{\sigma_3},\\
R_1 + R_2 &\geq I(U;R,B)_{\sigma_1} + I(V;R,A)_{\sigma_2} - I(U;V)_{\sigma_3} + \bar{I}(W;V)_{\sigma_3} + \bar{I}(W;U)_{\sigma_3} - \bar{I}(U;V)_{\sigma_3},\\
R_1 + C_1 &\geq I(U;R,Z)_{\sigma_3} + \bar{I}(W;V)_{\sigma_3} - I(U;V)_{\sigma_3},\\
R_2 + C_2 &\geq I(V;R,Z)_{\sigma_3} + \bar{I}(W;U)_{\sigma_3} - I(U;V)_{\sigma_3},\\
R_1 + R_2 + C_1 &\geq I(U;R,Z)_{\sigma_3} + I(V;R,A)_{\sigma_2} - I(U;V)_{\sigma_3} + \bar{I}(W;U)_{\sigma_3} + \bar{I}(W;V)_{\sigma_3} - \bar{I}(U;V)_{\sigma_3},\\
R_1 + R_2 + C_2 &\geq I(V;R,Z)_{\sigma_3} + I(U;R,B)_{\sigma_1} - I(U;V)_{\sigma_3} + \bar{I}(W;U)_{\sigma_3} + \bar{I}(W;V)_{\sigma_3} - \bar{I}(U;V)_{\sigma_3},\\
R_1 + R_2 + C_1 + C_2 &\geq I(U,V;R,Z)_{\sigma_3} + \bar{I}(W;U)_{\sigma_3} + \bar{I}(W;V)_{\sigma_3} - \bar{I}(U;V)_{\sigma_3}. \qquad (3.24)
\end{aligned}$$

where $\bar{I}(W;U)_{\sigma_3} = I(W;U)_{\sigma_3}, \bar{I}(W;V)_{\sigma_3} = I(W;V)_{\sigma_3}$, *and* $\bar{I}(U;V)_{\sigma_3} = I(U;V)_{\sigma_3}$ *if* $W = U + V$ *and* $\bar{I}(W;U)_{\sigma_3} = \bar{I}(W;V)_{\sigma_3} = \bar{I}(U;V)_{\sigma_3} = 0$ *if* $W = (U, V)$.

Proof. Observe that the theorem involves two different cases of W, one being equal to the sum $U + V$, and another being the pair (U, V). We provide a complete proof for the former case in Section 3.5. The proof for the latter follows from the proof of Theorem II.33 provided in Chapter

II. □

Remark III.23. Note that the rate-region obtained in Theorem II.33 using unstructured random code ensembles, contains the constraint $R_1 + R_2 + C_1 + C_2 \geq I(U,V;R,Z)_{\sigma_3}$. Hence when

$$I(W;U)_{\sigma_3} + I(W;V)_{\sigma_3} - I(U;V)_{\sigma_3} = 2S(U+V)_{\sigma_3} - S(U,V)_{\sigma_3} < 0,$$

the above theorem gives a lower sum rate constraint. As a result, the rate-region above contains points that are not contained within the rate-region provided by Theorem II.33. To illustrate this fact further, consider the following example.

Remark III.24. In the above theorem, we restrict our attention to prime finite fields for ease of exposition. The results can be generalized to arbitrary finite fields in a straight-forward manner.

Example III.25. Suppose the composite state ρ_{AB} is described using one of the Bell states on $\mathcal{H}_A \otimes \mathcal{H}_B$ as

$$\rho_{AB} = \frac{1}{2}\left(|00\rangle_{AB} + |11\rangle_{AB}\right)\left(\langle 00|_{AB} + \langle 11|_{AB}\right).$$

Since $\pi^A = \mathrm{Tr}_B\, \rho^{AB}$ and $\pi^B = \mathrm{Tr}_A\, \rho^{AB}$, 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 $\bar{M}_A$ and $\bar{M}_B$, given by $\bar{M}_A \triangleq \left\{\bar{\Lambda}_s^A\right\}_{s\in\mathcal{S}}, \bar{M}_B \triangleq \left\{\bar{\Lambda}_v^B\right\}_{t\in\mathcal{T}}$, where $\mathcal{S} = \mathcal{T} = \{0,1\}$, and

$$\Lambda_0^A = \Lambda_0^B \triangleq \begin{bmatrix} 0.9501 & 0.0826 + i0.1089 \\ 0.0826 - i0.1089 & 0.0615 \end{bmatrix}, \quad \Lambda_1^A = \Lambda_1^B \triangleq \begin{bmatrix} 0.0499 & -0.0826 - i0.1089 \\ -0.0826 + i0.1089 & 0.9385 \end{bmatrix}.$$

Alice and Bob together with Charlie are trying to simulate the action of $M_{AB} \triangleq \left\{\Gamma_z^{AB}\right\}_{z\in\mathcal{Z}}$, using classical communication and common randomness as the resources available, where $\mathcal{Z} = \{0,1\}$, and

$$\Gamma_z^{AB} \triangleq \sum_{s\in\{0,1\}} \sum_{t\in\{0,1\}} P_{Z|S,T}(z|s,t)\left(\Lambda_s^A \otimes \Lambda_t^B\right), \tag{3.25}$$

for $z \in \{0,1\}$, and $P_{Z|S,T}(0|0,0) = P_{Z|S,T}(0|1,1) = 1 - P_{Z|S,T}(0|0,1) = 1 - P_{Z|S,T}(0|1,0) = \lambda$, with

$\lambda \in (0,1)$. Note that the above POVM M_{AB} admits a separable decomposition as defined in the statement of Theorem III.22 with respect to the prime finite field $\mathbb{F}_2$, with $U = S$ and $V = T$, and

$$P_{Z|W}(0|0) = 1 - P_{Z|W}(0|1) = \lambda.$$

Hence the above theorem can be employed. This gives $S(U+V)_{\sigma_3} = 0.5155$, $S(U)_{\sigma_3} = S(V)_{\sigma_3} = 0.9999$, $S(U,V)_{\sigma_3} = 1.5154$, and $I(U,V)_{\sigma_3} = 0.4844$, where σ_3 is as defined in the statement of Theorem III.22. Since $S(U)_{\sigma_3} - S(U+V)_{\sigma_3} = S(V)_{\sigma_3} - S(U+V)_{\sigma_3} = I(U,V)_{\sigma_3}$, the constraints on R_1, R_2, $R_1 + C$ and $R_2 + C$ are the same as obtained in Theorem II.33. However, with $2S(U+V)_{\sigma_3} - S(U,V)_{\sigma_3} = -0.4844 < 0$, the constraint on $R_1 + R_2 + C_1 + C_2$ in the above theorem (3.24) is strictly weaker than the constraint obtained using random unstructured codes in Theorem II.33 of Chapter II. Therefore, the rate-region obtained above using random structured codes in Theorem III.22 is strictly larger than the rate-region obtained in Theorem II.33.

Example III.26. For the same state ρ_{AB} as in the above example, consider the following identical POVMs $M_A \triangleq \{\bar{\Lambda}_s^A\}_{s\in\mathcal{S}}$ and $M_B \triangleq \{\bar{\Lambda}_t^B\}_{t\in\mathcal{T}}$, where $\mathcal{S} = \mathcal{T} = \{0,1\}$, and

$$\Lambda_0^A = \Lambda_0^B \triangleq \begin{bmatrix} 0.4974 & 0.0471 + i0.4975 \\ 0.0471 - i0.4975 & 0.5026 \end{bmatrix}, \quad \Lambda_1^A = \Lambda_1^B \triangleq \begin{bmatrix} 0.5026 & -0.0471 - i0.4975 \\ -0.0471 + i0.4975 & 0.4974 \end{bmatrix}.$$

Let the joint measurement that Alice and Bob are trying to simulate be given by

$$\Gamma_z^{AB} \triangleq \sum_{s\in\{0,1\}} \sum_{t\in\{0,1\}} P_{Z|S,T}(z|s,t) \left(\Lambda_s^A \otimes \Lambda_t^B\right), \tag{3.26}$$

for $z \in \{0,1\}$ where $P_{Z|S,T} : \{0,1\} \to [0,1]$ is a conditional PMF on $\mathcal{Z} \times \mathcal{S} \times \mathcal{T}$ with $P_{Z|S,T}(0|0,0) = \delta_0 \in (0,1)$ and $P_{Z|S,T}(0|0,1) = P_{Z|S,T}(0|1,0) = P_{Z|S,T}(0|1,1) = \delta_1 \in (0,1)$. Note that $P_{Z|S,T}$ depends on the variables (s,t) only through $s \vee t$, the logical OR function. Now, we define the random variables U and V on the prime finite field $\mathbb{F}_3$ with the identity mappings $U = S$ and $V = T$, while noting that U and V take values in $\mathbb{F}_3$ with $P(U = 2) = P(V = 2) = 0$. Now with $W = U + V$, we identify the mapping $P_{Z|W}$ as $P_{Z|W}(0|0) = \delta_0$, and $P_{Z|W}(0|1) = P_{Z|W}(0|2) = \delta_1$. For this identification, we obtain $2S(U+V) - S(U,V) = -0.9039 < 0$, which gives the constraint on $R_1 + R_2 + C_1 + C_2$ in the above theorem (3.24) strictly weaker than the corresponding constraint

obtained using random unstructured codes in Theorem II.33. Since this is a biting constraint, the above rate-region is strictly larger than the former for this example.

Example III.27. Building upon Example III.26, we explore more points in the POVM space such that the above theorem provides constraints (3.24) that are strictly weaker than the corresponding constraint obtained in Theorem II.33 of Chapter II. For this, we consider the same state ρ_{AB}, as above and the following identical POVMs $M_A \triangleq \{\bar{\Lambda}_s^A\}_{s\in\mathcal{S}}$ and $M_B \triangleq \{\bar{\Lambda}_t^B\}_{t\in\mathcal{T}}$, where $\mathcal{S} = \mathcal{T} = \{0,1\}$, and

$$\Lambda_0^A = \Lambda_0^B = \begin{bmatrix} \theta_1 & \theta_2 + i\theta_3 \\ \theta_2 - i\theta_3 & 1-\theta_1 \end{bmatrix}, \quad \Lambda_1^A = \Lambda_1^B = I - \Lambda_0^A$$

for $\theta_i \in [-1,1]$[2]. Figure 3.2 illustrates the surface where $2S(U+V) = S(U,V)$ and therefore the region inside the surface has $2S(U+V) - S(U,V) < 0$, where the POVMs obtained provides the constraint on $R_1 + R_2 + C_1 + C_2$ in the above theorem (3.24) strictly weaker than the corresponding constraint obtained in Theorem II.33.

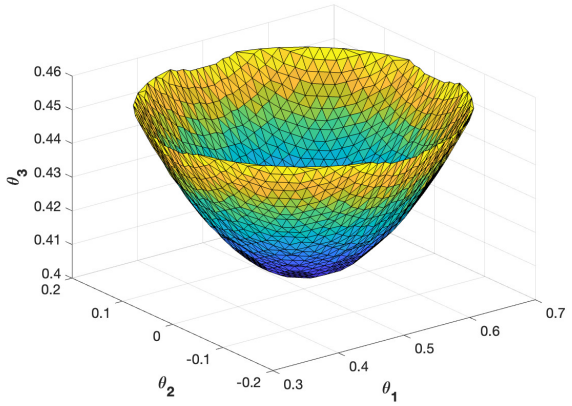

Figure 3.2: Shown above is a $(\theta_1, \theta_2, \theta_3)$-surface with POVMs satisfying $2S(U+V) = S(U,V)$. Although the surface is symmetric in θ_3, but for the ease of illustration only the upper half of the surface is shown.

Remark III.28. Note that for POVMs contained in the above $(\theta_1, \theta_2, \theta_3)$-surface of Example III.27, the sum rate constraint $R_1 + R_2 + C_1 + C_2$ is strictly weaker than the corresponding constraint obtained in Theorem II.33, and vice-versa outside.

[2]The above parametrization is only for illustrative purposes and do not constitute all the two dimensional POVMs.

3.5 Proof of the Inner Bound (Theorem III.22)

Suppose there exists a finite field $\mathbb{F}_p$, for a prime p, a pair of mappings $f_S : \mathcal{S} \to \mathbb{F}_p$ and $f_T : \mathcal{T} \to \mathbb{F}_p$, and a stochastic mapping $P_{Z|W} : \mathbb{F}_p \to \mathcal{Z}$ such that

$$P_{Z|S,T}(z|s,t) = P_{Z|W}(z|f_S(s) + f_T(t)), \quad \forall s \in \mathcal{S}, t \in \mathcal{T}, z \in \mathcal{Z},$$

yielding $U = f_S(S)$, and $V = f_T(T)$. This implies, we have POVMs $\bar{M}_A \triangleq \{\bar{\Lambda}_u^A\}_{u\in\mathcal{U}}$ and $\bar{M}_B \triangleq \{\bar{\Lambda}_v^B\}_{v\in\mathcal{V}}$ with $\mathcal{U} = \mathcal{V} \triangleq \mathbb{F}_p$ and a stochastic map $P_{Z|W} : \mathbb{F}_p \to \mathcal{Z}$, such that M_{AB} can be decomposed as

$$\Lambda_z^{AB} = \sum_{u,v} P_{Z|W}(z|u+v)\bar{\Lambda}_u^A \otimes \bar{\Lambda}_v^B, \ \forall z, \tag{3.27}$$

where W is defined as $W = U + V$. The coding strategy used here is based on Unionized Coset Codes, similar to the one employed in the point-to-point proof (Section 3.3.3), but extended to a distributed setting. Further, the structure in these codes provide a method to exploit the structure present in the stochastic processing applied by Charlie on the classical bits received, i.e., $P_{Z|U+V}$. Using this technique, we aim to strictly reduce the rate constraints compared to the ones obtained in Theorem II.33. Also note that, the results in the former are based on the assumption that approximating POVMs are all mutually independent. However, since the structured construction of the POVMs only guarantees pairwise independence among the operators of the POVM, the proofs below become significantly different from the proof of Theorem II.33.

We start by generating the canonical ensembles corresponding to $\bar{M}_A$ and $\bar{M}_B$, defined as

$$\begin{aligned} &\lambda_u^A \triangleq \mathrm{Tr}\{\bar{\Lambda}_u^A \rho_A\}, \quad \lambda_v^B \triangleq \mathrm{Tr}\{\bar{\Lambda}_v^B \rho_B\}, \quad \lambda_{uv}^{AB} \triangleq \mathrm{Tr}\{(\bar{\Lambda}_u^A \otimes \bar{\Lambda}_v^B)\rho_{AB}\}, \quad \text{and} \\ &\hat{\rho}_u^A \triangleq \frac{1}{\lambda_u^A}\sqrt{\rho_A}\bar{\Lambda}_u^A\sqrt{\rho_A}, \quad \hat{\rho}_v^B \triangleq \frac{1}{\lambda_v^B}\sqrt{\rho_B}\bar{\Lambda}_v^B\sqrt{\rho_B}, \quad \hat{\rho}_{uv}^{AB} \triangleq \frac{1}{\lambda_{uv}^{AB}}\sqrt{\rho_{AB}}(\bar{\Lambda}_u^A \otimes \bar{\Lambda}_v^B)\sqrt{\rho_{AB}}. \end{aligned} \tag{3.28}$$

With this notation, corresponding to each of the probability distributions, we can associate a δ-typical set. Let us denote $\mathcal{T}_\delta^{(n)}(U)$, $\mathcal{T}_\delta^{(n)}(V)$ and $\mathcal{T}_\delta^{(n)}(UV)$ as the δ-typical sets defined for $\{\lambda_u^A\}$, $\{\lambda_v^B\}$ and $\{\lambda_{uv}^{AB}\}$, respectively.

Let Π_{ρ_A} and Π_{ρ_B} denote the δ-typical projectors (as in (*Wilde*, 2013a, Def. 15.1.3)) for marginal density operators ρ_A and ρ_B, respectively. Also, for any $u^n \in \mathcal{U}^n$ and $v^n \in \mathcal{V}^n$, let $\Pi_{u^n}^A$ and $\Pi_{v^n}^B$

denote the strong conditional typical projectors (as in (*Wilde*, 2013a, Def. 15.2.4)) for the canonical ensembles $\{\lambda_u^A, \hat{\rho}_u^A\}$ and $\{\lambda_v^B, \hat{\rho}_v^B\}$, respectively.

For each $u^n \in \mathcal{T}_\delta^{(n)}(U)$ and $v^n \in \mathcal{T}_\delta^{(n)}(V)$ define

$$\tilde{\rho}_{u^n}^A \triangleq \Pi_{\rho_A} \Pi_{u^n}^A \hat{\rho}_{u^n}^A \Pi_{u^n}^A \Pi_{\rho_A}, \quad \tilde{\rho}_{v^n}^B \triangleq \Pi_{\rho_B} \Pi_{v^n}^B \hat{\rho}_{v^n}^B \Pi_{v^n}^B \Pi_{\rho_B},$$

and $\tilde{\rho}_{u^n}^A = 0$, and $\tilde{\rho}_{v^n}^B = 0$ for $u^n \notin \mathcal{T}_\delta^{(n)}(U)$ and $v^n \notin \mathcal{T}_\delta^{(n)}(V)$, respectively, with $\hat{\rho}_{u^n}^A \triangleq \bigotimes_i \hat{\rho}_{u_i}^A$ and $\hat{\rho}_{v^n}^B \triangleq \bigotimes_i \hat{\rho}_{v_i}^B$.

3.5.1 Construction of Structured POVMs

In what follows, we construct the random structured POVM elements. Fix a block length $n > 0$, positive integers N_1 and N_2, and a finite field $\mathbb{F}_p$. 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. Let $\tilde{\mu}_1 \in [1, \tilde{N}_1]$ and $\tilde{\mu}_2 \in [1, \tilde{N}_2]$ denote additional pairwise shared randomness used for random coding purposes. This randomness is only used to show the existence of a desired distributed protocol (as defined in Definition III.19), and is used only for bounding purposes. We denote $\bar{\mu}_i \triangleq (\mu_i, \tilde{\mu}_i)$, and $\bar{N}_i \triangleq N_i \cdot \tilde{N}_i$ for $i = 1, 2$. Further, let U and V be random variables defined on the alphabets $\mathcal{U}$ and $\mathcal{V}$, respectively, where $\mathcal{U} = \mathcal{V} = \mathbb{F}_p$. In building the code, we use the Unionized Coset Codes (UCCs) *Pradhan et al.* (2021) as defined above in Definition III.5.

For every $(\bar{\mu}_1, \bar{\mu}_2)$, consider two UCCs $(G, h_1^{(\bar{\mu}_1)})$ and $(G, h_2^{(\bar{\mu}_2)})$, each with parameters (n, k, l_1, p) and (n, k, l_2, p), respectively. Note that, for every $(\bar{\mu}_1, \bar{\mu}_2)$, they share the same generator matrix G.

For each $(\bar{\mu}_1, \bar{\mu}_2)$, the generator matrix G along with the function $h_1^{(\bar{\mu}_1)}$ and $h_2^{(\bar{\mu}_2)}$ generates p^{k+l_1} and p^{k+l_2} codewords, respectively. Each of these codewords are characterized by a triple $(a_i, m_i, \bar{\mu}_i)$, where $a_i \in \mathbb{F}_p^k$ and $m_i \in \mathbb{F}_p^{l_i}$ corresponds to the coarse code and the fine code indices, respectively, for $i \in [1, 2]$. Let $U^{n,(\bar{\mu}_1)}(a_1, i)$ and $V^{n,(\bar{\mu}_2)}(a_2, j)$ denote the codewords associated with Alice and Bob, generated using the above procedure, respectively, where

$$U^{n,(\bar{\mu}_1)}(a_1, i) \triangleq a_1 G + h_1^{(\bar{\mu}_1)}(i) \quad \text{and} \quad V^{n,(\bar{\mu}_2)}(a_2, j) \triangleq a_2 G + h_2^{(\bar{\mu}_2)}(j).$$

Now, construct the operators

$$\bar{A}_{u^n}^{(\bar{\mu}_1)} \triangleq \alpha_{u^n}\left(\sqrt{\rho_A}^{-1}\tilde{\rho}_{u^n}^{A}\sqrt{\rho_A}^{-1}\right) \quad \text{and} \quad \bar{B}_{v^n}^{(\bar{\mu}_2)} \triangleq \beta_{v^n}\left(\sqrt{\rho_B}^{-1}\tilde{\rho}_{v^n}^{B}\sqrt{\rho_B}^{-1}\right), \tag{3.29}$$

where

$$\alpha_{u^n} \triangleq \frac{1}{(1+\eta)}\frac{p^n}{p^{k+l_1}}\lambda_{u^n}^{A}, \qquad \text{and} \quad \beta_{v^n} \triangleq \frac{1}{(1+\eta)}\frac{p^n}{p^{k+l_2}}\lambda_{v^n}^{B}, \tag{3.30}$$

with $\eta \in (0,1)$ being a parameter to be determined. Having constructed the operators $\bar{A}_{u^n}^{(\bar{\mu}_1)}$ and $\bar{B}_{v^n}^{(\bar{\mu}_2)}$, we normalize these operators, so that they constitute a valid sub-POVM. To do so, we first define

$$\Sigma_A^{(\bar{\mu}_1)} \triangleq \sum_{u^n} \gamma_{u^n}^{(\bar{\mu}_1)}\bar{A}_{u^n}^{(\bar{\mu}_1)} \quad \text{and} \quad \Sigma_B^{(\bar{\mu}_2)} \triangleq \sum_{v^n} \zeta_{v^n}^{(\bar{\mu}_2)}\bar{B}_{v^n}^{(\bar{\mu}_2)},$$

where $\gamma_{u^n}^{(\bar{\mu}_1)}$ and $\zeta_{v^n}^{(\bar{\mu}_2)}$ are defined as

$$\gamma_{u^n}^{(\bar{\mu}_1)} \triangleq |\{(a_1,i) : U^{n,(\bar{\mu}_1)}(a_1,i) = u^n\}| \quad \text{and} \quad \zeta_{v^n}^{(\bar{\mu}_2)} \triangleq |\{(a_2,j) : V^{n,(\bar{\mu}_2)}(a_2,j) = v^n\}|.$$

Now, we define $\Pi_A^{\bar{\mu}_1}$ and $\Pi_B^{\bar{\mu}_2}$ as pruning operators for $\Sigma_A^{(\bar{\mu}_1)}$ and $\Sigma_B^{(\bar{\mu}_2)}$, with respect to Π_{ρ_A} and Π_{ρ_B}, respectively (see Definition III.2). Note that, these pruning operators, $\Pi_A^{\bar{\mu}_1}$ and $\Pi_B^{\bar{\mu}_2}$, depend on the triple $(G, h_1^{(\bar{\mu}_1)}, h_2^{(\bar{\mu}_2)})$. Using these pruning operators, for each $\bar{\mu}_1 \in [1, \bar{N}_1]$ and $\bar{\mu}_2 \in [1, \bar{N}_2]$, construct the sub-POVMs $M_1^{(n,\bar{\mu}_1)}$ and $M_2^{(n,\bar{\mu}_2)}$ as

$$M_1^{(n,\bar{\mu}_1)} \triangleq \{\gamma_{u^n}^{(\bar{\mu}_1)}A_{u^n}^{(\bar{\mu}_1)} : u^n \in \mathcal{U}^n\}, \quad \text{and} \quad M_2^{(n,\bar{\mu}_2)} \triangleq \{\zeta_{v^n}^{(\bar{\mu}_2)}B_{v^n}^{(\bar{\mu}_2)} : v^n \in \mathcal{V}^n\}, \tag{3.31}$$

where $A_{u^n}^{(\bar{\mu}_1)} = \Pi_A^{\mu_1}\bar{A}_{u^n}^{(\bar{\mu}_1)}\Pi_A^{\mu_1}$ and $B_{v^n}^{(\bar{\mu}_2)} = \Pi_B^{\mu_2}\bar{B}_{v^n}^{(\bar{\mu}_2)}\Pi_B^{\mu_2}$. Further, using these operators $\Pi_A^{\bar{\mu}_1}$ and $\Pi_B^{\bar{\mu}_2}$, we have $\sum_{u^n}\gamma_{u^n}^{(\bar{\mu}_1)}A_{u^n}^{(\bar{\mu}_1)} = \Pi_A^{\bar{\mu}_1}\Sigma_A^{(\bar{\mu}_1)}\Pi_A^{\bar{\mu}_1} \leq I$ and $\sum_{v^n}\zeta_{v^n}^{(\bar{\mu}_2)}B_{v^n}^{(\bar{\mu}_2)} = \Pi_B^{\bar{\mu}_2}\Sigma_B^{(\bar{\mu}_2)}\Pi_B^{\bar{\mu}_2} \leq I$, and thus $M_1^{(n,\bar{\mu}_1)}$ and $M_2^{(n,\bar{\mu}_2)}$ are valid sub-POVMs for all $\bar{\mu}_1 \in [1,\bar{N}_1]$ and $\bar{\mu}_2 \in [1,\bar{N}_2]$. Further, these collections $M_1^{(n,\bar{\mu}_1)}$ and $M_2^{(n,\bar{\mu}_2)}$ are completed using the operators $I - \sum_{u^n\in\mathcal{U}^n}\gamma_{u^n}^{(\bar{\mu}_1)}A_{u^n}^{(\bar{\mu}_1)}$ and $I - \sum_{v^n\in\mathcal{V}^n}\zeta_{v^n}^{(\bar{\mu}_2)}B_{v^n}^{(\bar{\mu}_2)}$.

3.5.2 Binning of POVMs

We next proceed to binning the above constructed collection of sub-POVMs. Since, UCC is already a union of several cosets, we associate a bin to each coset, and hence place all the codewords of a coset in the same bin. For each $i \in \mathbb{F}_p^{l_1}$ and $j \in \mathbb{F}_p^{l_2}$, let $\mathcal{B}_1^{(\bar{\mu}_1)}(i) \triangleq \mathbb{C}(G, h_1^{(\bar{\mu}_1)}(i))$ and $\mathcal{B}_2^{(\bar{\mu}_2)}(j) \triangleq \mathbb{C}(G, h_2^{(\bar{\mu}_2)}(j))$ denote the i^{th} and the j^{th} bins, respectively. Formally, we define the following operators:

$$\Gamma_i^{A,(\bar{\mu}_1)} \triangleq \sum_{u^n \in \mathcal{U}^n} \sum_{a_1 \in \mathbb{F}_p^k} A_{u^n}^{(\bar{\mu}_1)} \mathbb{1}_{\{a_1 G + h_1^{(\bar{\mu}_1)}(i) = u^n\}}, \quad \Gamma_j^{B,(\bar{\mu}_2)} \triangleq \sum_{v^n \in \mathcal{V}^n} \sum_{a_2 \in \mathbb{F}_p^{k_2}} B_{v^n}^{(\bar{\mu}_2)} \mathbb{1}_{\{a_2 G + h_2^{(\bar{\mu}_2)}(j) = v^n\}},$$

for all $i \in \mathbb{F}_p^{l_1}$ and $j \in \mathbb{F}_p^{l_2}$. Using these operators, we form the following collection:

$$M_A^{(n,\bar{\mu}_1)} \triangleq \{\Gamma_i^{A,(\bar{\mu}_1)}\}_{i \in \mathbb{F}_p^{l_1}}, \quad M_B^{(n,\bar{\mu})} \triangleq \{\Gamma_j^{B,(\bar{\mu}_2)}\}_{j \in \mathbb{F}_p^{l_2}}. \tag{3.32}$$

Note that if $M_1^{(n,\bar{\mu}_1)}$ and $M_2^{(n,\bar{\mu}_2)}$ are sub-POVMs, then so are $M_A^{(n,\bar{\mu}_1)}$ and $M_B^{(n,\bar{\mu}_2)}$, which is due to the relations

$$\sum_{i \in \mathbb{F}_p^{l_1}} \Gamma_i^{A,(\bar{\mu}_1)} = \sum_{u^n \in \mathcal{U}^n} \gamma_{u^n}^{(\mu_1)} A_{u^n}^{(\bar{\mu}_1)} \leq I, \quad \text{and} \quad \sum_{j \in \mathbb{F}_p^{l_2}} \Gamma_j^{B,(\bar{\mu}_2)} = \sum_{v^n \in \mathcal{V}^n} \zeta_{v^n}^{(\mu_2)} B_{v^n}^{(\bar{\mu}_2)} \leq I. \tag{3.33}$$

To make $M_A^{(n,\bar{\mu}_1)}$ and $M_B^{(n,\bar{\mu}_2)}$ complete, we define $\Gamma_0^{A,(\bar{\mu}_1)}$ and $\Gamma_0^{B,(\bar{\mu}_2)}$ as $\Gamma_0^{A,(\bar{\mu}_1)} = I - \sum_i \Gamma_i^{A,(\bar{\mu}_1)}$ and $\Gamma_0^{B,(\bar{\mu}_2)} = I - \sum_j \Gamma_j^{B,(\bar{\mu}_2)}$, respectively[3]. Now, we intend to use the completions $[M_A^{(n,\bar{\mu}_1)}]$ and $[M_B^{(n,\bar{\mu}_2)}]$ as the POVMs for encoders associated with Alice and Bob, respectively. Also, note that the effect of the binning is in reducing the communication rates from $(\frac{k+l_1}{n} \log p, \frac{k+l_2}{n} \log p)$ to (R_1, R_2), where $R_i \triangleq \frac{l_i}{n} \log p, i \in \{1, 2\}$. Now, we move on to describing the decoder.

3.5.3 Decoder mapping

We create a decoder that takes as an input a pair of bin numbers and produces a sequence $W^n \in \mathbb{F}_p^n$. More precisely, we define a mapping $F^{(\bar{\mu}_1, \bar{\mu}_2)}$, acting on the outputs of $[M_A^{(n,\bar{\mu}_1)}] \otimes [M_B^{(n,\bar{\mu}_2)}]$ as follows. On observing $(\bar{\mu}_1, \bar{\mu}_2)$ and the classical indices $(i, j) \in \mathbb{F}_p^{l_1} \times \mathbb{F}_p^{l_2}$ communicated by the

[3]Note that $\Gamma_0^{A,(\bar{\mu}_1)} = I - \sum_i \Gamma_i^{A,(\bar{\mu}_1)} = I - \sum_{u^n \in T_\delta^{(n)}(U)} A_{u^n}^{(\bar{\mu}_1)}$ and $\Gamma_0^{B,(\bar{\mu}_2)} = I - \sum_j \Gamma_j^{B,(\bar{\mu}_2)} = I - \sum_{v^n \in T_\delta^{(n)}(V)} B_{v^n}^{(\bar{\mu}_2)}$.

encoder, the decoder constructs $D^{(\bar{\mu}_1,\bar{\mu}_2)}$ and $F^{(\bar{\mu}_1,\bar{\mu}_2)}(\cdot,\cdot)$ as,

$$\begin{aligned} D_{i,j}^{(\bar{\mu}_1,\bar{\mu}_2)} &\triangleq \Big\{\tilde{a} \in \mathbb{F}_p^k : \tilde{a}G + h_1^{(\bar{\mu}_1)}(i) + h_2^{(\bar{\mu}_2)}(j) \in \mathcal{T}_{\hat{\delta}}^{(n)}(W)\Big\}, \\ F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j) &\triangleq \begin{cases} \tilde{a}G + h_1^{(\bar{\mu}_1)}(i) + h_2^{(\bar{\mu}_2)}(j) & \text{if } D_{i,j}^{(\bar{\mu}_1,\bar{\mu}_2)} \equiv \{\tilde{a}\} \\ w_0^n & \text{otherwise}\ , \end{cases} \end{aligned} \tag{3.34}$$

where $\hat{\delta} = p\delta$ and w_0^n is an arbitrary sequence in $\mathbb{F}_p^n \backslash \mathcal{T}_{\hat{\delta}}^{(n)}(W)$. Further, $F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j) = w_0^n$ for $i = 0$ or $j = 0$. Given this, we obtain the sub-POVM $\tilde{M}_{AB}$ with the following operators.

$$\tilde{\Lambda}_{w^n}^{AB} \triangleq \frac{1}{\bar{N}_1\bar{N}_2} \sum_{\bar{\mu}_1=1}^{\bar{N}_1} \sum_{\bar{\mu}_2=1}^{\bar{N}_2} \sum_{(i,j):F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j)=w^n} \Gamma_i^{A,(\bar{\mu}_1)} \otimes \Gamma_j^{B,(\bar{\mu}_2)},$$

$\forall w^n \in \mathbb{F}_p^n \bigcup \{w_0^n\}$. Now, we use the stochastic mapping $\mathsf{P}_{Z|W}$ to define the approximating sub-POVM $\hat{M}_{AB}^{(n)} \triangleq \{\hat{\Lambda}_{z^n}\}$ as

$$\hat{\Lambda}_{z^n}^{AB} \triangleq \sum_{w^n} \tilde{\Lambda}_{w^n}^{AB} P_{Z|W}^n(z^n|w^n),\ \forall z^n \in \mathcal{Z}^n.$$

Note that $\tilde{\Lambda}_{w^n}^{AB} = 0$ for $w^n \notin \mathcal{T}_{\delta}^{(n)}(W) \bigcup \{w_0^n\}$.

UCC Grand Ensemble: The generator matrix G and the functions $h_1^{(\bar{\mu}_1)}$ and $h_2^{(\bar{\mu}_2)}$ are chosen randomly uniformly and independently, for $\bar{\mu}_1 \in [1, \bar{N}_1]$ and $\bar{\mu}_2 \in [1, \bar{N}_2]$.

3.5.4 Trace Distance

In what follows, we show that $\hat{M}_{AB}^{(n)}$ is ϵ-faithful to $M_{AB}^{\otimes n}$ with respect to $\rho_{AB}^{\otimes n}$ (according to Def. I.1), where $\epsilon > 0$ can be made arbitrarily small. More precisely, using (3.27), we show that, $\mathbb{E}[K] \leq \epsilon$, where

$$K \triangleq \sum_{z^n} \left\| \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}}(\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B)\sqrt{\rho_{AB}^{\otimes n}} P_{Z|W}^n(z^n|u^n+v^n) - \sqrt{\rho_{AB}^{\otimes n}}\hat{\Lambda}_{z^n}^{AB}\sqrt{\rho_{AB}^{\otimes n}} \right\|_1, \tag{3.35}$$

and the expectation is with respect to the codebook generation.

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

Consider the second term within K, which can be written as

$$\sum_{w^n}\sqrt{\rho_{AB}^{\otimes n}}\tilde{\Lambda}_{w^n}^{AB}\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^n(z^n|w^n)$$
$$=\frac{1}{\bar{N}_1\bar{N}_2}\sum_{\bar{\mu}_1,\bar{\mu}_2}\sum_{i,j}\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_i^{A,(\bar{\mu}_1)}\otimes\Gamma_j^{B,(\bar{\mu}_2)}\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^n(z^n|F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j))\underbrace{\sum_{w^n}\mathbb{1}_{\{F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j)=w^n\}}}_{=1},$$

which can be written as $T+\widetilde{T}$, where

$$T\triangleq\frac{1}{\bar{N}_1\bar{N}_2}\sum_{\bar{\mu}_1,\bar{\mu}_2}\sum_{\{i>0\}\bigcap\{j>0\}}\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_i^{A,(\bar{\mu}_1)}\otimes\Gamma_j^{B,(\bar{\mu}_2)}\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^n(z^n|F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j)),$$
$$\widetilde{T}\triangleq\frac{1}{\bar{N}_1\bar{N}_2}\sum_{\bar{\mu}_1,\bar{\mu}_2}\sum_{\{i=0\}\bigcup\{j=0\}}\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_i^{A,(\bar{\mu}_1)}\otimes\Gamma_j^{B,(\bar{\mu}_2)}\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^n(z^n|w_0^n).$$

Hence, we have $K\leq S+\widetilde{S}$, where

$$S\triangleq\sum_{z^n}\left\|\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\bar{\Lambda}_{u^n}^A\otimes\bar{\Lambda}_{v^n}^BP_{Z|W}^n(z^n|u^n+v^n)\right)\sqrt{\rho_{AB}^{\otimes n}}-T\right\|_1,\tag{3.36}$$

and $\widetilde{S}\triangleq\sum_{z^n}\|\widetilde{T}\|_1$. Note that $\widetilde{S}$ captures the error induced by not covering the state $\rho_{AB}^{\otimes n}$.

Remark III.29. The terms corresponding to the operators that complete the sub-POVMs $M_A^{(n,\bar{\mu}_1)}$ and $M_B^{(n,\bar{\mu}_2)}$, i.e., $I-\sum_{u^n\in\mathcal{T}_\delta^{(n)}(U)}\gamma_{u^n}^{(\bar{\mu}_1)}A_{u^n}^{(\bar{\mu}_1)}$ and $I-\sum_{v^n\in\mathcal{T}_\delta^{(n)}(V)}\zeta_{v^n}^{(\bar{\mu}_2)}B_{v^n}^{(\bar{\mu}_2)}$ are taken care in $\widetilde{T}$. The expression T excludes these completing operators.

Step 2: Isolating the effect of error induced by binning

We begin by simplifying T as

$$T=\frac{1}{\bar{N}_1\bar{N}_2}\sum_{\bar{\mu}_1,\bar{\mu}_2}\sum_{u^n,v^n}\sum_{\substack{i>0,\\ j>0}}\sqrt{\rho_{AB}^{\otimes n}}\Bigg(\sum_{a_1\in\mathbb{F}_p^k}\sum_{a_2\in\mathbb{F}_p^k}A_{u^n}^{(\bar{\mu}_1)}\otimes B_{v^n}^{(\bar{\mu}_2)}\mathbb{1}_{\{a_1G+h_1^{(\bar{\mu}_1)}(i)=u^n\}}$$
$$\times\mathbb{1}_{\{a_2G+h_2^{(\bar{\mu}_2)}(j)=v^n\}}\Bigg)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^n(z^n|F^{(\bar{\mu}_1,\bar{\mu}_2)}(i,j)).$$

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 (3.36), and add and subtract an appropriate term within S and apply the triangle

inequality to isolate the effect of binning as $S \leq S_1 + S_2$, where

$$S_1 \triangleq \sum_{z^n} \left\| \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \frac{1}{\bar{N}_1 \bar{N}_2} \sum_{\bar{\mu}_1, \bar{\mu}_2} \gamma_{u^n}^{(\bar{\mu}_1)} A_{u^n}^{(\bar{\mu}_1)} \otimes \zeta_{v^n}^{(\bar{\mu}_2)} B_{v^n}^{(\bar{\mu}_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|W}^n(z^n | u^n + v^n) \right\|_1,$$

$$S_2 \triangleq \sum_{z^n} \left\| \frac{1}{\bar{N}_1 \bar{N}_2} \sum_{\bar{\mu}_1, \bar{\mu}_2} \sum_{\substack{i>0 \\ j>0}} \sum_{a_1, a_2} \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(A_{u^n}^{(\bar{\mu}_1)} \otimes B_{v^n}^{(\bar{\mu}_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \mathbb{1}_{\{a_1 G + h_1^{(\bar{\mu}_1)}(i) = u^n, a_2 G + h_2^{(\bar{\mu}_2)}(j) = v^n\}} \right.$$
$$\left. \times \left(P_{Z|W}^n(z^n | u^n + v^n) - P_{Z|W}^n \left(z^n | F^{(\bar{\mu}_1, \bar{\mu}_2)}(i,j) \right) \right) \right\|_1.$$

Note that the term S_1 characterizes the error introduced by approximation of the original POVM with the collection of approximating sub-POVMs $M_1^{(n,\bar{\mu}_1)}$ and $M_2^{(n,\bar{\mu}_2)}$, 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 III.30 (Mutual Packing)**.** *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[S_2] \leq \epsilon$, if $\frac{k+l_1}{n} \log p > I(U;RB)_{\sigma_1} - S(U)_{\sigma_3} + \log p$, $\frac{k+l_2}{n} \log p > I(V;RA)_{\sigma_2} - S(V)_{\sigma_3} + \log p$, $\frac{k+l_1}{n} \log p + \frac{1}{n} \log \bar{N}_1 > \log p$, $\frac{k+l_2}{n} \log p + \frac{1}{n} \log \bar{N}_2 > \log p$, $\frac{k}{n} \log p < \log p - S(W)_{\sigma_3}$, where σ_1, σ_2 and σ_3 are the auxiliary states as defined in the statement of the theorem.*

Proof. The proof is provided in Appendix B.7 □

Since averaged over $\tilde{\mu}_1 \in [1, \tilde{N}_1], \tilde{\mu}_2 \in [1, \tilde{N}_2]$, the quantity $\mathbb{E}[S_2]$ can be made arbitrarily small, there must exist a pair $(\tilde{\mu}_1, \tilde{\mu}_2)$ such that $\mathbb{E}[S_2]$ is small for this pair of $(\tilde{\mu}_1, \tilde{\mu}_2)$. For the rest of the proof, we fix $(\tilde{\mu}_1, \tilde{\mu}_2)$ to be this pair. The dependence of functions defined in the sequel on this pair is not made explicit for ease of notation. For the term corresponding to $\widetilde{S}$, we prove the following result.

Proposition III.31. *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[\widetilde{S}] \leq \epsilon$, if $\frac{k+l_1}{n} \log p > I(U;RB)_{\sigma_1} - S(U)_{\sigma_1} + \log p$ and $\frac{k+l_2}{n} \log p > I(V;RA)_{\sigma_2} - S(V)_{\sigma_2} + \log p$, where σ_1 and σ_2 are auxiliary states defined in the statement of the theorem.*

Proof. The proof is provided in Appendix B.8. □

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

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

$$Q_2 \triangleq \sum_{z^n} \left\| \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\gamma_{u^n}^{(\mu_1)} A_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B - \frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \gamma_{u^n}^{(\mu_1)} A_{u^n}^{(\mu_1)} \otimes \zeta_{v^n}^{(\mu_2)} B_{v^n}^{(\mu_2)} \right) \right.$$
$$\left. \times \sqrt{\rho_{AB}^{\otimes n}} P_{Z|W}^n(z^n|u^n + v^n) \right\|_1. \tag{3.37}$$

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 prove Q_1 is small, we characterize the rate constraints which ensure that an upper bound to Q_1 can be made to vanish in an expected sense. In addition, this upper bound becomes lucrative 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

$$J \triangleq \sum_{z^n, v^n} \left\| \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \gamma_{u^n}^{(\mu_1)} A_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|W}^n(z^n|u^n + v^n) \right\|_1. \tag{3.38}$$

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 σ_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., $\frac{k+l_1}{n} \log p + \frac{1}{n} \log N_1$, to be larger than $I(U; RZV)_{\sigma_3}$. We prove this in the following.

Note that from the triangle inequality, we have $Q_1 \leq J$. Further, we add and subtract appro-

priate terms within J, and again use the triangle inequality to obtain $J \leq J_1 + J_2$, where

$$J_1 \triangleq \sum_{z^n, v^n} \left\| \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \gamma_{u^n}^{(\mu_1)} \bar{A}_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|W}^n(z^n|u^n + v^n) \right\|_1,$$

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

Now we use the following proposition to bound the term corresponding to J_1.

Proposition III.32. *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[J_1] \leq \epsilon$ if $\frac{k+l_1}{n} \log p + \frac{1}{n} \log N_1 > I(U;RZV)_{\sigma_3} + \log p - S(U)_{\sigma_3}$, where σ_3 is the auxiliary state defined in the statement of the theorem.*

Proof. The proof of proposition is provided in Appendix B.9. □

Now we move on to bounding the term corresponding to J_2. We start by applying triangle inequality followed by Lemma II.11 on J_2 to obtain

$$\begin{aligned}
J_2 \leq & \sum_{z^n} \sum_{u^n, v^n} P_{Z|W}^n(z^n|u^n + v^n) \left\| \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sqrt{\rho_A^{\otimes n}} \left(\left(\gamma_{u^n}^{(\mu_1)} \bar{A}_{u^n}^{(\mu_1)} - \gamma_{u^n}^{(\mu_1)} A_{u^n}^{(\mu_1)} \right) \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_A^{\otimes n}} \right\|_1 \\
= & \sum_{u^n, v^n} \left\| \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sqrt{\rho_A^{\otimes n}} \left(\left(\gamma_{u^n}^{(\mu_1)} \bar{A}_{u^n}^{(\mu_1)} - \gamma_{u^n}^{(\mu_1)} A_{u^n}^{(\mu_1)} \right) \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_A^{\otimes n}} \right\|_1 \\
\leq & \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sum_{u^n} \gamma_{u^n}^{(\mu_1)} \left\| \sqrt{\rho_A^{\otimes n}} \left(\bar{A}_{u^n}^{(\mu_1)} - A_{u^n}^{(\mu_1)} \right) \sqrt{\rho_A^{\otimes n}} \right\|_1. \qquad (3.39)
\end{aligned}$$

Now we use the following proposition to bound the term corresponding to J_2.

Proposition III.33. *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[J_2] \leq \epsilon$ if $\frac{k+l_1}{n} \log p > I(U;RB)_{\sigma_1} + \log p - S(U)_{\sigma_3}$, where σ_1 and σ_3 are the auxiliary states defined in the statement of the theorem.*

Proof. The proof is provided in Appendix B.10. □

Since $Q_1 \leq J \leq J_1 + J_2$, hence $\mathbb{E}[J]$, and consequently $\mathbb{E}[Q_1]$, can be made arbitrarily small for sufficiently large n, if $\frac{k+l_1}{n} \log p + \frac{1}{n} \log N_1 > I(U;RZV)_{\sigma_3} - S(U)_{\sigma_3} + \log p$ and $\frac{k+l_1}{n} \log p > I(U;RB)_{\sigma_1} - S(U)_{\sigma_3} + \log p$. 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 III.34. *For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[Q_2] \leq \epsilon$, if $\frac{k+l_1}{n}\log p + \frac{1}{n}\log N_1 > I(U;RZV)_{\sigma_3} - S(U)_{\sigma_3} + \log p$, $\frac{k+l_2}{n}\log p + \frac{1}{n}\log N_2 > I(V;RZ)_{\sigma_3} - S(V)_{\sigma_3} + \log p$, $\frac{k+l_1}{n}\log p > I(U;RB)_{\sigma_1} - S(U)_{\sigma_3} + \log p$, and $\frac{k+l_2}{n}\log p > I(V;RA)_{\sigma_2} - S(V)_{\sigma_3} + \log p$ where $\sigma_1, \sigma_2, \sigma_3$ are the auxiliary states defined in the statement of the theorem.*

Proof. The proof is provided in Appendix B.11. □

3.5.5 Rate Constraints

To sum-up, we showed $\mathbb{E}[K] \leq \epsilon$ holds for sufficiently large n if the following bounds hold:

$$\tilde{R} + R_1 > I(U;RB)_{\sigma_1} - S(U)_{\sigma_3} + \log p, \tag{3.40a}$$

$$\tilde{R} + R_2 > I(V;RA)_{\sigma_2} - S(V)_{\sigma_3} + \log p, \tag{3.40b}$$

$$\tilde{R} + R_1 + C_1 > I(U;RZV)_{\sigma_3} - S(U)_{\sigma_3} + \log p, \tag{3.40c}$$

$$\tilde{R} + R_2 + C_2 > I(V;RZ)_{\sigma_3} - S(V)_{\sigma_3} + \log p, \tag{3.40d}$$

$$0 \leq \tilde{R} < \log p - S(U+V)_{\sigma_3}, \tag{3.40e}$$

$$C_1 \geq 0, \quad C_2 \geq 0, \tag{3.40f}$$

where $C_i \triangleq \frac{1}{n}\log_2 N_i, i \in \{1,2\}$ and $\tilde{R} \triangleq \frac{k}{n}\log p$. Therefore, there exists a distributed protocol with parameters $(n, 2^{nR_1}, 2^{nR_2}, 2^{nC_1}, 2^{nC_2})$ such that its overall POVM $\hat{M}_{AB}$ is ϵ-faithful to $M_{AB}^{\otimes n}$ with respect to $\rho_{AB}^{\otimes n}$.

Let us denote the above achievable rate-region by $\mathcal{R}_1$. By doing an exact symmetric analysis, but by replacing the first encoder by a product distribution instead of the second encoder in S_1 (as performed in (3.37)), all the constraints remain the same, except that the constraints on $\tilde{R}+R_1+C_1$

and $\tilde{R} + R_2 + C_2$ change as follows

$$\begin{aligned}\tilde{R} + R_1 + C_1 &\geq I(U;RZ)_{\sigma_3} - S(U)_{\sigma_3} + \log p,\\ \tilde{R} + R_2 + C_2 &\geq I(V;RZU)_{\sigma_3} - S(V)_{\sigma_3} + \log p.\end{aligned} \tag{3.41}$$

Let us denote the above achievable rate-region by $\mathcal{R}_2$. By time sharing between the any two points of $\mathcal{R}_1$ and $\mathcal{R}_2$ one can achieve any point in the convex closure of $(\mathcal{R}_1 \bigcup \mathcal{R}_2)$. The following lemma gives a symmetric characterization of the closure of convex hull of the union of the above achievable rate-regions.

Lemma III.35. *For the above defined rate regions $\mathcal{R}_1$ and $\mathcal{R}_2$, we have*

$$\mathcal{R}_3 = \textit{Convex Closure}(\mathcal{R}_1 \bigcup \mathcal{R}_2),$$

where $\mathcal{R}_3$ is given by the set of all the quintuples $(\tilde{R}, R_1, R_2, C_1, C_2)$ satisfying the following constraints:

$$\begin{aligned}\tilde{R} + R_1 &\geq I(U;RB)_{\sigma_1} - S(U)_{\sigma_3} + \log p,\\ \tilde{R} + R_2 &\geq I(V;RA)_{\sigma_2} - S(V)_{\sigma_3} + \log p,\\ \tilde{R} + R_1 + C_1 &\geq I(U;RZ)_{\sigma_3} - S(U)_{\sigma_3} + \log p,\\ \tilde{R} + R_2 + C_2 &\geq I(V;RZ)_{\sigma_3} - S(V)_{\sigma_3} + \log p,\\ 2\tilde{R} + R_1 + R_2 + C_1 + C_2 &\geq I(UV;RZ)_{\sigma_3} - S(U,V)_{\sigma_3} + 2\log p,\\ 0 \leq \tilde{R} &\leq \log p - S(U+V)_{\sigma_3},\\ R_1 \geq 0, R_2 \geq 0 \quad C_1 \geq 0, C_2 &\geq 0.\end{aligned} \tag{3.42}$$

Proof. The proof follows from elementary convex analysis. □

Lastly, we complete the proof of the theorem using the following lemma.

Lemma III.36. *Let $\bar{\mathcal{R}}_3$ denote the set of all quadruples (R_1, R_2, C_1, C_2) for which there exists $\tilde{R}$ such that the quintuple $(R_1, R_2, C_1, C_2, \tilde{R})$ satisfies the inequalities in (3.42). Let $\mathcal{R}_F$ denote the set of all quadruples (R_1, R_2, C_1, C_2) that satisfy the inequalities in (3.24) given in the statement*

of the theorem. Then, $\bar{\mathcal{R}}_3 = \mathcal{R}_F$.

Proof. The proof follows from Fourier-Motzkin elimination *Ziegler* (2012). □

3.6 Conclusion

In this chapter, we developed a technique of randomly generating structured POVMs using algebraic codes. Using this technique, we demonstrated a new achievable information-theoretic rate-region for the task of faithfully simulating a distributed quantum measurement and function computation. We further devised a Pruning Trace inequality which is a tighter version of the known operator Markov inequality, and a covering lemma which is independent of the operator Chernoff inequality, so as to analyse pairwise-independent POVM elements. Finally, combining these techniques, we demonstrated rate gains for this problem over traditional coding schemes, and provided a multi-party distributed faithful simulation and function computation protocol.

CHAPTER IV

Multi-party Purity Distillation

4.1 Introduction

In Chapters II and III the focus of our work was mainly QC setups. In this chapter, we will investigate distilling a quantum resource called purity. A primary task in quantum information theory is to quantify the amount of local and non-local information present within a quantum information source. For instance, the task of entanglement distillation aims at capturing the non-local correlations to transform a noisy shared state ρ^{AB} into pure bell states (in particular, the ebit $|\Phi^+\rangle$), in an aymptotic sense. A complementary notion to this task is the paradigm of local purity distillation, where pure ancilla qubits are distilled from a distributed state ρ^{AB} using local unitary operations.

Although it may seem unusual, local pure states cannot be considered as a free resource. One may argue that pure states can be obtained from a mixed state by performing a measurement, but this is only true after a measurement apparatus is initialized in a pure state. For this reason, the second law of thermodynamics recognizes purity as indeed a resource *Alicki et al.* (2004); *Horodecki et al.* (2005a). In this regard, the idea of distilling of local purity was first introduced in *Oppenheim et al.* (2002); *Horodecki et al.* (2003a) where the aim was to manipulate the qubits and concentrate the existing diluted form of purity. Two version of this problem have been introduced, (i) a single-party variant and (ii) a distributed version. In the former single-party scenario, also called as *local purity concentration*, many copies of a noisy state ρ^A are provided to Alice, and she aims at concentrating or extracting purity using only unitary operations. The authors in *Horodecki et al.* (2003b) characterized the asymptotic performance limit of this protocol ($\kappa(\rho^A)$) as the difference

between the number of qubits describing the system and the von Neumann entropy of the state ρ^A. For the latter case of distilling purity from a non-local distributed state, commonly termed as *local purity distillation*, two parties, Alice and Bob, share many copies of the noisy state ρ^{AB} and aim at jointly distilling pure ancilla qubits. Again, they are allowed to perform only local unitaries and but can communicate classically (LOCC), possibly through the use of a dephasing channel *Oppenheim et al.* (2002). Further, the protocols for both the variants require isolation (Closed-LOCC) from the environment which eliminated the possibility of unlimited consumption of the pure ancilla qubits. The authors in *Horodecki et al.* (2003a) provided bounds for this problem in the one-way and the two-way classical communication scenarios.

Later, Devetak in *Devetak* (2005a) considered a new paradigm called 1-CLOCC$'$, which was defined as an extension of Closed-LOCC, with (i) the allowance of using additional catalytic pure ancilla as long as these are returned back to the system, and (ii) the unlimited bidirectional classical communication replaced by unlimited one-way communication from Alice to Bob. Devetak obtained an information theoretic characterization of the distillable purity in the 1-CLOCC$'$ setting (allowing additional catalysts) and highlighted its connection to the earlier known one-way distillable common randomness measure *Devetak and Winter* (2004). The usage of catalytic resource to improve the quantum information tasks was first introduced in *Jonathan and Plenio* (1999). This further was extensively studied in a multitude of works, including but not limited to *Daftuar and Klimesh* (2001); *van Dam and Hayden* (2002); *Turgut* (2007); *Aubrun and Nechita* (2008); *Sanders and Gour* (2009); *Brandao et al.* (2015); *Duarte et al.* (2016); *Bu et al.* (2016); *Shiraishi and Sagawa* (2021); *Lipka-Bartosik and Skrzypczyk* (2021); *Ding et al.* (2021); *Takagi and Shiraishi* (2021). Building upon the work of *Devetak* (2005a), the authors in *Krovi and Devetak* (2007) extended the result to a setting with bounded one-way classical communication, again allowing for the additional catalytic resource. They improved upon the classical communication rate by using the Winter's approximate measurement *Winter* (2004), instead of an n-letter product measurement, and extracted purity for the states obtained thereby.

In this work, we revisit the task of distilling purity and consider a three-party setup. We ask the question of how many ancilla qubits can be distilled from a noisy state ρ^{ABC}, shared among three parties, Alice, Bob and Charlie. Similar to earlier problem formulation, we only allow local unitary operations at each party in a closed setting but permit the use of additional catalytic

ancillas with the promise of returning them at the end of the protocol. In addition, similar to *Krovi and Devetak* (2007), we only allow limited classical communication, which we model using a one-way multiple-access dephasing channel, with Alice and Bob as the senders and Charlie as the centralized receiver.

The contributions of this chapter can be summarized as follows. We first formulate a three-party purity distillation problem, and develop a 1-CLOCC$'$ multi-party purity distillation protocol for this problem capable of extracting purity from n copies of the noisy shared state $\rho_{ABC}^{\otimes n}$, using only local unitary operations and a one-way multiple-access dephasing channel. Further, for $\rho_{ABC}^{\otimes n}$, we define the asymptotic performance limit of the problem as the set of all triples (P, R_1, R_2), where P denotes the amount of purity that can be distilled from ρ^{ABC}, using R_1 and R_2 bits of classical communication. Then we characterize a quantum-information theoretic inner bound to the achievable rate region in terms of computable single-letter information quantities (see Theorem IV.4).

Toward the development of the results, we encounter two main challenges. The first challenge is in the compression of the joint measurements. Since the classical communication allowed by the protocol is limited, the joint measurements, that Alice and Bob employ, are required to be compressed. Although a distributed measurement compression protocol for compressing a joint measurement was developed in Chapter II, one cannot directly use this protocol as a complete black box. The reason for this is that the measurement compression protocol also requires additional resource of common randomness which the current purity distillation protocol does not allow. Apart from this, the measurement compression protocols provided in *Winter* (2004); *Wilde et al.* (2012) and Chapters II and III shows the "faithfulness" of the post-measurement state of the reference along with the classical-quantum register storing the measurement outcome. These protocols remain unconcerned about the post-measurement state of the system on which the measurement is performed. However, in the current problem the closeness of the latter is needed. To overcome this, we identify appropriate purifications of the post-measurement reference states and argue an existence of a collection of unitary operations achieving the latter (see Lemma IV.8 in the sequel).

The second major challenge is that after the application of the compressed measurement, the states across the three parties are not necessary separable. This is because a compressed measurement is usually not a "sharp" rank-one measurement. In *Devetak* (2005a) rank-one measurements

are employed which makes the states separable and hence eases the analysis. To handle this, we develop a result that captures how much a purity extracting protocol can disturb other correlated subsystems (see Lemma IV.10).

4.2 Preliminaries and Problem Formulation

4.2.1 Notation

We supplement the notations of the earlier chapters with the following. Recall that, given any natural number M, the finite set $\{1, 2, \cdots, M\}$ is denoted by $[1, M]$. Let $\mathcal{B}(\mathcal{H}), \mathcal{D}(\mathcal{H})$ denote the algebra of all bounded and density operators acting on $\mathcal{H}$, respectively. 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. Let $\kappa(\rho^A)$ denote the asymptotic purity distillable by local purity concentration protocols from ρ^A *Oppenheim et al.* (2002); *Horodecki et al.* (2003a). We know $\kappa(\rho^A) = \log \dim(\mathcal{H}_A) - S(\rho^A)$.

4.2.2 Problem Formulation

In the following we describe the problem statement. Let ρ^{ABC} be a density operator acting on $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. Consider two measurements M_A and M_B on sub-systems A and B, respectively. Imagine that we have three parties, named Alice, Bob and Charlie, trying to distill local purity from the noisy joint state ρ^{ABC}. The resources available to these parties are (i) the classical communication links of specified rates between Alice and Charlie, and Bob and Charlie, modelled as a multiple-access dephasing channel, and (ii) an additional triple of pure catalytic quantum systems A_C, B_C and C_C available to Alice, Bob and Charlie, respectively. Given the distributed nature of the problem, no communication is possible between Alice and Bob. The problem is formally defined in the following.

Definition IV.1. For a given finite set $\mathcal{Z}$, and a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, a distributed purity distillation protocol with parameters $(n, \Theta_1, \Theta_2, \kappa_1, \kappa_2, \kappa_3, \iota_1, \iota_2, \iota_3)$ is characterized by

1. a unitary operation on Alice's system $U_A \colon \mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_{A_C} \to \mathcal{H}_{A_p} \otimes \mathcal{H}_{X_1} \otimes \mathcal{H}_{A_g}$, with $\dim(\mathcal{H}_{A_p}) = \kappa_1$, $\dim(\mathcal{H}_{A_C}) = \iota_1$, and $\dim(\mathcal{H}_{X_1}) = \Theta_1$.

2. a unitary operation on Bob's system $U_B \colon \mathcal{H}_B^{\otimes n} \otimes \mathcal{H}_{B_C} \to \mathcal{H}_{B_p} \otimes \mathcal{H}_{X_2} \otimes \mathcal{H}_{B_g}$, with $\dim(\mathcal{H}_{B_p}) = \kappa_2$, $\dim(\mathcal{H}_{B_C}) = \iota_2$, and $\dim(\mathcal{H}_{X_2}) = \Theta_2$.

3. a multiple access dephasing channel $\mathcal{N} \colon \mathcal{H}_{X_1} \otimes \mathcal{H}_{X_2} \to \mathcal{H}_{X_1} \otimes \mathcal{H}_{X_2}$.

4. a unitary operation on Charlie's system $U_C \colon \mathcal{H}_C^{\otimes n} \otimes \mathcal{H}_{C_C} \otimes \mathcal{H}_{X_1} \otimes \mathcal{H}_{X_2} \to \mathcal{H}_{C_p} \otimes \mathcal{H}_{C_g} \otimes \mathcal{H}_{X_1} \otimes \mathcal{H}_{X_2}$, with $\dim(\mathcal{H}_{C_C}) = \iota_3$ and $\dim(\mathcal{H}_{C_p}) = \kappa_3$.

Definition IV.2. Given a quantum state $\rho^{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{C})$, a triple (P, R_1, R_2) is said to be achievable, if for all $\epsilon > 0$ and for all sufficiently large n, there exists a distributed purity distillation protocol with parameters $(n, \Theta_1, \Theta_2, \kappa_1, \kappa_2, \kappa_3, \iota_1, \iota_2, \iota_3)$ such that

$$G \triangleq \| \xi^{A_p B_p C_p} - |0\rangle\langle 0|^{A_p} \otimes |0\rangle\langle 0|^{B_p} \otimes |0\rangle\langle 0|^{C_p} \|_1 \leq \epsilon,$$
$$\frac{1}{n} \log_2 \Theta_i \leq R_i + \epsilon \colon i \in [2], \quad \frac{1}{n} \sum_{i \in [3]} (\log_2 \kappa_i - \log_2 \iota_i) \geq P - \epsilon,$$

where $|\xi\rangle \triangleq U_C \mathcal{N} U_B U_A |\Psi_\rho^{\otimes n}\rangle^{ABCR}$, and $|\Psi_\rho^{\otimes n}\rangle^{ABCR}$ is a purification of $(\rho^{ABC})^{\otimes n}$. The set of all achievable triples (P, R_1, R_2) is called the achievable rate region.

Given a POVM $M \triangleq \{\Lambda_x^A\}_{x \in \mathcal{X}}$ acting on ρ, the post-measurement state of the reference together with the classical outputs is represented by $(\mathrm{id} \otimes M)(\Psi_{RA}^\rho) \triangleq \sum_{x \in \mathcal{X}} |x\rangle\langle x| \otimes \mathrm{Tr}_A\{(I^R \otimes \Lambda_x^A)\Psi_{RA}^\rho\}$.

Definition IV.3. Consider a quantum state $\rho^{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$, and a POVM $M_{AB} = \bar{M}_A \otimes \bar{M}_B$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ where $\bar{M}_A = \{\bar{\Lambda}_s^A\}_{s \in \mathcal{S}}$ and $\bar{M}_B = \{\bar{\Lambda}_t^B\}_{t \in \mathcal{T}}$. Define the auxiliary states

$$\sigma_1^{RBCS} \triangleq (id_R \otimes \bar{M}_A \otimes id_{BC})(\Psi_\rho^{RABC}),$$
$$\sigma_2^{RACT} \triangleq (id_R \otimes id_{AC} \otimes \bar{M}_B)(\Psi_\rho^{RABC}), \quad \text{and}$$
$$\sigma_3^{RST} \triangleq \sum_{s,t} \sqrt{\rho^{AB}} \left(\bar{\Lambda}_s^A \otimes \bar{\Lambda}_t^B\right) \sqrt{\rho^{AB}} \otimes |s\rangle\langle s| \otimes |t\rangle\langle t|,$$

for some orthonormal sets $\{|s\rangle\}_{s \in \mathcal{S}}$ and $\{|t\rangle\}_{t \in \mathcal{T}}$, where Ψ_ρ^{RABC} is a purification of ρ^{ABC}. Let $\mathcal{R}_b(\rho^{ABC}, M_{AB})$ be defined as the set of all pairs (R_1, R_2) such that there exists finite sets $\mathcal{U}$ and $\mathcal{V}$ and a pair of mappings $f_S : \mathcal{S} \to \mathcal{U}$ and $f_T : \mathcal{T} \to \mathcal{V}$, yielding $U = f_S(S)$, $V = f_T(T)$, and

$W = (U, V)$, and the following inequalities are satisfied:

$$\begin{aligned} R_1 &\geq I(U;RBC)_{\sigma_1} - I_b(U;V)_{\sigma_3}, \\ R_2 &\geq I(V;RAC)_{\sigma_2} - I_b(U;V)_{\sigma_3}, \\ R_1 + R_2 &\geq I(U;RBC)_{\sigma_1} + I(V;RAC)_{\sigma_2} - I_b(U;V)_{\sigma_3}, \end{aligned}$$

where $I_b(\cdot)_\sigma = b \times I(\cdot)_\sigma$

4.3 Achievable purity and communication rates: An Inner Bound

Theorem IV.4. *Given a quantum state $\rho^{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$, a triple (R_1, R_2, P) is achievable if there exists a POVM $M_{AB} = \bar{M}_A \otimes \bar{M}_B$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ with POVMs $\bar{M}_A = \{\Lambda_s^A\}_{s\in\mathcal{S}}$ and $\bar{M}_B = \{\Lambda_t^B\}_{t\in\mathcal{T}}$ $\mathcal{H}_A \otimes \mathcal{H}_B$ and a real number $b \in [0,1]$ such that the following holds:*

$$P \leq \kappa(\rho_A) + \kappa(\rho_B) + \kappa(\rho_C) + I(C;U,V)_\sigma - I_b(U;V)_{\sigma_3},$$

and $(R_1, R_2) \in \mathcal{R}_b(\rho^{ABC}, M_{AB})$, where

$$\sigma^{RCST} \triangleq (\mathrm{id}_R \otimes \mathrm{id}_C \otimes \bar{M}_A \otimes \bar{M}_B)(\Psi_\rho^{RABC}).$$

Proof. The proof is provided in Section 4.4. □

Definition IV.5. Given a quantum state $\rho^{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$, and a dephasing channel with communication links of rates R_1 and R_2 define 1-way distillable distributed local purity $\kappa_\rightarrow(\rho^{ABC}, R_1, R_2)$ as the supremum of the sum of all the locally distillable purity.

Corollary IV.6. *Given a quantum state $\rho^{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$, let*

$$\begin{aligned} \kappa^I_\rightarrow(\rho^{ABC}, R_1, R_2) &\triangleq \kappa(\rho^A) + \kappa(\rho^B) + \kappa(\rho^C) + P^D_\rightarrow(\rho^{ABC}, R_1, R_2), \\ P^D_\rightarrow(\rho^{ABC}, R_1, R_2) &\triangleq \frac{1}{n} \lim_{n\to\infty} \bar{P}^D_\rightarrow((\rho^{ABC})^{\otimes n}, nR_1, nR_2), \\ \bar{P}^D_\rightarrow(\rho^{ABC}, R_1, R_2) &\triangleq \max_{M_{AB}, b\in[0,1]} \{I(C;U,V)_\sigma - I_b(U;V)_\sigma : (R_1, R_2) \in \mathcal{R}_b(\rho^{ABC}, M_{AB})\}. \end{aligned}$$

With the above definitions, we have $\kappa^I_\rightarrow(\rho^{ABC}, R_1, R_2) \leq \kappa_\rightarrow(\rho^{ABC}, R_1, R_2)$. In other words, for

any communication rates (R_1, R_2), $\kappa^I_{\rightarrow}(\rho^{ABC}, R_1, R_2)$ *amount of purity can be jointly distilled from the three parties using the protocol defined in Def. IV.1.*

Proof. The proof follows from Theorem IV.4 and performing regularization. □

4.4 Proof of the Inner Bound (Theorem IV.4)

The proof is mainly composed of two parts. In the first part, we construct a protocol by developing all the actions of the three parties, and describe them as unitary evolution (as these are the only actions allowed by the protocol, Def. IV.1). Simultaneously, we also provide necessary lemmas needed for the next part. The second part deals with characterizing the action of the developed unitary operators on the shared quantum state ρ^{ABC} and then bounding the error between the final state and the desired pure state. Since our result is derived for a bounded communication channel, we start by approximating the measurements to achieve a decreased outcome set, while preserving the statistics of the measurement. Let $M_A \triangleq \{\Lambda^A_u\}$ and $M_B \triangleq \{\Lambda^B_v\}$ denote the POVMs after the maps f_S and f_T, respectively.

4.4.1 Approximation of the measurement $M_A \otimes M_B$

We start by generating the canonical ensembles corresponding to M_A and M_B, defined as

$$\lambda^A_u \triangleq \mathrm{Tr}\{\Lambda^A_u \rho^A\}, \quad \lambda^B_v \triangleq \mathrm{Tr}\{\Lambda^B_v \rho^B\}, \quad \lambda^{AB}_{uv} \triangleq \mathrm{Tr}\{(\Lambda^A_u \otimes \Lambda^B_v)\rho^{AB}\}, \quad \text{and}$$
$$\hat{\rho}^A_u \triangleq \frac{1}{\lambda^A_u}\sqrt{\rho^A}\Lambda^A_u\sqrt{\rho^A}, \quad \hat{\rho}^B_v \triangleq \frac{1}{\lambda^B_v}\sqrt{\rho^B}\Lambda^B_v\sqrt{\rho^B}, \quad \hat{\rho}^{AB}_{uv} \triangleq \frac{1}{\lambda^{AB}_{uv}}\sqrt{\rho^{AB}}(\Lambda^A_u \otimes \Lambda^B_v)\sqrt{\rho^{AB}}. \tag{4.2}$$

Let Π_{ρ_A} and Π_{ρ_B} denote the δ-typical projectors (as in (*Wilde*, 2013a, Def. 15.1.3)) for marginal density operators ρ^A and ρ^B, respectively. Also, for any $u^n \in \mathcal{U}^n$ and $v^n \in \mathcal{V}^n$, let $\Pi^A_{u^n}$ and $\Pi^B_{v^n}$ denote the strong conditional typical projectors (as in (*Wilde*, 2013a, Def. 15.2.4)) for the canonical ensembles $\{\lambda^A_u, \hat{\rho}^A_u\}$ and $\{\lambda^B_v, \hat{\rho}^B_v\}$, respectively.

For each $u^n \in \mathcal{T}^{(n)}_\delta(U)$ and $v^n \in \mathcal{T}^{(n)}_\delta(V)$ define

$$\tilde{\rho}^{A'}_{u^n} \triangleq \Pi_{\rho_A}\Pi^A_{u^n}\hat{\rho}^A_{u^n}\Pi^A_{u^n}\Pi_{\rho_A}, \quad \tilde{\rho}^{B'}_{v^n} \triangleq \Pi_{\rho_B}\Pi^B_{v^n}\hat{\rho}^B_{v^n}\Pi^B_{v^n}\Pi_{\rho_B},$$

and $\tilde{\rho}^A_{u^n} = 0$, and $\tilde{\rho}^B_{v^n} = 0$ for $u^n \notin \mathcal{T}^{(n)}_\delta(U)$ and $v^n \notin \mathcal{T}^{(n)}_\delta(V)$, respectively, with $\hat{\rho}^A_{u^n} \triangleq \bigotimes_i \hat{\rho}^A_{u_i}$

and $\hat{\rho}^B_{v^n} \triangleq \bigotimes_i \hat{\rho}^B_{v_i}$. Randomly and independently select $2^{n\tilde{R}_1}$ and $2^{n\tilde{R}_2}$ sequences $(U^n(l), V^n(k))$ according to the pruned distributions, i.e.,

$$\mathbb{P}\left((U^n(l), V^n(k)) = (u^n, v^n)\right) = \begin{cases} \dfrac{\lambda^A_{u^n}}{(1-\varepsilon)} \dfrac{\lambda^B_{v^n}}{(1-\varepsilon')} & \text{for} \quad u^n \in \mathcal{T}^{(n)}_\delta(U), v^n \in \mathcal{T}^{(n)}_\delta(V) \\ 0 & \text{otherwise} \end{cases}, \tag{4.3}$$

where $\varepsilon = \sum_{u^n \in \mathcal{T}^{(n)}_\delta(U)} \lambda^A_{u^n}$ and $\varepsilon' = \sum_{v^n \in \mathcal{T}^{(n)}_\delta(V)} \lambda^B_{v^n}$. Let $\mathcal{C}$ denote the codebook containing all pairs of codewords $(U^n(l), V^n(k))$. Further, define $\sigma^{A'}$ and $\sigma^{B'}$ as

$$\sigma^{A'} \triangleq \sum_{u^n \in \mathcal{T}^{(n)}_\delta(U)} \frac{\lambda^A_{u^n}}{(1-\varepsilon)} \tilde{\rho}^{A'}_{u^n}, \quad \sigma^{B'} \triangleq \sum_{v^n \in \mathcal{T}^{(n)}_\delta(V)} \frac{\lambda^B_{v^n}}{(1-\varepsilon')} \tilde{\rho}^{B'}_{v^n}, \tag{4.4}$$

where $\varepsilon = \sum_{u^n \in \mathcal{T}^{(n)}_\delta(U)} \lambda^A_{u^n}$ and $\varepsilon' = \sum_{v^n \in \mathcal{T}^{(n)}_\delta(V)} \lambda^B_{v^n}$. Note that $\sigma^{A'}$ and $\sigma^{B'}$ defined above are expectations with respect to the pruned distribution *Wilde* (2013a). Let $\hat{\Pi}^A$ and $\hat{\Pi}^B$ be the projectors onto the subspaces spanned by the eigenstates of $\sigma^{A'}$ and $\sigma^{B'}$ corresponding to eigenvalues that are larger than $\varepsilon 2^{-n(S(\rho_A)+\delta_1)}$ and $\varepsilon' 2^{-n(S(\rho_B)+\delta_1)}$, where $\delta_1 > 0$ is such that $\mathrm{Tr}(\Pi_{\rho_A}) \leq 2^{n(S(\rho_A)+\delta_1)}$, and $\mathrm{Tr}(\Pi_{\rho_B}) \leq 2^{n(S(\rho_B)+\delta_1)}$, and $\delta_1 \searrow 0$ as $\delta \searrow 0$. Lastly, define

$$\tilde{\rho}^A_{u^n} \triangleq \hat{\Pi}^A \tilde{\rho}^{A'}_{u^n} \hat{\Pi}^A, \quad \text{and} \quad \tilde{\rho}^B_{v^n} \triangleq \hat{\Pi}^B \tilde{\rho}^{B'}_{v^n} \hat{\Pi}^B.$$

Using these definitions, for any given $\epsilon \in (0,1)$, and sufficiently large n and sufficiently small δ, we have

$$\sum_{u^n \in \mathcal{U}^n} \lambda^A_{u^n} \mathrm{Tr}\{\tilde{\rho}^A_{u^n}\} \geq 1-\epsilon, \quad \sum_{u^n \in \mathcal{U}^n} \lambda^A_{u^n} \|\hat{\rho}^A_{u^n} - \tilde{\rho}^A_{u^n}\|_1 \leq \epsilon, \quad \sum_{v^n \in \mathcal{V}^n} \lambda^B_{v^n} \|\hat{\rho}^B_{v^n} - \tilde{\rho}^B_{v^n}\|_1 \leq \epsilon. \tag{4.5}$$

(A detailed proof of the statement can be found in (*Wilde et al.*, 2012, Eqs 28 and 35).) Using these definitions, construct operators

$$A_{u^n} \triangleq \gamma_{u^n}\left(\sqrt{\rho_A}^{-1} \tilde{\rho}^A_{u^n} \sqrt{\rho_A}^{-1}\right) \quad \text{and} \quad B_{v^n} \triangleq \zeta_{v^n}\left(\sqrt{\rho_B}^{-1} \tilde{\rho}^B_{v^n} \sqrt{\rho_B}^{-1}\right), \tag{4.6}$$

where

$$\gamma_{u^n} \triangleq \frac{1-\varepsilon}{1+\eta} 2^{-n\tilde{R}_1} |\{l : U^n(l) = u^n\}| \quad \text{and } \zeta_{v^n} \triangleq \frac{1-\varepsilon'}{1+\eta} 2^{-n\tilde{R}_2} |\{k : V^n(k) = v^n\}|, \tag{4.7}$$

and $\eta \in (0,1)$ is a parameter that determines the probability of not obtaining sub-POVMs. Then construct $M_1^{(n)}$ and $M_2^{(n)}$ as in the following

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

We show later that $M_1^{(n)}$ and $M_2^{(n)}$ form sub-POVMs, with high probability, These collections $M_1^{(n)}$ and $M_2^{(n)}$ are completed using the operators $I - \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} A_{u^n}$ and $I - \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} B_{v^n}$, and these operators are associated with sequences u_0^n and v_0^n, which are chosen arbitrarily from $\mathcal{U}^n \backslash \mathcal{T}_\delta^{(n)}(U)$ and $\mathcal{V}^n \backslash \mathcal{T}_\delta^{(n)}(V)$, respectively. Let $\mathbb{1}_{\{\text{sP-i}\}}$ denote the indicator random variable corresponding to the event that $M_i^{(n)}$ form sub-POVM for $i = 1, 2$. We use the trivial POVM $\{I\}$ in the case of the complementary event and associate it with u_0^n and v_0^n as the case maybe. In summary, the POVMs are given by $\{\mathbb{1}_{\{\text{sP-1}\}} A_{u^n} + (1 - \mathbb{1}_{\{\text{sP-1}\}}) \mathbb{1}_{\{u^n = u_0^n\}} I\}_{u^n \in \mathcal{U}^n}$, and $\{\mathbb{1}_{\{\text{sP-2}\}} B_{v^n} + (1 - \mathbb{1}_{\{\text{sP-2}\}}) \mathbb{1}_{\{v^n = v_0^n\}} I\}_{v^n \in \mathcal{V}^n}$.

Now, we intend to use the completions $[M_1^{(n,\bar{\mu}_1)}]$ and $[M_2^{(n,\bar{\mu}_2)}]$ in constructing the unitaries U_A and U_B, as described in the protocol (Def. IV.1), for Alice and Bob, respectively. Before concluding the discussion on the POVMs, we provide two lemmas which would be useful in the sequel. The first lemma deals with bounding from below the probability that the constructed collection of operators indeed form a sub-POVM. Toward this, observe that the collections of approximating POVMs, $\{A_{U^n(l)}\}$ and $\{B_{V^n(k)}\}$, constructed in this work are identical to the ones employed in *Winter* (2004) (c.f. Chapter II) which allows us to use the Operator Chernoff Bound to establish that these collections form a sub-POVM with high probability. Toward this consider, the following proposition.

Proposition IV.7. *For any $\epsilon \in (0,1)$, any $\eta \in (0,1)$, any $\delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have*

$$\mathbb{E}\left[\mathbb{1}_{\{sP\text{-}1\}} \mathbb{1}_{\{sP\text{-}2\}}\right] > 1 - \epsilon,$$

if $\tilde{R}_1 > I(U;RB)_{\sigma_1}$ and $\tilde{R}_2 > I(V;RA)_{\sigma_2}$, where σ_1, σ_2 are defined as in the statement of the

theorem.

Proof. Observe that the collections $\{A_{U^n(l)}\}$ and $\{B_{V^n(k)}\}$ satisfy all the hypotheses of the operator Chernoff bound *Ahlswede and Winter* (2002) after identifying Π as the cut-off operator employed in *Winter* (2004). Now by following identical steps as in *Winter* (2004), the result follows. □

The second lemma provides a unitary to show closeness of the post-measurement states obtained from approximating measurements and the actual measurements. Note that the faithful simulation results *Winter* (2004); *Wilde et al.* (2012) and of Chapters II and III show the closeness of states in the reference system, but the current result proves the closeness of the post-measurement states. The main elements of the proof is in identifying appropriate purifications and using the Uhlmann's Theorem *Wilde* (2013a). Before stating the lemma, consider the following definitions. Let,

$$A_{U^n(l)} \triangleq \gamma \sqrt{\rho_A}^{-1} \tilde{\rho}^A_{U^n(l)} \sqrt{\rho_A}^{-1} \quad \text{and} \quad B_{V^n(k)} \triangleq \zeta \sqrt{\rho_B}^{-1} \tilde{\rho}^B_{V^n(k)} \sqrt{\rho_B}^{-1},$$

for $\gamma \triangleq \frac{1-\varepsilon}{1+\eta} 2^{-n\tilde{R}_1}$ and $\zeta \triangleq \frac{1-\varepsilon'}{1+\eta} 2^{-n\tilde{R}_2}$. The lemma is as follows.

Lemma IV.8. *Using the above definitions, for all $l \in [2^{n\tilde{R}_1}]$ and $k \in [2^{n\tilde{R}_2}]$, let*

$$\hat{\sigma}^{AE}_l \triangleq \frac{(I^E \otimes \sqrt{\Lambda^A_{U^n(l)}}) \left|\Psi_{\rho^{\otimes n}}\right\rangle^{ABCR}}{\sqrt{\lambda_{U^n(l)}}} \quad \textit{and} \quad \tilde{\sigma}^{AE}_l \triangleq \frac{(I^E \otimes \sqrt{A_{U^n(l)}}) \left|\Psi_{\rho^{\otimes n}}\right\rangle^{ABCR}}{\sqrt{\gamma}\sqrt{\mathrm{Tr}\left\{\tilde{\rho}^A_{U^n(l)}\right\}}},$$

($\hat{\sigma}^{BF}_k$ and $\tilde{\sigma}^{BF}_k$ defined analogously) where E and F denotes the system BCR and ACR, respectively, then for each $l \in [2^{n\tilde{R}_1}]$ and $k \in [2^{n\tilde{R}_2}]$, there exists a pair of unitaries $U^A_r(l)$ and $U^B_r(k)$, such that

$$\sqrt{F(\hat{\sigma}^{AE}_l, (I^E \otimes U^A_r(l))\tilde{\sigma}^{AE}_l)} \geq 1 - \frac{1}{2}\|\hat{\rho}^A_l - \tilde{\rho}^A_l\|_1 - \frac{1}{2}\left|1 - \mathrm{Tr}\{\tilde{\rho}^A_l\}\right|, \textit{ and}$$
$$\sqrt{F(\hat{\sigma}^{BF}_k, (I^F \otimes U^B_r(k))\tilde{\sigma}^{BF}_k)} \geq 1 - \frac{1}{2}\|\hat{\rho}^B_k - \tilde{\rho}^B_k\|_1 - \frac{1}{2}\left|1 - \mathrm{Tr}\{\tilde{\rho}^B_k\}\right|.$$

Proof. The proof is provided in Appendix C.1. □

We now move on to characterizing the unitaries U_A and U_B.

4.4.2 Action of Alice and Bob

Using the approximating POVMs constructed above, as a first unitary operation, Alice and Bob perform a coherent version of the approximating POVM. This is defined as

$$U_M^{AL} \triangleq \sum_{l\in[2^{n\tilde{R}_1}]} \sqrt{A_{U^n(l)}} \otimes |l\rangle\,, \quad U_M^{BK} \triangleq \sum_{k\in[2^{n\tilde{R}_2}]} \sqrt{B_{V^n(k)}} \otimes |k\rangle\,.$$

From now on, for ease of notation, we use $\Lambda_l^A, \Lambda_k^B, \lambda_l^A, \lambda_k^B, A_l, B_k, \hat{\sigma}_l^A, \hat{\sigma}_k^B, \gamma_l$, and ζ_k to denote the corresponding $n-$letter objects constructed for the codewords $U^n(l)$ and $V^n(k)$, respectively.

Although the operators defined above are isometry operators, but with the help of additional catalyst qubits, these can be implemented as unitary operators. The rate at which these catalytic ancilla are used will be characterized and subtracted from the total rate at which the protocol produces them. Now, to extract purity from the states obtained after performing the measurements we employ the approach of *Krovi and Devetak* (2007). More formally, we define the collection of unitaries $\{U_p^A(l)\}_{l\in[2^{n\tilde{R}_1}]}$ and $\{U_p^B(k)\}_{k\in[2^{n\tilde{R}_2}]}$ as those that can extract purity for the collection of states $\{\hat{\sigma}_l^A\}_{l\in[2^{n\tilde{R}_1}]}$ and $\{\hat{\sigma}_k^B\}_{k\in[2^{n\tilde{R}_2}]}$, respectively. Note that since $\hat{\sigma}_l^A$ and $\hat{\sigma}_k^B$ are product states, we use a type based construction (similar to one proposed in *Krovi and Devetak* (2007)) in designing the unitary operators $U_p^A(l)$ and $U_p^B(k)$.

Now we characterize the complete action at Alice and Bob as

$$U_A \triangleq U_P^{AL} U_R^{AL} U_M^{AL} \;\text{ and }\; U_B \triangleq U_P^{BK} U_R^{BK} U_M^{BK}, \tag{4.8}$$

where U_P^{AL} and U_R^{AL} are controlled unitary operators defined as

$$U_P^{AL} \triangleq \sum_{l\in[2^{n\tilde{R}_1}]} U_p^A(l) \otimes |l\rangle\langle l|\,, \quad U_R^{AL} \triangleq \sum_{l\in[2^{n\tilde{R}_1}]} U_r^A(l) \otimes |l\rangle\langle l|\,, \tag{4.9}$$

and similar is true for U_P^{BK} and U_R^{BK}. This gives

$$U_A = \sum_{l\in[2^{n\tilde{R}_1}]} U_p^A(l) U_r^A(l) \sqrt{A_l} \otimes |l\rangle\,, \quad U_B = \sum_{k\in[2^{n\tilde{R}_2}]} U_p^B(k) U_r^B(k) \sqrt{B_k} \otimes |k\rangle\,.$$

Before introducing the action of Charlie, we provide two useful lemmas. The first lemma

characterizes the purity that can be extracted individually by Alice and Bob, and the second lemma integrates the action of the two parties and provides a way to cumulatively analyze the effect of the two parties on the overall state. The first lemma can be stated as follows:

Lemma IV.9. *Using the above defined unitary operations, let $\mathcal{N}_A : \mathcal{H}_A^{\otimes n} \to \mathcal{H}_{A_p}$ correspond to the protocol that extracts purity from ρ_A using the above defined unitaries, i.e.,*

$$\mathcal{N}_A(\rho_A) \triangleq \mathrm{Tr}_{A_g L}\left\{(I^{RBC} \otimes U_A)\rho_A^{\otimes n}(I^{RBC} \otimes U_A)^\dagger\right\}.$$

Then, for any given $\epsilon \in (0,1)$, we have

$$\mathbb{E}_A\left[\left\|\mathcal{N}_A(\rho_A) - |0\rangle\langle 0|_{A_p}\right\|_1 \mathbb{1}_{\{sP\}}\right] \leq \epsilon, \tag{4.10}$$

where $\mathcal{H}_A^{\otimes n} = \mathcal{H}_{A_g} \otimes \mathcal{H}_{A_p}$ and $\frac{1}{n}\log\dim\mathcal{H}_{A_p} < \log\dim\mathcal{H}_A - \sum_u \lambda_u^A S(\rho_u^A)$, for all sufficiently large n, and sufficiently small $\eta, \delta > 0$.

Proof. The proof is provided in Appendix C.2. □

The second lemma is stated as follows:

Lemma IV.10. *Given a protocol $\mathcal{N}_A : \mathcal{H}_A^{\otimes n} \to \mathcal{H}_{A_p}$ (as described above) capable of extracting purity from ρ_A, let $V_A : \mathcal{H}_A^{\otimes n} \to \mathcal{H}_{A_p} \otimes \mathcal{H}_{A_g} \otimes \mathcal{H}_L$ be a Stinespring's dilation of $\mathcal{N}_A$. Then,*

$$\mathbb{E}_A\left[\left\|(I^{A_R} \otimes V_A)\Psi_{\rho_A^{\otimes n}}^{A_R A}(I^{A_R} \otimes V_A)^\dagger - \rho_{A_R A_g L} \otimes |0\rangle\langle 0|_{A_p}\right\|_1\right] \leq 4\sqrt{\mathbb{E}_A\left[\left\|\mathcal{N}_A(\rho_A) - |0\rangle\langle 0|_{A_p}\right\|_1\right]}, \tag{4.11}$$

where

$$\rho_{A_R A_g L} = \mathrm{Tr}_{A_p}\left\{(I^{A_R} \otimes V_A)\Psi_{\rho_A^{\otimes n}}^{A_R A}(I^{A_R} \otimes V_A)^\dagger\right\},$$

and $\Psi_{\rho_A^{\otimes n}}^{A_R A}$ is a purification of $\rho_A^{\otimes n}$.

Proof. The proof is provided in Appendix C.3. □

4.4.3 Transmission over the Dephasing Channel $\mathcal{N}$

Before we proceed to employ the dephasing channel, observe that the classical registers created by the coherent measurement contains correlations across Alice and Bob. These correlations could be exploited which can further reduce the communication needed over the dephasing channel. For this, we employ the traditional binning operation. Begin by fixing the binning rates (R_1, R_2), with $R_1 \leq \tilde{R}_1$ and $R_2 \leq \tilde{R}_2$. 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 $[1, 2^{nR_2}]$. For each $i \in [1, 2^{nR_1}]$ and $j \in [1, 2^{nR_2}]$, let $\mathcal{B}_1(i)$ and $\mathcal{B}_2(j)$ denote the i^{th} and the j^{th} bins, respectively. More precisely, $\mathcal{B}_1(i)$ is the set of all u^n sequences with assigned index equal to i, and similar is $\mathcal{B}_2(j)$. 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). Moreover, let $\iota_1 : \mathcal{T}_\delta^{(n)}(U) \to [1, 2^{nR_1}]$, and $\iota_2 : \mathcal{T}_\delta^{(n)}(V) \to [1, 2^{nR_2}]$, denote the corresponding random binning functions. With this, we can denote $|l\rangle$ for $l \in [2^{n\tilde{R}_1}]$ as $|l\rangle_L = |\iota_1(l)\rangle_{L_1} |\beta_U(l)\rangle_{L_2}$ and similarly, $|k\rangle$ for $k \in [2^{n\tilde{R}_2}]$ as $|k\rangle_K = |\iota_2(k)\rangle_{K_1} |\beta_V(k)\rangle_{K_2}$[1], where the functions β_U and β_V describe the remaining $\tilde{R}_1 - R_1$ and $\tilde{R}_2 - R_2$ qubits, respectively. Now the qubits in the state $|\iota_1(\cdot)\rangle$ and $|\iota_2(\cdot)\rangle$ are sent over the multiple-access dephasing channel $\mathcal{N}$, each requiring rates of R_1 and R_2 qubits, respectively. Let

$$\sigma^{ABCRLK} \triangleq \mathcal{N}({\Psi_1}^{ABCRLK}).$$

With this, we move on to describing the action of Charlie.

4.4.4 Action of Charlie

Charlie begins by undoing the binning operation. For this, let

$$D_{i,j} \triangleq \big\{(l,k) : (U^n(l), V^n(k)) \in \mathcal{T}_\delta^{(n)}(UV) \text{ and } (U^n(l), V^n(k)) \in \mathcal{B}_1(i) \times \mathcal{B}_2(j)\big\}.$$

For every $i \in [1, 2^{nR_1}]$ and $j \in [1, 2^{nR_2}]$ define the function $F(i,j) = (l,k)$ if (l,k) is the only element of $D_{i,j}$; otherwise $F(i,j) = (0,0)$ Further, $F(i,j) = (0,0)$ for $i = 0$ or $j = 0$.

[1]Note that $\iota_1(l) = \iota_1(U^n(l))$, and similar holds for the functions $\iota_2, \beta_U, \beta_V$.

Using the qubits received from Alice and Bob, and the above definition of $F(i,j)$, Charlie aims at undoing the binning operations. This can be characterized as an isometric map $U_F^C : \mathcal{H}_{Y_1} \otimes \mathcal{H}_{Y_2} \to \mathcal{H}_{Y_1} \otimes \mathcal{H}_{Y_2} \otimes \mathcal{H}_F$ defined as

$$U_F^C \triangleq \sum_{i \in [2^{nR_1}]} \sum_{j \in [2^{nR_2}]} |F(i,j)\rangle \langle i,j| , \tag{4.12}$$

where $F()$ is such that $\dim(\mathcal{H}_F) = R_{tb} \triangleq \tilde{R}_1 - R_1 + \tilde{R}_2 - R_2$. Note that, since binning decreased the total number of qubits transmitted by R_{tb}, to implement the above isometry, Charlie would need R_{tb} number of additional catalytic qubits present in the pure state. As the protocol allows for the use of additional catalysts, as long as they are returned successfully, such an isometry can be implemented as a unitary.

Remark IV.11. As will be shown in the sequel, the error analysis gives an upper bound on R_{tb}. As this is only an upper bound, one can choose to not bin at the maximum rate and can save on the catalytic qubits needed. However, this would increase the communication rates by equivalent factors. This is modelled in the theorem statement using the real number $b \in [0,1]$.

After the complete identification of the measurement outcomes of Alice and Bob, Charlie now extracts the purity from his state, conditioned on these outcomes. For this, he develops a collection of unitary operations $\{U_p^C(l,k)\}_{l \in [2^{n\tilde{R}_1}], k \in [2^{n\tilde{R}_2}]}$, analogous to the earlier ones, with respect to the state $\hat{\sigma}_{l,k}^C$ defined as

$$\hat{\sigma}_{l,k}^C \triangleq \frac{\mathrm{Tr}_{RAB}\left\{ (I^{RC} \otimes \Lambda_l^A \otimes \Lambda_k^B) \Psi_{\rho^{\otimes n}}^{ABCR} (I^{RC} \otimes \Lambda_l^A \otimes \Lambda_k^B)^\dagger \right\}}{\lambda_{l,k}^{AB}}. \tag{4.13}$$

Although $\hat{\sigma}_{l,k}^C$ is defined as an n-letter state, it can be written as $\otimes_{i=1}^n \hat{\sigma}_{u_i,v_i}^C$, where $\hat{\sigma}_{u,v}^C$ is defined correspondingly. Further, he constructs the controlled unitary U_P^{CLK} defined as

$$U_P^{CLK} \triangleq \sum_{l \in [2^{n\tilde{R}_1}]} \sum_{k \in [2^{n\tilde{R}_2}]} U_p^C(l,k) \otimes |l,k\rangle\langle l,k| . \tag{4.14}$$

This characterizes Charlie's unitary as $U_C = U_P^{CLK} U_F^C$, and gives

$$\xi^{ABCRLK} \triangleq (I \otimes U_P^{CLK} U_F^C) \sigma^{ABCRLK} (I \otimes U_P^{CLK} U_F^C)^\dagger.$$

At this point, we have the characterized the actions of all the three parties as unitary operations. The next step is to measure the distance between the obtained state and the desired pure state, and establish the G can be made arbitrary small. Let $\mathbb{1}_{\{sP\}} \triangleq \mathbb{1}_{\{sP-1\}}\mathbb{1}_{\{sP-2\}}$. Using the boundedness of the trace distance, it suffices to show

$$\mathbb{E}[G] \triangleq \mathbb{E}\left[\|\xi^{A_pB_pC_p}\mathbb{1}_{\{sP\}} - |0\rangle\langle 0|^{A_pB_pC_p}\|_1\right] \leq \epsilon,$$

where

$$\xi^{A_pB_pC_p} \triangleq \mathrm{Tr}_{RA_gB_gC_gLK}\{(I^{ABR}\otimes U_C)\mathcal{N}\left(I^{RC}\otimes U_A\otimes U_B)\Psi_{\rho^{\otimes n}}(I^{RC}\otimes U_A\otimes U_B)\right)(I^{RAB}\otimes U_C)^\dagger\}. \tag{4.15}$$

4.4.5 Analysis of Trace Distance

We first provide a proof for the case assuming the encoders do not perform any binning (i.e, $b = 0$), and later incorporate the analysis for the setting when b is non-zero. With this assumption, we define

$$\xi_b^{T_p} \triangleq \mathrm{Tr}_{RT_gLK}\{(I^R\otimes U_P^C\otimes U_A\otimes U_B)\Psi_{\rho^{\otimes n}}^{ABCR}(I^R\otimes U_P^C\otimes U_A\otimes U_B)^\dagger\}, \tag{4.16}$$

where we have replaced U_F^C with an identity transformation, and $T_p \triangleq \mathcal{H}_{A_p}\otimes\mathcal{H}_{B_p}\otimes\mathcal{H}_{C_p}$ and $T_g \triangleq \mathcal{H}_{A_g}\otimes\mathcal{H}_{B_g}\otimes\mathcal{H}_{C_g}$. Now we show the closeness of this state with $\xi_1^{T_p}$ defined as

$$\xi_1^{T_p} \triangleq \mathrm{Tr}_{C_gRLK}\left\{(I^R\otimes U_P^C)\Phi^{RCLK}(I^R\otimes U_P^C)^\dagger\right\}\otimes|0\rangle\langle 0|_{A_pB_p}, \tag{4.17}$$

where

$$\Phi^{RABCLK} \triangleq (I^{RC}\otimes U_A\otimes U_B)\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RC}\otimes U_A\otimes U_B)^\dagger,$$

and $\Phi^{RCLK} \triangleq \mathrm{Tr}_{AB}\left\{\Phi^{RABCLK}\right\}$. Let $\mathbb{1}_{\{sP\}} \triangleq \mathbb{1}_{\{sP-1\}}\cdot\mathbb{1}_{\{sP-2\}}$.

Step 1: Closeness of $\xi_b^{T_p}$ and $\xi_1^{T_p}$: As a first step, we show that $\xi_1^{T_p}$ can be made arbitrary close to $\xi_b^{T_p}$, in trace distance, for sufficiently large n, in the expected sense. Toward this, we

perform the following analysis under the event $\left\{\mathbb{1}_{\{sP\}}=1\right\}$.

$$\begin{aligned}\|\xi_1^{T_p}-\xi_b^{T_p}\|_1 &\leq \left\|\mathrm{Tr}_{A_gB_g}\left\{(I^{RC}\otimes U_A\otimes U_B)\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RC}\otimes U_A\otimes U_B)^\dagger\right\}-\Phi^{RCLK}\otimes|0\rangle\langle 0|_{A_pB_p}\right\|_1\\ &\leq S_A+S_B,\end{aligned} \tag{4.18}$$

where

$$\begin{aligned}S_A &\triangleq \Big\| \mathrm{Tr}_{A_gB_g}\left\{(I^{RC}\otimes U_A\otimes U_B)\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RC}\otimes U_A\otimes U_B)^\dagger\right\}\\ &\qquad -\mathrm{Tr}_{AB_g}\left\{(I^{RC}\otimes U_A\otimes U_B)\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RC}\otimes U_A\otimes U_B)^\dagger\right\}\otimes|0\rangle\langle 0|_{A_p}\Big\|_1,\\ S_B &\triangleq \left\|\mathrm{Tr}_{AB_g}\left\{(I^{RC}\otimes U_A\otimes U_B)\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RC}\otimes U_A\otimes U_B)^\dagger\right\}\otimes|0\rangle\langle 0|_{A_p}-\Phi^{RCLK}\otimes|0\rangle\langle 0|_{A_pB_p}\right\|_1,\end{aligned} \tag{4.19}$$

Now our objective is to show that S_A and S_B can be made arbitrarily small for all sufficiently large n, and sufficiently small $\eta,\delta>0$, in the expected sense, under the event $\left\{\mathbb{1}_{\{sP\}}=1\right\}$. Again using the monotonicity and isometric invariance of trace distance, we obtain

$$\begin{aligned}\mathbb{E}\left[S_A\,\mathbb{1}_{\{sP\}}\right] \leq \mathbb{E}\Big[\Big\| &\mathrm{Tr}_{A_g}\left\{(I^{RBC}\otimes U_A)\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RC}\otimes U_A)^\dagger\right\}\\ &-\mathrm{Tr}_{A}\left\{(I^{RBC}\otimes U_A)\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RBC}\otimes U_A)^\dagger\right\}\otimes|0\rangle\langle 0|_{A_p}\,\mathbb{1}_{\{sP-1\}}\Big\|_1\Big]\leq 4\sqrt{\epsilon},\end{aligned} \tag{4.20}$$

where the last inequality follows by using the results of Lemma IV.9 and Lemma IV.10. Similarly, we have

$$\mathbb{E}\left[S_B\,\mathbb{1}_{\{sP\}}\right]\leq\mathbb{E}\Big[\Big\|\mathrm{Tr}_{B_g}\left\{(I^{RAC}\otimes U_B)\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RAC}\otimes U_B)^\dagger\right\}-\Phi^{RACLK}\otimes|0\rangle\langle 0|_{B_p}\,\mathbb{1}_{\{sP-2\}}\Big\|_1\Big],$$

where $\Phi^{RACLK}\triangleq\mathrm{Tr}_B\left\{\Phi^{RABCLK}\right\}$. Again using the results of Lemma IV.9 and Lemma IV.10, we get $\mathbb{E}\left[S_B\,\mathbb{1}_{\{sP\}}\right]\leq 4\sqrt{\epsilon}$. This implies, $\mathbb{E}\left[\|\xi_1^{T_p}-\xi_b^{T_p}\|_1\mathbb{1}_{\{sP\}}\right]\leq 8\sqrt{\epsilon}$ for all sufficiently large n and

sufficiently small $\eta, \delta > 0$, while obtaining purity that satisfies

$$\frac{1}{n}\log\dim\mathcal{H}_{A_p} < \log\dim\mathcal{H}_A - \sum_u \lambda_u^A S(\rho_u^A), \quad \frac{1}{n}\log\dim\mathcal{H}_{B_p} < \log\dim\mathcal{H}_B - \sum_v \lambda_v^A S(\rho_v^B), \tag{4.21}$$

With the above result, we now move on to the next step. For this define,

$$\xi_2^{T_p} \triangleq \mathrm{Tr}_{RABC_gLK}\{(I^R \otimes U_P^C \otimes U_A' \otimes U_B')\Psi_{\rho^{\otimes n}}(I^R \otimes U_P^C \otimes U_A' \otimes U_B')^\dagger\} \otimes |0\rangle\langle 0|_{A_pB_p},$$

where

$$U_A' \triangleq \sum_{l\in[2^{n\tilde{R}_1}]} U_p^A(l)\sqrt{\frac{\gamma}{\lambda_l^A}}\sqrt{\Lambda_l^A} \otimes |l\rangle, \quad U_B' \triangleq \sum_{k\in[2^{n\tilde{R}_2}]} U_p^B(k)\sqrt{\frac{\zeta}{\lambda_k^B}}\sqrt{\Lambda_k^B} \otimes |k\rangle.$$

Note that, although U_A' and U_B' are not unitary operators, they are introduced for analysis purposes. In other words, the non-product unitary corresponding to the coherent measurement U_M^A followed by the n-letter rotation operator U_R^A is replaced by its corresponding product measurement, and similarly for U_M^B. In the next step, it will become clear how this operator is used in the analysis.

Step 2: Closeness of $\xi_1^{T_p}$ and $\xi_2^{T_p}$: consider the following set of inequalities:

$$\begin{aligned}
&\|\xi_2^{T_p} - \xi_1^{T_p}\|_1 \mathbb{1}_{\{sP\}} \\
&\quad\le \Big\| \mathrm{Tr}_{ABLK}\Big\{(I^{RAB}\otimes U_P^C)\Big[(I^{RC}\otimes U_A\otimes U_B)\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RC}\otimes U_A\otimes U_B) \\
&\qquad\qquad - (I^{RC}\otimes U_A'\otimes U_B')\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RC}\otimes U_A'\otimes U_B')\Big](I^{RAB}\otimes U_P^C)^\dagger\Big\}\Big\|_1 \mathbb{1}_{\{sP\}} \\
&\quad\le \sum_{l,k}\Big\|(I^R\otimes U_p^C(l,k))\Big[\mathrm{Tr}_{AB}\Big\{(I^{RC}\otimes\sqrt{A_l}\otimes\sqrt{B_k})\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RC}\otimes\sqrt{A_l}\otimes\sqrt{B_k})^\dagger\Big\} \\
&\qquad - \frac{\gamma\zeta}{\lambda_l^A\lambda_k^B}\mathrm{Tr}_{AB}\Big\{(I^{RC}\otimes\sqrt{\Lambda_l^A}\otimes\sqrt{\Lambda_k^B})\Psi_{\rho^{\otimes n}}^{ABCR}(I^{RC}\otimes\sqrt{\Lambda_l^A}\otimes\sqrt{\Lambda_k^B})^\dagger\Big\}\Big](I^R\otimes U_p^C(l,k))^\dagger\Big\|_1 \\
&\quad= \sum_{l,k}\Big\|\sqrt{\rho_{AB}^{\otimes n}}(A_l\otimes B_k)\sqrt{\rho_{AB}^{\otimes n}} - \frac{\gamma\zeta}{\lambda_l^A\lambda_k^B}\sqrt{\rho_{AB}^{\otimes n}}(\Lambda_l^A\otimes\Lambda_k^B)\sqrt{\rho_{AB}^{\otimes n}}\Big\|_1 \le T_A + T_B,
\end{aligned} \tag{4.22}$$

where

$$T_A \triangleq \sum_{l,k} \left\| \sqrt{\rho_{AB}^{\otimes n}}(A_l \otimes B_k)\sqrt{\rho_{AB}^{\otimes n}} - \sqrt{\rho_{AB}^{\otimes n}} \left(A_l \otimes \frac{\zeta}{\lambda_k^B}\Lambda_k^B \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1,$$
$$T_B \triangleq \sum_{l,k} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(A_l \otimes \frac{\zeta}{\lambda_k^B}\Lambda_k^B \right) \sqrt{\rho_{AB}^{\otimes n}} - \frac{\gamma\zeta}{\lambda_l^A\lambda_k^B}\sqrt{\rho_{AB}^{\otimes n}}(\Lambda_l^A \otimes \Lambda_k^B)\sqrt{\rho_{AB}^{\otimes n}} \right\|_1,$$

To bound T_A and T_B, we use the following lemma from Chapter II.

Lemma IV.12 (Lemma II.11 Chapter II)**.** *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 Γ^A acting on $\mathcal{H}_A$, we have*

$$\sum_{y\in\mathcal{Y}} \left\| \sqrt{\rho_{AB}} \left(\Gamma^A \otimes \Lambda_y^B \right) \sqrt{\rho_{AB}} \right\|_1 \leq \left\| \sqrt{\rho_A}\, \Gamma^A \sqrt{\rho_A} \right\|_1, \tag{4.23}$$

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

Using Lemma IV.12, we can now simplify T_A and T_B as

$$\mathbb{E}[T_A] \leq \mathbb{E}\left[\sum_k \left\| \sqrt{\rho_B^{\otimes n}} B_k \sqrt{\rho_B^{\otimes n}} - \sqrt{\rho_B^{\otimes n}} \left(\frac{\zeta}{\lambda_k^B}\Lambda_k^B \right) \sqrt{\rho_B^{\otimes n}} \right\|_1 \right] = \mathbb{E}\left[\sum_k \zeta \|\tilde{\rho}_l^A - \hat{\rho}_l^A\|_1 \right] \leq \epsilon, \tag{4.24}$$

for all sufficiently large n, and all sufficiently small $\eta, \delta > 0$, where the last inequality follows from (4.5). For T_B, we obtain

$$\begin{aligned}
&\mathbb{E}[T_B] \\
&= \mathbb{E}_A\left[\sum_{l,k} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \mathbb{E}_B[\mathbb{1}_{\{V^n(k)=v^n\}}] \frac{\zeta}{\lambda_{v^n}^B} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(A_l \otimes \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} - \frac{\gamma}{\lambda_l^A} \sqrt{\rho_{AB}^{\otimes n}} (\Lambda_l^A \otimes \Lambda_{v^n}^B) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \right] \\
&\leq \frac{1}{(1+\eta)} \mathbb{E}_A\left[\sum_l \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(A_l \otimes \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} - \frac{\gamma}{\lambda_l^A} \sqrt{\rho_{AB}^{\otimes n}} (\Lambda_l^A \otimes \Lambda_{v^n}^B) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \right] \\
&\leq \frac{1}{(1+\eta)} \mathbb{E}_A\left[\sum_l \left\| \sqrt{\rho_A^{\otimes n}} A_l \sqrt{\rho_A^{\otimes n}} - \frac{\gamma}{\lambda_l^A} \sqrt{\rho_A^{\otimes n}} \Lambda_l^A \sqrt{\rho_A^{\otimes n}} \right\|_1 \right] \\
&\leq \frac{1}{(1+\eta)} \mathbb{E}_A\left[\sum_l \gamma \left\| \tilde{\rho}_l^A - \hat{\rho}_l^A \right\|_1 \right] \leq \epsilon,
\end{aligned} \tag{4.25}$$

for all sufficiently large n, and all sufficiently small $\eta, \delta > 0$, where the last inequality follows from (4.5). This implies, $\mathbb{E}[\|\xi_2^{T_p} - \xi_1^{T_p}\|_1 \mathbb{1}_{\{sP\}}] \leq 2\epsilon$, for all sufficiently large n, and all sufficiently small $\eta, \delta > 0$.

Step 3: Closeness of $\xi_2^{T_p}$ and $|0\rangle\langle 0|_{T_p}$:

$$\|\xi_2^{T_p} - |0\rangle\langle 0|_{T_p}\|_1 = \left\| \sum_{l,k} \frac{\gamma\zeta}{\lambda_l^A \lambda_k^B} \lambda_{l,k}^{AB} \operatorname{Tr}_{C_p} \left\{ U_p^C(l,k) \hat{\sigma}_{l,k}^C (U_p^C(l,k))^\dagger \right\} - |0\rangle\langle 0|_{C_p} \right\|_1 \leq Q_1 + Q_2, \tag{4.26}$$

where $\hat{\sigma}_{l,k}^C$ is as defined in 4.13, and

$$Q_1 \triangleq \sum_{l,k} \frac{\gamma\zeta}{\lambda_l^A \lambda_k^B} \lambda_{l,k}^{AB} \left\| \operatorname{Tr}_{C_p} \left\{ U_p^C(l,k) \hat{\sigma}_{l,k}^C (U_p^C(l,k))^\dagger \right\} - |0\rangle\langle 0|_{C_p} \right\|_1,$$
$$Q_2 \triangleq \left\| \sum_{l,k} \frac{\gamma\zeta}{\lambda_l^A \lambda_k^B} \lambda_{l,k}^{AB} |0\rangle\langle 0|_{C_p} - |0\rangle\langle 0|_{C_p} \right\|_1 = \left| \sum_{l,k} \frac{\gamma\zeta}{\lambda_l^A \lambda_k^B} \lambda_{l,k}^{AB} - 1 \right|.$$

From the definition of $U_p^C(l,k)$, it follows that

$$\begin{aligned} \mathbb{E}[Q_1] \leq & \sum_{(u^n, v^n) \in \mathcal{T}_\delta^{(n)}(U,V)} \frac{(1-\varepsilon_1)(1-\varepsilon_2)}{(1+\eta)^2} \lambda_{u^n v^n}^{AB} \left\| \operatorname{Tr}_{C_p} \left\{ U_p^C(u^n, v^n) \hat{\sigma}_{u^n, v^n}^C (U_p^C(u^n, v^n))^\dagger \right\} - |0\rangle\langle 0|_{C_p} \right\|_1 \\ & + 2 \sum_{(u^n, v^n) \notin \mathcal{T}_\delta^{(n)}(U,V)} \lambda_{u^n v^n}^{AB} \\ \leq & 3\epsilon, \end{aligned} \tag{4.27}$$

for all sufficiently large n and sufficiently small $\eta, \delta > 0$, and for $\frac{1}{n} \log \dim \mathcal{H}_{C_p} < \log \dim \mathcal{H}_C - \sum_{u,v} \lambda_{uv}^{AB} S(\hat{\sigma}_{uv}^C)$. For the term corresponding to Q_2, we provide the following proposition.

Proposition IV.13. *For any $\epsilon \in (0,1)$, any $\eta \in (0,1)$, any $\delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[Q_2] \leq \epsilon$, if $\tilde{R}_1 + \tilde{R}_2 > I(U;V)$.*

Proof. The proof follows from using a similar set of arguments developed in (*Cuff*, 2009, Lemma 19). □

This implies, $\mathbb{E}\left[\|\xi_2^{T_p} - |0\rangle\langle 0|_{T_p}\|_1\right] \leq 4\epsilon$, for any given $\epsilon \in (0,1)$, and for sufficiently large n, and sufficiently small $\eta, \delta > 0$.

Now as a final step, we consider the case when Alice and Bob chooses to bin their measurement outcomes before sending over the dephasing channel, i.e., the case when $b > 0$. We term the error introduced by this process as the binning error.

Step 4: Closeness of $\xi_b^{T_p}$ and ξ^{T_p}: In this step, we bound the error that is introduced when Charlie tries to undo the binning operation by performing the unitary U_F. We show that Charlie will be successful if the rate at which binning is performed is constrained by a non-trivial bound (to be obtained in Proposition IV.14), and hence the error involved in undoing the binning operation can be made arbitrary small, in an expected sense, for all sufficiently large n.

For $l \in l \in [2^{n\tilde{R}_1}]$ and $k \in k \in [2^{n\tilde{R}_2}]$, define $d(l,k) \triangleq F(i,j)$, such that $(U^n(l), V^n(k)) \in \mathcal{B}_1(i) \times \mathcal{B}_2(j)$. Note that $d(\cdot,\cdot)$ captures the overall effect of the binning followed by the decoding function F. Further, for $l \in l \in [2^{n\tilde{R}_1}]$ and $k \in k \in [2^{n\tilde{R}_2}]$, define

$$\tilde{\Psi}_{\rho^{\otimes n}} \triangleq \frac{(I^{RC} \otimes \sqrt{A_l \otimes B_k})\Psi_{\rho^{\otimes n}}(I^{RC} \otimes \sqrt{A_l \otimes B_k})}{\gamma_l \zeta_k}.$$

Using these definitions, we obtain

$$\begin{aligned}
\|\xi^{T_p} - \xi_b^{T_p}\|_1 \leq \sum_{l,k} \gamma_l \zeta_k \Big\| &(I^R \otimes U_p^A(l) U_r^A(l) \otimes U_B(k) U_r^B(k)) \Big(U_p^C(l,k) \tilde{\Psi}_{\rho^{\otimes n}} (U_p^C(l,k))^\dagger \\
&- U_p^C(d(l,k)) \tilde{\Psi}_{\rho^{\otimes n}} \otimes U_p^C(d(l,k))^\dagger \Big) (I^R \otimes U_p^A(l) U_r^A(l) \otimes U_B(k) U_r^B(k))^\dagger \Big\| \\
= \sum_{l,k} \gamma_l \zeta_k \Big\| &U_p^C(l,k) \tilde{\Psi}_{\rho^{\otimes n}} (U_p^C(l,k))^\dagger - U_p^C(d(l,k)) \tilde{\Psi}_{\rho^{\otimes n}} \otimes U_p^C(d(l,k))^\dagger \Big\|. \qquad (4.28)
\end{aligned}$$

Now, consider the following proposition.

Proposition IV.14. *For any $\epsilon \in (0,1)$, and sufficiently large n and sufficiently small $\eta, \delta > 0$, we have $\mathbb{E}[\|\xi^{T_p} - \xi_b^{T_p}\|_1] \leq \epsilon$ if $\tilde{R}_1 - R_1 + \tilde{R}_2 - R_2 \geq I(U;V)_{\sigma_3}$.*

Proof. The proof is provided in Appendix C.4. □

Finally, we complete the proof by combining the results from all the above steps in the following. Using this, we have

$$\|\xi^{T_p} \mathbb{1}_{\{sP\}} - |0\rangle\langle 0|^{T_p}\|_1 \leq \Big[\|\xi^{T_p} - \xi_b^{T_p}\|_1 + \|\xi_b^{T_p} - \xi_1^{T_p}\|_1 + \|\xi_1^{T_p} - \xi_2^{T_p}\|_1 \Big] \mathbb{1}_{\{sP\}} + \|\xi_2^{T_p} - |0\rangle\langle 0|^{T_p}\|_1,$$

Taking expectation of the above inequality and using (i) the closeness of trace norm proved in each of the steps, and (ii) the result from Proposition IV.7, we have the desired result. This completes the proof.

4.5 Conclusion

In this chapter, we considered the task of distilling local purity from a noisy quantum state ρ^{ABC}. We provided a protocol for three parties, Alice, Bob and Charlie, to distill local purity (at a rate P) from many independent copies of a given quantum state ρ^{ABC}. The three parties have access to their respective subsystems of ρ^{ABC}, and are provided with pure ancilla catalytically, i.e., with the promise of returning them unaltered after the end of the protocol. In addition, Alice and Bob can communicate with Charlie using a one-way multiple-access dephasing channel of link rates R_1 and R_2, respectively. The objective of the protocol was to minimize the usage of the dephasing channel (in terms of rates R_1 and R_2) while maximizing the asymptotic purity that can be jointly distilled from ρ^{ABC}. To achieve this, we employed ideas from distributed measurement compression protocols, and in turn, characterized a set of sufficient conditions on (P, R_1, R_2) in terms of quantum information theoretic quantities such that P amount of purity can be distilled using rates R_1 and R_2. Moreover, we observed a continuous trade-off between communication rate and purity, expressed in terms of the variable b. This variable captures the correlations present in the measurement outcomes across the distributed subsystems. We provided a way to exploit these correlations by a technique similar to classical binning, undoing which results in a tradeoff between distillable purity and communication rates.

CHAPTER V

Lossy Quantum Source Coding

5.1 Introduction

In Chapters II and III, apart from our objective of measurement compression, we also obtained QC distributed compression results. In this chapter, we turn our attention to a fully quantum problem. A fundamental problem from an quantum information theoretic perspective is the asymptotic characterization of the rate required to compress a quantum source that can be recovered to a certain measurable degree. Such a problem is referred to as quantum source coding or a quantum data compression problem. In the lossless regime, Schumacher *Schumacher* (1995); *Jozsa and Schumacher* (1994) proved that a quantum source could be compressed at a rate given by von Neumann entropy while incurring a very small error between the reconstruction and the source state. The error in this model is defined for the entire block, also called as block error or *global error*. Considering the block error, a strong converse was also proved in the lossless regime *Winter* (1999), which states that it is impossible to achieve any rate below von Neumann entropy even when the asymptotic probability of block error is relaxed from being (almost) zero.

As for the lossy regime, where the objective is to further reduce the rate at the expense of increased but bounded error, Barnum *Barnum* (2000) conjectured minimal coherent information as a candidate in characterizing the asymptotic performance limit. Generalizing the formulation from the classical rate-distortion theory *Shannon et al.* (1959), Barnum in *Barnum* (2000) introduced a *local* distortion criterion as averaged symbol-wise entanglement fidelity based on marginal operations (partial trace) between the reconstruction and the reference of the original source. In *Datta et al.* (2013a), Datta *et. al* obtained a regularized expression for the quantum rate-distortion

distortion function in terms of the entanglement of purification. Further, the authors also formulated the entanglement-assisted quantum rate-distortion problem and characterized its asymptotic performance limit using a single-letter expression. Wilde *et. al* further refined the characterization of the quantum rate-distortion function in terms of regularized entanglement of formation, and also generalized the problem setup to various scenarios, including side information in *Wilde et al.* (2013). Works toward the asymptotic simulation of a memoryless quantum channel in *Berta et al.* (2011); *Bennett et al.* (2014) have shown to be useful in achieving the above results, in particular, the entanglement-assisted formulations. Authors in *Datta et al.* (2013b) formulated a quantum-to-classical rate-distortion problem and provided a single-letter formula. A rate-distortion version of the quantum state redistribution task *Devetak and Yard* (2008); *Luo and Devetak* (2009) was considered in *Khanian and Winter* (2021). Investigations on a rate-distortion framework of generic mixed quantum sources have been the focus of *Khanian and Winter* (2022); *Baghali Khanian et al.* (2022). Other works that addressed related problems include *Koashi and Imoto* (2001); *Devetak and Berger* (2002); *Winter* (2002); *Datta et al.* (2013c); *Hsieh and Watanabe* (2016); *Salek et al.* (2018); *Anshu et al.* (2019).

In this work, we consider a new formulation of the problem of lossy quantum source coding, and characterize a rate function, no larger than von Neumann entropy, while allowing for bounded error in the reconstruction. We use a global error criterion as opposed to the approach of local symbol-wise error studied in the literature. The problem we consider is without any shared entanglement resources between the encoder and the decoder. We motivate this formulation with the following observations.

An important component of any source coding problem is the error criterion employed in the formulation of the problem. The local error criterion in the quantum rate-distortion framework is inspired by the corresponding additive local single-letter distortion criterion in the classical source coding formulation of Shannon *Shannon et al.* (1959), where a single-letter characterization is available. The motivation for considering a local criterion is the strong converse of the lossless source coding theorem which states that the entropy bound cannot be breached even when the asymptotic probability of block error is relaxed to any number in $(0, 1)$ (*Csiszár and Körner*, 2011, Theorem 1.1). Although the strong converse precludes any relaxation of probability of block error criterion, it does not prevent adopting alternative global error definitions for formulating a lossy

source coding problem.

In *Shannon et al.* (1959); *Csiszár and Körner* (2011), a duality connection between the source coding problem and the channel coding problem was observed. These problems were interpreted in terms of a covering versus packing perspective. In both problems, the same information measure, namely the mutual information, captures the asymptotic performance limits. A similar duality connection exists between the classical-quantum communication problem *Holevo* (1998); *Schumacher and Westmoreland* (1997) and the quantum-classical source coding problem *Winter* (2004); *Datta et al.* (2013b), with the performance limits of both problems characterized in terms of single-letter Holevo information quantities *Holevo* (2019). This has been further explored in *Cheng et al.* (2019). In the fully quantum setting, from this standpoint, its well known that the quantum channel coding problem has an asymptotic performance limit characterized using regularized coherent information *Lloyd* (1997); *Shor* (2002); *Devetak* (2005b); *Hayden et al.* (2008). Among others, Devetak developed a proof of this result by employing a coherent approach to covering and packing, and combined them cohesively, inspired by his work on the private channel capacity problem *Devetak* (2005b). Coherent information can be interpreted in terms of packing of subspaces as elucidated in *Lloyd* (1997). Quantum error-correcting codes have been extensively studied along these lines in the coding theory literature, e.g., quantum Hamming bound *Nielsen and Chuang* (2002). This leads us to the question: why is such a limit based on coherent information absent for the lossy quantum source compression problem?

Toward answering this question, we take a closer look at the classical discrete memoryless setting. We find that in addition to Shannon's pioneering work of characterizing the rate-distortion problem *Shannon et al.* (1959); *Berger* (1975), there have been several works discussing the lossy source compression problem. A concept that has received particular attention is the notion of a backward channel (*Csiszár and Körner*, 2011, Problem 8.3), which characterizes the posterior distribution of the source given the reconstruction. The structure of this channel has been studied in *Gallager* (1968); *Berger* (1971); *Gerrish* (1963). Although the forward channel, relating the reconstruction to the source, achieving the rate-distortion function need not be unique, the resulting backward channel is indeed unique. Moreover, the rate-distortion achievability result in (*Csiszár and Körner*, 2011, Theorem 2.3) is shown by constructing a channel code for a backward channel with a large probability of error and by using the encoder of the latter as a decoder of the former

and vice versa. Highlighting this duality further, inspired by results on the output statistics of good channel codes *Shamai and Verdú* (1997), the following was shown in *Pradhan* (2004). The n-letter actual posterior conditional distribution of the source vector given the reconstruction vector of any rate-distortion achieving code converges in normalized divergence to the n-product of the unique minimum-mutual-information backward channel conditional distribution. In other words, although the encoder and decoder are block operations, the induced posterior n-letter channel becomes discrete memoryless in the asymptotic limit for a rate-distortion achieving code. For further developments on this concept see *Weissman and Ordentlich* (2005); *Cuff et al.* (2010); *Schieler and Cuff* (2013); *Kostina and Verdú* (2015). This channel also plays a fundamental role in Bayesian estimation and detection theory *Poor* (1998), e.g., maximum a posteriori (MAP) estimation. Therefore, we ask the question, can we use such a channel to formulate a lossy source coding problem?

Contributions of this work: In light of this, in this chapter, we explore a new formulation of the source compression problem in the memoryless setting. This formulation is based on the notion of a posterior channel that produces the reference of the source from that of the reconstruction. Instead of a single-letter distortion function, now, we are given a single-letter posterior channel that characterizes the nature of the loss incurred in the encoding and decoding operations. More precisely, we want to construct an encoder and a decoder such that the joint effect of encoding and decoding – to produce a reconstruction sequence from the source sequence – is close to the effect of the n-product posterior channel acting on the non-product reconstruction sequence.The closeness is measured using the trace distance in the quantum case and the total variation in the classical case, manifesting as a global error constraint. A related concept is the Petz recovery map which has found significant relevance in information-theoretic problems *Petz* (1986); *Barnum and Knill* (2002); *Hayden et al.* (2004). However, we take a different approach and consider a quantum channel, i.e., a CPTP map, acting on the reference of the reconstruction to produce the reference of the source, whose existence is guaranteed using Uhlmann's theorem. We refer to this as a posterior reference map.

As one of the main contributions of our work, we provide a single-letter characterization of the asymptotic performance limit of this source coding problem using the minimal coherent information of the posterior reference map, where the minimization is over all reconstructions (see Theorem

V.9). Furthermore, our work establishes a duality connection between quantum lossy compression and the quantum channel coding problem. Our proof is based on the coherent application of two fundamental tools of quantum information theory, namely, packing and covering, implying a duality relationship with Devetak's proof for the channel coding problem *Devetak* (2005b) (also see *Shor* (2002); *Hayden et al.* (2008)).

At one end of the spectrum, when the posterior reference map is specified as the identity transformation, our rate expression in the quantum case reduces to the von Neumann entropy of the given quantum source, demonstrating the connection with the Schumacher's lossless compression *Schumacher* (1995). In fact, with Schumacher's formulation also based on a global error criterion, the latter and the current formulations enjoys a stronger relationship of equivalence at this extreme. On the other end, when the specified posterior reference map is such that coherent information is negative for some reference of the reconstruction, we characterize the asymptotic performance limit of the lossy quantum source coding problem to be zero.

The techniques employed to prove our results can be summarized as follows. For the achievability of the Theorem V.9, we first construct a posterior reference isometry V (as in Definition V.1) and decompose it as a coherent measurement. We then make use of Winter's measurement compression protocol *Winter* (2004), and apply it in a coherent fashion to compress the output of the above isometry. This involves using the Uhlmann's Theorem *Uhlmann* (1976) (or (*Wilde*, 2011, Theorem 9.2.1)) followed by incorporating additional phases to achieve a coherent faithful simulation of the posterior reference map. To further decrease the compression rate, we exploit the fact that a noiseless quantum channel can preserve arbitrary superpositions. Therefore, we perform additional encoding to embed the information at the output of V as superpositions within itself. This requires availing the HSW classical communication result *Holevo* (1998); *Schumacher and Westmoreland* (1997) to construct information decoding POVMs, and Naimark's extension theorem to construct a unitary from POVM elements. The method used for expurgation is another interesting feature of the proof. The protocol as it stands only permits operations that are unitary or isometric, followed by partial tracing. It can be challenging to guarantee this when there are repeated codewords in a code. A similar phenomenon was observed in the Devetak's proof *Devetak* (2005b). For the converse of Theorem V.9, we use the quantum data processing inequality for coherent information, the Fannes-Audenart inequality, and monotonicity results. such as the quan-

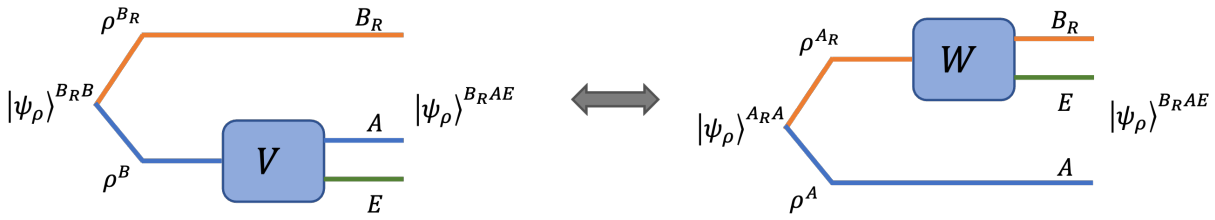

Figure 5.1: Figure demonstrating the construction of the posterior reference map W from the isometry V (the Stinespring's dilation of $\mathcal{N}_V$) and the source state ρ^B.

tum data processing inequality, the concavity of conditional quantum entropy, and the continuity of quantum mutual information (AFW inequality).

The chapter is organized as follows. We provide some necessary definitions and useful lemmas in Section 5.2. In Section 5.3, we formulate the problems and provide the main results (Theorem V.9). We provide examples corresponding to the main result in section 5.4. In Sections 5.6 and 5.7 we provide proofs of the main results. Finally, Section 5.8 concludes the chapter.

5.2 Preliminaries and Notations

We supplement our notations so far with the following. Let I_A denote the identity operator acting on a Hilbert space $\mathcal{H}_A$. The set of density operators on $\mathcal{H}_A$ are denoted by $\mathcal{D}(\mathcal{H}_A)$, and linear operators by $\mathcal{L}(\mathcal{H}_A)$. We denote $\mathcal{H}_{A_R}$ as the Hilbert space associated with the reference space of $\mathcal{H}_A$, with $\dim \mathcal{H}_{A_R} = \dim \mathcal{H}_A$. In this work, we focus exclusively on references obtained from canonical purifications of quantum states (*Winter*, 2004, Lemma 14 (Pretty Good Purifications)), and define canonical purification $|\psi_\rho\rangle^{A_R A}$ of ρ^A as $|\psi_\rho\rangle^{A_R A} \triangleq (I_{A_R} \otimes \sqrt{\rho^A})\Gamma_{A_R A}$, where $\Gamma_{A_R A}$ is defined as the unnormalized maximally entangled state. We use $\Psi_\rho^{A_R A}$ to denote the density operator corresponding to $|\psi_\rho\rangle^{A_R A}$. As is the convention, for two states acting on the same Hilbert space, we use the same Γ when defining their canonical purifications. We denote the finite alphabet of a source as X, and the set of probability distributions on the finite alphabet X as $\mathcal{P}(\mathsf{X})$. Let $[\Theta] \triangleq \{1, 2, \cdots, \Theta\}$. For a CPTP map $\mathcal{N} : \mathcal{H}_A \to \mathcal{H}_B$, and an input density operator $\rho^A \in \mathcal{D}(\mathcal{H}_A)$, we use $I_c(\mathcal{N}, \rho^A)$ to denote the coherent information of $\mathcal{N}$ with respect to ρ^A.

Definition V.1 (Posterior Reference Map)**.** Given a source $\rho^B \in \mathcal{D}(\mathcal{H}_B)$ and a channel $\mathcal{N}_V : \mathcal{H}_B \to \mathcal{H}_A$, let $\rho^A \triangleq \mathcal{N}_V(\rho^B)$. Let $V : \mathcal{H}_B \to \mathcal{H}_A \otimes \mathcal{H}_E$ be a Stinespring's isometry corresponding

to the CPTP map $\mathcal{N}_V$ with $\dim(\mathcal{H}_E) \geq \dim(\mathcal{H}_A)$, such that $\mathcal{N}_V(\cdot) = \mathrm{Tr}_E\{V(\cdot)V^\dagger\}$. As shown in Figure 5.1, define the "posterior reference map" of V with respect to ρ^A as the CPTP map $\mathcal{N}_W : \mathcal{H}_{A_R} \to \mathcal{H}_{B_R}$ corresponding to the isometry $W : \mathcal{H}_{A_R} \to \mathcal{H}_{B_R} \otimes \mathcal{H}_E$ satisfying $(W \otimes I_A) |\psi_\rho\rangle^{A_R A} = (I_{B_R} \otimes V) |\psi_\rho\rangle^{B_R B}$ where $|\psi_\rho\rangle^{A_R A}$ and $|\psi_\rho\rangle^{B_R B}$ are the canonical purifications of ρ^A and ρ^B, respectively.

Remark V.2 (Existence of a Posterior Reference Map). Using the equivalence of purifications, one can guarantee the existence of such a posterior reference isometry $W : \mathcal{H}_{A_R} \to \mathcal{H}_{B_R} \otimes \mathcal{H}_E$. Since V is an isometry with $\dim(\mathcal{H}_E) \geq \dim(\mathcal{H}_A)$, and since $|\psi_\rho\rangle^{A_R A}$ and $|\psi_\rho\rangle^{B_R AE} \triangleq (I_{B_R} \otimes V) |\psi_\rho\rangle^{B_R B}$ are purifications of ρ^A (as $\mathrm{Tr}_E(V\rho^B V^\dagger) = \rho^A$), from (*Wilde*, 2011, Theorem 5.1.1), there exists an isometry $W : \mathcal{H}_{A_R} \to \mathcal{H}_{B_R} \otimes \mathcal{H}_E$ such that $(W \otimes I_A) |\psi_\rho\rangle^{A_R A} = |\psi_\rho\rangle^{B_R AE}$.

5.2.1 Useful Lemmas

Lemma V.3 (*Fuchs and Van De Graaf* (1999), Theorem 9.3.1 *Wilde* (2011))**.** *Given two states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, we have*

$$1 - \sqrt{F(\rho,\sigma)} \leq \frac{1}{2}\|\rho - \sigma\|_1 \leq \sqrt{1 - F(\rho,\sigma)}.$$

Lemma V.4. *For $\rho^B, \sigma^B \in \mathcal{D}(\mathcal{H}_B)$, the following inequality holds:*

$$F(|\psi_\rho\rangle, |\psi_\sigma\rangle) \geq \left(1 - \frac{1}{2}\left\|\rho^B - \sigma^B\right\|_1\right)^2, \tag{5.1}$$

where $|\psi_\rho\rangle$ and $|\psi_\sigma\rangle$ are the canonical purifications of ρ^B and σ^B, respectively.

Proof. We provide a proof in Appendix D.1. □

The above lemma is a slight tightening of the Lemma 14 ("Pretty good purifications") of *Winter* (2004).

Lemma V.5 (Naimark's extension theorem *Naimark* (1940), (*Wilde*, 2013b, Theorem 2.1))**.** *Given a POVM $\{\Gamma_x\}_{x\in\mathsf{X}}$ acting on the system $\mathcal{H}_A$, there exists a unitary $U_{AA'}$ acting on the system $\mathcal{H}_A$ and auxiliary system $\mathcal{H}_{A'}$ and an orthonormal basis $\{|x\rangle^{A'}\}_{x\in\mathsf{X}}$ such that*

$$\mathrm{Tr}\left\{\overline{\Gamma}_x(\rho^A \otimes |0\rangle\langle 0|_{A'})\right\} = \mathrm{Tr}\big(\Gamma_x \rho^A\big),$$

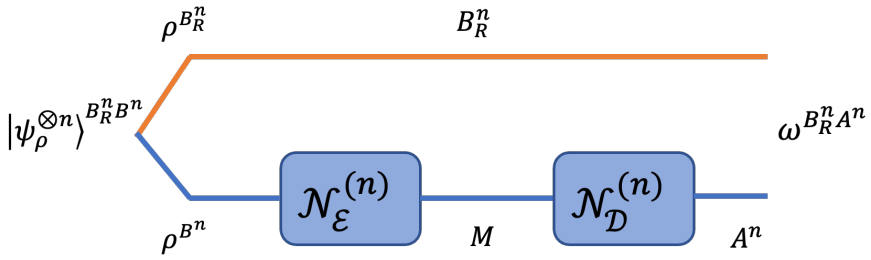

Figure 5.2: Illustration of Lossy Quantum Compression protocol

where $\{\overline{\Gamma}_x \triangleq U_{AA'}^\dagger(I_A \otimes |x\rangle\langle x|^{A'})U_{AA'}\}$ *are orthogonal projectors acting on system* $\mathcal{H}_A \otimes \mathcal{H}_{A'}$*. And,* $|0\rangle^{A'}$ *is some fixed state in* $\mathcal{H}_{A'}$*, independent of* Γ_x *and* ρ^A*.*

5.3 Main Result: Characterization of the Achievable Rate Region

We first formulate a quantum source coding problem as follows. For any memoryless quantum information source, characterized by $\rho^B \in \mathcal{D}(\mathcal{H}_B)$, denote its canonical purification by $|\psi_\rho\rangle^{BB_R}$. Let $\rho^{B_R} \triangleq \mathrm{Tr}_B[\Psi_\rho^{B_R B}]$.

Definition V.6 (Quantum Source Coding Setup)**.** A quantum source coding setup is characterized by a triple $(\rho^B, \mathcal{H}_A, \mathcal{N}_W)$, where $\rho^B \in \mathcal{D}(\mathcal{H}_B)$ is a density operator, $\mathcal{H}_A$ is a reconstruction Hilbert space, and $\mathcal{N}_W$ is a single-letter CPTP map from $\mathcal{H}_{A_R}$ to $\mathcal{H}_{B_R}$, where $\mathcal{H}_{A_R}$ and $\mathcal{H}_{B_R}$ are reference spaces corresponding to $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively.

Definition V.7 (Lossy Quantum Compression Protocol)**.** For a given input and reconstruction Hilbert spaces $(\mathcal{H}_B, \mathcal{H}_A)$, an (n, Θ) lossy quantum compression protocol consists of a encoding CPTP map $\mathcal{N}_{\mathcal{E}}^{(n)} : \mathcal{H}_{B^n} \to \mathcal{H}_M$ and a decoding CPTP map $\mathcal{N}_{\mathcal{D}}^{(n)} : \mathcal{H}_M \to \mathcal{H}_{A^n}$, such that $\dim(\mathcal{H}_M) = \Theta$, as shown in Figure 5.2.

Definition V.8 (Achievability)**.** For a quantum source coding setup $(\rho^B, \mathcal{H}_A, \mathcal{N}_W)$, a rate R is said to be achievable, if for all $\epsilon > 0$ and all sufficiently large n, there exists an (n, Θ) lossy quantum compression protocol satisfying

$$\left\| \omega^{B_R^n A^n} - (\mathcal{N}_W^{\otimes n} \otimes I_{A^n}) \Psi_\omega^{A_R^n A^n} \right\|_1 \leq \epsilon, \tag{5.2}$$

and $\frac{1}{n}\log\Theta \leq R+\epsilon$, where $\omega^{B_R^n A^n} \triangleq (I\otimes\mathcal{N}_{\mathcal{D}}^{(n)})(I\otimes\mathcal{N}_{\mathcal{E}}^{(n)})(\Psi_\rho^{B_R^n B^n})$, and $\Psi_\rho^{B_R^n B^n}$ and $\Psi_\omega^{A_R^n A^n}$ are the canonical purifications of $\rho^{B^{\otimes n}}$ and ω^{A^n}, respectively.

In other words, the protocol ensures that the joint state of the reconstruction on $\mathcal{H}_A^{\otimes n}$ and the original reference $\mathcal{H}_{B_R}^{\otimes n}$ is close to the effect of the n-product posterior channel acting on the reference of the non-product reconstruction sequence. Our objective is to characterize the set of all achievable rates using single-letter quantum information quantities.

Theorem V.9 (Lossy Quantum Compression Theorem)**.** *For a $(\rho^B,\mathcal{H}_A,\mathcal{N}_W)$ quantum source coding setup, a rate R is achievable if and only if $S(\rho^B,\mathcal{N}_W)$ is non empty, and*

$$R \geq \min_{\rho^{A_R}\in\mathcal{S}(\rho^B,\mathcal{N}_W)} I_c^+(\mathcal{N}_W,\rho^{A_R}),$$

where for any real x, $x^+ \triangleq \max(x,0)$ and

$$\mathcal{S}(\rho^B,\mathcal{N}_W) \triangleq \{\rho^{A_R}\in\mathcal{D}(\mathcal{H}_{A_R}) : \mathcal{N}_W(\rho^{A_R}) = \rho^{B_R}\}.$$

Proof. A proof of the achievability is provided in Sections 5.5 and 5.6, and a proof of converse is provided in Section 5.7. □

Remark V.10 (Covering of Subspaces). The asymptotic rate obtained in the statement of Theorem V.9 can be interpreted using a subspace covering argument. Let us assume we are given a source ρ^B and a CPTP map $\mathcal{N}_W$ whose coherent information is positive for all $\rho^{A_R}\in\mathcal{S}(\rho^B,\mathcal{N}_W)$. Let $W : \mathcal{H}_{A_R}\to\mathcal{H}_{B_R}\otimes\mathcal{H}_E$ be a Stinespring's dilation of $\mathcal{N}_W$. This implies $I_c(\mathcal{N}_W,\rho^{A_R}) = S(B_R)_\sigma - S(E)_\sigma$, where $\sigma^{B_R E} \triangleq W\rho^{A_R}W^\dagger$, for $\rho^{A_R}\in\mathcal{S}(\rho^B,\mathcal{N}_W)$. We know that the n-product source state $\rho^{B^{\otimes n}}$ can be compressed using Schumacher compression to a subspace of normalized logarithmic dimension $S(B_R)_\sigma$ with high probability. In order to further reduce the rate, we use the posterior reference map of W with respect to ρ^{B_R} such that its action on the source produces the state ρ^A. Each basis vector in the reconstruction space can be thought of as covering a subspace of normalized logarithmic dimension of $S(E)_\sigma$ in the reference space. Therefore, one needs a rate of coherent information (which is the difference of the two entropies) to cover the entire source space with high probability. A similar observation was made for the quantum channel coding problem in *Lloyd*

(1997).

Remark V.11 (Comparison with Schumacher's lossless compression). Schumuacher's compression *Schumacher* (1995) requires $\lim_{n\to\infty} \|\omega^{B_R^n A^n} - \Psi_\rho^{B_R^n B^n}\| = 0$. In the current formulation, if one chooses the identity map as the posterior reference map, i.e., $\mathcal{N}_W = I_{A_R \to B_R}$, we require the condition $\lim_{n\to\infty} \|\omega^{B_R^n A^n} - \Psi_\omega^{A_R^n A^n}\| = 0$. Using Lemma V.4, monotonicity of the trace norm, and the triangle inequality, one can show that the two conditions are equivalent. Subsequently, both formulations yield the same asymptotic performance limit of von Neumann entropy. Observe that the standard source coding formulation using the average single-letter distortion criterion at zero distortion level is not equivalent to Schumacher's compression.

Remark V.12 (Comparison with average single-letter rate distortion). Given any sequence of (n, Θ) lossy quantum compression protocol for a quantum source coding setup $(\rho^B, \mathcal{H}_A, \mathcal{N}_W)$ that achieves the optimality in Theorem V.9, we observe that the following is true. Let $\omega^{B_R^n A^n} \triangleq (I \otimes \mathcal{N}_{\mathcal{D}}^{(n)})(I \otimes \mathcal{N}_{\mathcal{E}}^{(n)})(\Psi_\rho^{B_R^n B^n})$ be the induced state of the n-letter reference and the reconstruction by the protocol. Since the protocol satisfies (5.2), by monotonicity of trace distance, we obtain

$$\lim_{n\to\infty} \|\omega^{B_{R_i} A_i} - (\mathcal{N}_W \otimes I_A)(\Psi_\omega^{A_{R_i} A_i})\|_1 = 0, \quad \forall\, 1 \le i \le n,$$

where $\Psi_\omega^{A_{R_i} A_i} \triangleq \mathrm{Tr}_{A^{n\setminus i} A_R^{n\setminus i}}[\Psi_\omega^{A_R^n A^n}]$. It is worth noting that $\Psi_\omega^{A_{R_i} A_i}$ is not necessarily a pure state. Moreover, this does not necessarily provide any guarantee on the average single-letter distortion between the reference and the reconstruction as considered in the standard formulation of the problem (*Datta et al.*, 2013a, Lemma 1), where a single-letter purification of the source is taken into account. From this perspective, the current formulation is more "optimistic" in terms of measuring the quality of the reconstruction.

Remark V.13 (Comparison with Entanglement Assistance). We note that

$$I_c(\mathcal{N}_W, \rho^{A_R}) = \frac{1}{2}\left[I(B_R; A)_\sigma - I(A; E)_\sigma\right] \le \frac{1}{2} I(B_R; A)_\sigma,$$

where $\sigma^{B_R A E} \triangleq (I \otimes V)\Psi_\rho^{B_R B}(I \otimes V)^\dagger$, and $V : \mathcal{H}_B \to \mathcal{H}_A \otimes \mathcal{H}_E$ is a posterior reference map of W with respect of ρ^{B_R}. It was shown in *Datta et al.* (2013a) that $\frac{1}{2} I(B_R; A)_\sigma$ characterizes the asymptotic performance limit for the rate-distortion problem (with a local single-letter distortion

function) with unlimited entanglement assistance. Hence, this also provides a lower bound on the asymptotic performance limit for the corresponding problem in the unassisted case. Fortunately, this does not lead to any contradiction, as the current formulation differs from the former by being more optimistic.

5.4 Illustrative Examples

Example V.14 (Quantum Source Coding using Bit-Flip Channel)**.** In this example, we analyze the performance of a lossy quantum compression protocol corresponding to a quantum source coding setup $(\rho^B, \mathcal{H}_A, \mathcal{N}_W)$, where ρ^B is chosen as the maximally mixed state ($\rho^B = I_B/2$), and $\mathcal{N}_W \colon \mathcal{H}_{A_R} \to \mathcal{H}_{B_R}$ is specified as a bit-flip channel. An isometry $W \colon \mathcal{H}_{A_R} \to \mathcal{H}_{B_R} \otimes \mathcal{H}_E$ for $\mathcal{N}_W$ can be specified as

$$W = \sqrt{1-p} I \otimes |0\rangle^E + \sqrt{p} X \otimes |1\rangle^E,$$

where $\mathcal{N}_W(\rho^{A_R}) = \mathrm{Tr}_E(W \rho^{A_R} W^\dagger)$ for all $p \in (0, 1/2)$. Note that the canonical purification $|\psi_\rho\rangle^{B_R B}$ of ρ^B is given by

$$|\psi_\rho\rangle^{B_R B} = \frac{1}{\sqrt{2}} \left(|0\rangle^{B_R} |0\rangle^B + |1\rangle^{B_R} |1\rangle^B \right), \tag{5.3}$$

where $|0\rangle^{B_R} \triangleq (I \otimes \langle 0|^B) |\Gamma\rangle^{B_R B}$. This implies, $\rho^{B_R} = I_{B_R}/2$. To compute the asymptotic performance of the protocol for this source coding setup, as characterized by Theorem V.9, we first need to identify a ρ^{A_R} such that $\mathcal{N}_W(\rho^{A_R}) = \rho^{B_R}$. A simple computation reveals $\mathcal{S}(\rho^B, \mathcal{N}_W) = \{I_{A_R}/2\}$. This gives

$$\min_{\rho^{A_R} \in \mathcal{S}(\rho^B, \mathcal{N}_W)} I_c^+(\mathcal{N}_W, \rho^{A_R}) = I_c(\mathcal{N}_W, I_{A_R}/2) = S(B_R)_\sigma - S(E)_\sigma,$$

where $\sigma^{B_R E} = W \rho^{A_R} W^\dagger$. Note that $\sigma^{B_R} = I_{B_R}/2$ and $\sigma^E = (1-p) |0\rangle\langle 0|^E + p |1\rangle\langle 1|^E$, which gives $I_c(\mathcal{N}_W, I_{A_R}/2) = 1 - h_b(p)$, where $h_b(p) \triangleq -p \log(p) - (1-p) \log(1-p)$. Therefore, a maximally mixed source can be compressed at a rate $1 - h_b(p)$ while satisfying the error criterion as defined in (5.2).

Example V.15 (Quantum Source Coding using Depolarizing Channel)**.** In this example, we study the performance of another candidate channel, namely a depolarising channel. We again proceed

with the objective of compressing a maximally mixed state $\rho^{B_R} = \frac{I_{B_R}}{2}$, with $\mathcal{N}_W$ defined as

$$\mathcal{N}_W(\rho^{A_R}) = \left(1 - \frac{3p}{4}\right)\rho^{A_R} + \frac{p}{4}(X\rho^{A_R}X^\dagger + Y\rho^{A_R}Y^\dagger + Z\rho^{A_R}Z^\dagger).$$

for some $p \in [0,1]$. A simple calculation to satisfy $\mathcal{N}_W(\rho^{A_R}) = \rho^{B_R} = \frac{I_{B_R}}{2}$ reveals $\mathcal{S}(\rho^B, \mathcal{N}_W) = \{I_{A_R}/2\}$, for all $p \in (0,1)$. Analogous to the above example, finding an isometric extension of $\mathcal{N}_W$ gives

$$\min_{\rho^{A_R} \in \mathcal{S}(\rho^B, \mathcal{N}_W)} I_c^+(\mathcal{N}_W, \rho^{A_R}) = I_c^+(\mathcal{N}_W, I_{A_R}/2) = \max\left\{0, 1 - h_b(3p/4) - \frac{3p}{4}\log(3)\right\}.$$

Example V.16 (Hamming codes for quantum source compression)**.** In this example, we look at how Hamming codes perform when evaluated using the standard single-letter (local) entanglement fidelity criterion. Hamming codes are perfect codes, and achieve the Delsarte upper bound on the covering radius *Mattson Jr* (2012). Again, let $\rho^B = \frac{I_B}{2}$. Let a maximally entangled bipartite state $|\psi_m\rangle^{B_R B}$, defined as

$$|\psi_m\rangle^{B_R B} = \frac{1}{\sqrt{2}}\left(|00\rangle^{B_R B} + |11\rangle^{B_R B}\right), \tag{5.4}$$

be the purification of ρ^B. Let $\mathbb{F}_2$ denote a binary finite field, and let $G \in \mathbb{F}_2^{k\times n}$ be the generator matrix of a Hamming code. To encode ρ^B, we appeal to the duality perspective, and use the decoder of a Hamming code. Then the encoding is defined as $\mathcal{E}(x^n) \triangleq \operatorname{argmin}_{u^k}\{w_H(u^k\, G \oplus x^n)\}$, for all $x^n \in \mathbb{F}_2^n$, where w_H denotes the Hamming weight. Similarly, the decoder can be described as mapping $\mathcal{D} \circ \mathcal{E}((x^n)) = \mathcal{E}(x^n)G$. We describe this encoding as an isometric action $V_H : \mathcal{H}_B^{\otimes n} \to \mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_E^{\otimes n}$ taking the basis $|x^n\rangle^{B^n}$ to a vector $|\mathcal{E}(x^n)\rangle^{A^n} \otimes |x^n \oplus \mathcal{E}(x^n)\rangle^{E^n} \in \mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_E^{\otimes n}$, where the subsystem $\mathcal{H}_A^{\otimes n}$ stores the reconstruction and $\mathcal{H}_E^{\otimes n}$ is eventually traced out, and $\mathcal{H}_A$ is assumed to be an isomorphic copy of $\mathcal{H}_B$. This implies that the encoded state can be characterized as

$$\rho^{B_R^n A^n} = \operatorname{Tr}_{E^n}\left\{V_H \left|\psi_m^{\otimes n}\right\rangle\!\!\left\langle\psi_m^{\otimes n}\right|^{B_R^n B^n} V_H^\dagger\right\}.$$

Using

$$V_H \left|\psi_m^{\otimes n}\right\rangle^{B_R^n B^n} = \frac{1}{\sqrt{2^n}} \sum_{x^n} |x^n\rangle_{B_R^n} |\mathcal{E}(x^n)\rangle_{A^n} |x^n \oplus \mathcal{E}(x^n)\rangle_{E^n}$$
$$= \frac{1}{\sqrt{2^n}} \sum_{c^n \in \mathcal{C}} \sum_{e^n \in \mathbb{F}_2^n : w_H(e^n) \leq 1} |c^n \oplus e^n\rangle_{B_R^n} |c^n\rangle_{A^n} |e^n\rangle_{E^n},$$

we can simplify $\rho^{B_R^n A^n}$ as

$$\rho^{B_R^n A^n} = \frac{1}{2^n} \sum_{c^n, c'^n, e^n} |c^n \oplus e^n\rangle \langle c'^n \oplus e^n| \otimes |c^n\rangle \langle c'^n|, \tag{5.5}$$

where $\mathcal{C}$ denotes the set of codewords of the Hamming code. To compute the single-letter entanglement fidelity, we compute

$$\rho^{B_{R_i} A_i} = \mathrm{Tr}_{B_R^{n\setminus i} A^{n\setminus i}} \left\{\rho^{B_R^n A^n}\right\} = \frac{1}{2^n} \sum_{c^n, e^n} |c_i \oplus e_i\rangle \langle c_i \oplus e_i| \otimes |c_i\rangle \langle c_i|, \tag{5.6}$$

where tracing is performed on all the subsystems except corresponding to $B_R^{n\setminus i} A^{n\setminus i}$, and the second equality follows from using the fact that minimum Hamming distance of any Hamming code is three. This gives,

$$\langle \psi_m^{B_R B} | \rho^{B_{R_i} A_i} | \psi_m^{B_R B} \rangle = \frac{1}{2} \frac{1}{2^n} \sum_{c^n e^n} \left[\mathbb{1}_{\{c_i \oplus e_i = 0, c_i = 0\}} + \mathbb{1}_{\{c_i \oplus e_i = 1, c_i = 1\}}\right] = \frac{1}{2^{n+1}} \sum_{c^n e^n} \mathbb{1}_{\{e_i = 0\}}. \tag{5.7}$$

Therefore,

$$\frac{1}{n} \sum_{i=1}^{n} \langle \psi_m{}^{B_R B} | \rho^{B_{R_i} A_i} | \psi_m{}^{B_R B} \rangle = \frac{1}{2^{n+1} n} \sum_{c^n e^n} \sum_{i=1}^{n} \mathbb{1}_{\{e_i = 0\}} = \frac{|\mathcal{C}| n^2}{2^{n+1} n} = \frac{n}{2 \cdot 2^{n-k}}. \tag{5.8}$$

We know that for Hamming codes $k = 2^r - r - 1$ and $n = 2^r - 1$, which simplifies as

$$\frac{1}{n} \sum_{i=1}^{n} \langle \psi_m{}^{B_R B} | \rho^{B_{R_i} A_i} | \psi_m{}^{B_R B} \rangle = \frac{2^r - 1}{2 \cdot 2^r}, \tag{5.9}$$

and goes to half as r goes to infinity. Note that $r \to \infty$ serves as both a demonstration of the code's asymptotic performance and the condition for the rate k/n to reach unity. This results in a discontinuous asymptotic performance, since at rate exactly one, trivial identity encoding can be

used to achieve the average single-letter fidelity of unity. Further, note that $S(E^n) = \log(n+1) = r$. Hence the normalized amount of qubits that is dissipated, given by $\frac{S(E^n)}{n}$, approaches zero as $r \to \infty$, indicating that there is significant entanglement between the reconstruction and the reference.

As was demonstrated in Example V.14, it is possible to compress a maximally mixed source in a continuous fashion, when the error is measured in accordance with the suggested definition in (5.2),

5.5 Achievability Proof Overview

We provide a brief overview of the achievability proof before formally presenting one. The proof we present here is inspired by Devetak's work in *Devetak* (2005b) for the quantum channel communication problem (also detailed in (*Wilde*, 2011, Chapter 24)). An integral component of that work is the decomposition of coherent information as the difference of two Holevo information quantities. We intend to perform a similar decomposition, but from the perspective of the given map $\mathcal{N}_W$. Toward this, for the given source ρ^B, we first search for a $\rho^{A_R} \in \mathcal{D}(\mathcal{H}_{A_R})$, satisfying $\mathcal{N}_W(\rho^{A_R}) = \rho^{B_R}$. Once found, using the spectral decomposition, we expand ρ^{A_R} as $\rho^{A_R} = \sum_{a\in\mathcal{A}} \lambda_a^A |a\rangle\langle a|^{A_R}$, for some finite set $\mathcal{A}$. Observe that since $|a\rangle\langle a|^{A_R}$ is pure, $S(\mathcal{N}_W(|a\rangle\langle a|^{A_R})) = S(\mathcal{N}_W^c(|a\rangle\langle a|^{A_R}))$, where $\mathcal{N}_W^c : \mathcal{H}_{A_R} \to \mathcal{H}_E$ is a complementary CPTP map of $\mathcal{N}_W$, defined using the Stinespring's dilation $W : \mathcal{H}_{A_R} \to \mathcal{H}_{B_R} \otimes \mathcal{H}_E$ corresponding to $\mathcal{N}_W$. This also means that

$$\sum_{a\in\mathcal{A}} \lambda_a^A S(\mathcal{N}_W(|a\rangle\langle a|^{A_R})) = \sum_{a\in\mathcal{A}} \lambda_a^A S(\mathcal{N}_W^c(|a\rangle\langle a|^{A_R})).$$

Furthermore, from the linearity of CPTP maps, we see

$$\sum_{a\in\mathcal{A}} \lambda_a^A \mathcal{N}_W(|a\rangle\langle a|^{A_R}) = \mathcal{N}_W(\rho^{A_R}) \quad \text{and} \quad \sum_{a\in\mathcal{A}} \lambda_a^A \mathcal{N}_W^c(|a\rangle\langle a|^{A_R}) = \mathcal{N}_W^c(\rho^{A_R}).$$

This implies, we can rewrite $I_c(\mathcal{N}_W, \rho^{A_R})$ as

$$I_c(\mathcal{N}, \rho^{A_R}) = S(\mathcal{N}_W(\rho^{A_R})) - S(\mathcal{N}_W^c(\rho^{A_R}))$$

$$= \left[S(\mathcal{N}_W(\rho^{A_R})) - \sum_{a\in\mathcal{A}} S(\lambda_a^A \mathcal{N}_W(|a\rangle\langle a|^{A_R})) \right] - \left[S(\mathcal{N}_W^c(\rho^{A_R})) - \sum_{a\in\mathcal{A}} \lambda_a^A S(\mathcal{N}_W^c(|a\rangle\langle a|^{A_R})) \right]$$
$$= \chi\left(\left\{ \lambda_a^A, \mathcal{N}_W(|a\rangle\langle a|^{A_R}) \right\} \right) - \chi\left(\left\{ \lambda_a^A, \mathcal{N}_W^c(|a\rangle\langle a|^{A_R}) \right\} \right). \tag{5.10}$$

Now our aim is to show the achievability of a rate equal to the above difference. After obtaining a similar decomposition, Devetak achieved the performance limit by applying a coherent version of the CQ packing lemma (*Wilde*, 2011, Chapter 16) followed by an application of the QC covering lemma (*Wilde*, 2011, Chapter 17). Inspired by this, and the duality connections between the two problems, we achieve the difference obtained in (5.10). In particular, we start with the objective of applying a coherent version of the QC covering lemma (or the measurement compression result *Winter* (2004)). Toward this, as shown in Figure 5.1, we first obtain a posterior reference map V corresponding to the isometry W. Then we identify the action of V on the state ρ^B as a coherent quantum measurement. Now, using the approximating POVMs constructed in *Winter* (2004), we perform a coherent covering that allows us to compress the obtained measurement, and in turn the output of V, at rate given by the first Holevo information. The compression is performed while *faithfully simulating* the action of V, giving a reconstruction satisfying the error criterion (as in (5.2)). This procedure is delineated in Step 1.1 where an encoder is constructed to perform coherent covering and in Step 2.1 where the effect of covering is analyzed, and a rate corresponding to the first Holevo information is achieved.

To get the needed coherent information, the rate corresponding to the second Holevo information must be further decreased. This entails diffusing more data or qubits into the environment (partial tracing). However, as will be demonstrated in the proof below, such an action would destroy quantum correlations present in the source, possibly turning it into a classical mixture. Therefore, before such partial tracing operation, in Step 1.2 (Section 5.6) we construct a unitary operation that can condense the information into fewer qubits in the form of entanglement, and thus allowing for further decrease in the rate. This includes using the coherent post-measurement state of the subsystem E as side information available at the encoder. The Step 2.2 of Section 5.6 details this procedure and achieves the desired rate. Finally, an additional step (Step 2.3) is required to show the intended closeness as required in (5.2).

Another intriguing aspect of the proof is the technique used for expurgation. As clear from

the definition of the protocol, it only allows unitary or isometric operations, followed by partial tracing. When a code contains repeated codewords, it can be difficult to guarantee this. An approach to removing all repetitions is to perform expurgations. This is achieved by finding a good code (satisfying all its constraints) while allowing a small fraction of repeats and then expurgating just this fraction of the code. However, if there are exponentially many constraints, it becomes challenging to finding a good code. The exponentially many covering constraints in Devetak's problem have a doubly exponential decreasing probability of error, which Devetak was able to take advantage of. In the current problem we instead have exponentially many packing constraints which only have an exponential decay. In order to combat this, we construct our proof to just require one packing constraint: the average of all exponentially many packing constraints. This enables us to find a good code and successfully expurgate it. We now formally construct the arguments toward proving the statement of the theorem.

5.6 Proof of Achievability

The proof is mainly composed of four parts. In the first part, we develop the necessary single-letter ensembles required in the proof. In the next part, we provide the random coding setup and the distributions on the ensembles with which the codewords are generated. We also state here the constraints that a good code must satisfy and argue the existence of one code with non-zero probability. We further use an expurgation strategy to make all the codewords distinct. In the third part, we construct a protocol by developing all the actions of the encoder and the decoder and describing them as unitary (or isometry) evolutions. Note that the only actions allowed by the protocol (Definition V.7) are quantum channels which can be described as unitary or isometric evolutions followed by partial trace operations. In parallel, we also provide the necessary lemmas needed for the next part. The last part deals with analyzing the action of encoding and decoding operations on the source ρ^B, and then bounding the trace distance as in Definition V.7.

Toward this, fix two positive integers M and K, and $\epsilon \in (0,1)$. Let $\mathcal{M}$ and $\mathcal{K}$ denote the sets $[0, M-1]$ and $[0, K-1]$, respectively. Given a quantum source coding setup $(\rho^B, \mathcal{H}_A, \mathcal{N}_W)$, let $|\psi_\rho\rangle^{B_R B}$ be the canonical purification of ρ^B and $\rho^{B_R} \triangleq \mathrm{Tr}_B\{\Psi_\rho{}^{B_R B}\}$. Moreover, let $\mathcal{H}_{A_R}$ be the reference space associated with $\mathcal{H}_A$. Now choose $\rho^{A_R} \in S(\rho^B, \mathcal{N}_W)$. Let $\mathcal{H}_E$ denote the Hilbert

space such that $W : \mathcal{H}_{A_R} \to \mathcal{H}_{B_R} \otimes \mathcal{H}_E$ forms an isometric extension (or Stinespring's dilation) of $\mathcal{N}_W$ according to (*Wilde*, 2011, Definition 5.2.1) with $\dim(\mathcal{H}_E) \geq \dim(\mathcal{H}_{B_R})$. As shown in Figure 5.1, define the posterior reference isometry of W with respect to ρ^{B_R} (according to Definition V.1) as the isometry $V : \mathcal{H}_B \to \mathcal{H}_A \otimes \mathcal{H}_E$ satisfying $(W \otimes I_A) |\psi_\rho\rangle^{A_R A} = (I_{B_R} \otimes V) |\psi_\rho\rangle^{B_R B}$ where $|\psi_\rho\rangle^{A_R A}$ is the canonical purification of ρ^{A_R}. Let $\rho^A \triangleq \mathrm{Tr}_{B_R E}\{(I \otimes V)\Psi_\rho{}^{B_R B}(I \otimes V)^\dagger\}$.

5.6.1 Defining the ensembles

In this section, we construct the single-letter ensembles corresponding to two Holevo information quantities used in the decomposition of coherent information discussed in Section 5.5. We begin by using the definition of W to obtain,

$$(I_{B_R} \otimes V) |\psi_\rho\rangle^{B_R B} = (W \otimes I_A) |\psi_\rho\rangle^{A_R A} = \sum_{a \in \mathcal{A}} \sqrt{\lambda_a^A} W |a\rangle^{A_R} \otimes |a\rangle^A , \tag{5.11}$$

where we use $\rho^A = \sum_{a \in \mathcal{A}} \lambda_a^A |a\rangle\langle a|^A$ as its spectral decomposition, and define $|a\rangle^{A_R} \triangleq (I_{A_R} \otimes \langle a|^A) |\Gamma\rangle^{A_R A}$ for $a \in \mathcal{A}$, for some finite set $\mathcal{A}$. This also gives,

$$W |a\rangle^{A_R} = \frac{(\langle a|^A \otimes I_{B_R E})(I_{B_R} \otimes V) |\psi_\rho\rangle^{B_R B}}{\sqrt{\lambda_a^A}}. \tag{5.12}$$

Using the spectral decomposition of ρ^B as $\rho^B = \sum_{b \in \mathcal{B}} \lambda_b^B |b\rangle\langle b|^B$, for $b \in \mathcal{B}$ for some finite set $\mathcal{B}$, we can rewrite the action of V on ρ^B as

$$\begin{aligned}
(I_{B_R} \otimes V) |\psi_\rho\rangle^{B_R B} &= \sum_{b \in \mathcal{B}} \sqrt{\lambda_b^B} |b\rangle^{B_R} \otimes V |b\rangle^B \\
&= \left(I_{B_R} \otimes I_E \otimes \sum_{a \in \mathcal{A}} |a\rangle\langle a|^A \right) \sum_{b \in \mathcal{B}} \sqrt{\lambda_b^B} |b\rangle^{B_R} \otimes V |b\rangle^B \\
&= \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} \sqrt{\lambda_b^B} |b\rangle^{B_R} M_a |b\rangle^B \otimes |a\rangle^A ,
\end{aligned} \tag{5.13}$$

where we define $|b\rangle^{B_R} \triangleq (I_{B_R} \otimes \langle b|^B) |\Gamma\rangle^{B_R B}$, and $M_a : \mathcal{H}_B \to \mathcal{H}_E$ as

$$M_a \triangleq \left(I_E \otimes \langle a|^A\right) V. \tag{5.14}$$

By defining a POVM $\Lambda \triangleq \{M_a^\dagger M_a\}_{a\in\mathcal{A}}$, we can identify a coherent measurement (isometry) U_Λ corresponding to Λ with $U_\Lambda \triangleq \sum_{a\in\mathcal{A}} M_a \otimes |a\rangle^A$, and therefore express the action of V as

$$(I_{B_R} \otimes V)\,|\psi_\rho\rangle^{B_R B} = (I_{B_R} \otimes U_\Lambda)\,|\psi_\rho\rangle^{B_R B}\,. \tag{5.15}$$

Now our objective is to *faithfully simulate* the action of the isometry (or the coherent measurement) U_Λ while using an exponentially smaller subspace in $\mathcal{H}_{A^n}$. Equivalently, we intend to minimize the amount of qubits needed to represent the quantum state in the Hilbert space $\mathcal{H}_{A^n}$. Employing Schumacher's compression *Schumacher* (1995), one can only achieve a rate of Von-Neumann entropy while faithfully simulating U_Λ. However, since U_Λ is a coherent measurement, we employ a coherent version of the measurement compression protocol *Winter* (2004) and demonstrate a faithful simulation of the isometry while further decreasing the resource requirement. In particular, an approximating coherent measurement (henceforth referred to as the covering isometry) $U_\mathcal{M}$ is constructed to faithfully simulate the action of U_Λ while requiring the rate equal to Holevo quantity corresponding to the canonical ensemble $\{\lambda_a^A, \hat{\rho}_a^{B_R}\}$, where

$$\hat{\rho}_a^{B_R} \triangleq \frac{\sqrt{\rho^{B_R}}(M_a^\dagger M_a)^T\sqrt{\rho^{B_R}}}{\lambda_a^A} \quad \text{and} \quad (M_a^\dagger M_a)^T \triangleq \sum_{b,b'} |b\rangle\,\langle b'|^{B_R}\langle b'|(M_a^\dagger M_a)\,|b\rangle^B\,. \tag{5.16}$$

Observe that using the definition of M_a from (5.14), it follows

$$\begin{aligned}
\mathrm{Tr}\Big\{M_{a'}^\dagger M_a \rho^B\Big\} &= \mathrm{Tr}\Big\{(I_{B_R E}\otimes\langle a|)V\,|\psi_\rho\rangle\langle\psi_\rho|^{B_R B}\,V^\dagger(|a'\rangle\otimes I_{B_R E})\Big\} \\
&= \sum_b \sqrt{\lambda_a^A\lambda_{a'}^A}\,\mathrm{Tr}\Big\{\langle b|\,W|a'\rangle\,\langle a|\,W^\dagger\,|b\rangle\Big\} \\
&= \sum_b \sqrt{\lambda_a^A\lambda_{a'}^A}\,\langle a|\,W^\dagger\,|b\rangle\langle b|\,W|a'\rangle = \lambda_a^A\cdot\mathbb{1}_{\{a=a'\}},
\end{aligned} \tag{5.17}$$

for all $a, a' \in \mathcal{A}$, where the first equality uses the definition of M_a, and the second follows from using the relation (5.11). Using the simplification from (5.12), it is useful to note

$$W\,|a\rangle^{A_R} = \frac{(I_{B_R}\otimes M_a)\,|\psi_\rho\rangle^{B_R B}}{\sqrt{\lambda_a^A}}. \tag{5.18}$$

For the second Holevo information, we define the packing ensemble $\{\lambda_a^E, \tau_a^E\}$ as

$$\tau_a^E \triangleq \frac{\mathrm{Tr}_{B_R}\left(I_{B_R}\otimes M_a\right)\Psi_\rho^{B_R B}\left(I_{B_R}\otimes M_a\right)^\dagger}{\lambda_a^E} = \frac{M_a \rho^B M_a^\dagger}{\lambda_a^E}, \quad \text{and} \quad \lambda_a^E \triangleq \lambda_a^A. \tag{5.19}$$

The discussion on how this ensemble is employed to reduce the rate follows in the sequel.

5.6.2 Random Coding and Expurgation

In this section, we construct the random coding argument, and simultaneously, define all the conditions that pertain to the construction of a good random code. Subsequently, we randomly generate one code that satisfies these constraints. We then expurgate this code to ensure no repetitions are present. Toward constructing an approximating coherent measurement $U_{\mathcal{M}}$, randomly and independently select $|\mathcal{M}| \times |\mathcal{K}|$ sequences $A^n(m,k)$ according to the following pruned distribution

$$\mathbb{P}\left(A^n(m,k) = a^n\right) = \begin{cases} \dfrac{\lambda_{a^n}^A}{(1-\varepsilon)} & \text{for} \quad a^n \in \mathcal{T}_\delta^{(n)}(A) \\ 0 & \text{otherwise,} \end{cases} \tag{5.20}$$

where $\varepsilon = \sum_{a^n \notin \mathcal{T}_\delta^{(n)}(A)} \lambda_{a^n}^A$, $\mathcal{T}_\delta^{(n)}(A)$ is the δ-typical set corresponding to the distribution λ_a^A on the set $\mathcal{A}$, and $\lambda_{a^n}^A \triangleq \Pi_{i=1}^n \lambda_{a_i}^A$. Let $\mathcal{C}^{(m)}$ denote the codebook $\{A^n(m,k)\}_{k\in\mathcal{K}}$ for a given m, and $\mathcal{C}$ denote the collection of all codebooks $\{\mathcal{C}^{(m)}\}_{m\in\mathcal{M}}$. Further, for each $a^n \in \mathcal{T}_\delta^{(n)}(A)$ define

$$\tilde{\rho}_{a^n}^{B_R} \triangleq \hat{\pi}\pi_{\rho^{B_R}}\pi_{a^n}\hat{\rho}_a^{B_R}\pi_{a^n}\pi_{\rho^{B_R}}\hat{\pi}, \tag{5.21}$$

and $\tilde{\rho}_{a^n}^{B_R} = 0$, for $a^n \notin \mathcal{T}_\delta^{(n)}(A)$, where $\hat{\rho}_a^{B_R} \triangleq \bigotimes_i \hat{\rho}_{a_i}^{B_R}$, $\pi_{\rho^{B_R}}$ and π_{a^n} are the $\delta-$typical and conditionally typical projectors defined as in (*Wilde*, 2011, Def. 15.1.3) and (*Wilde*, 2011, Def. 15.2.4), with respect to $\rho^{B_R} = \sum_{a\in\mathcal{A}} \lambda_a^A \hat{\rho}_a^{B_R}$ and $\hat{\rho}_a^{B_R}$, respectively, and $\hat{\pi}$ is the cut-off projector as defined in *Winter* (2004). Using the Average Gentle Measurement Lemma (*Wilde*, 2011, Lemma 9.4.3), for any given $\epsilon \in (0,1)$, and all sufficiently large n and all sufficiently small δ, we have

$$\sum_{a^n \in \mathcal{A}^n} \lambda_{a^n}^A \|\hat{\rho}_{a^n}^{B_R} - \tilde{\rho}_{a^n}^{B_R}\|_1 \leq \epsilon. \tag{5.22}$$

A detailed proof of the above statement can be found in (*Wilde et al.*, 2012, Eq. 35). Using these

definitions, construct operators

$$A_{a^n}^{B_R^n} \triangleq \gamma_{a^n}\left(\sqrt{\rho^{B_R\otimes n}}^{-1}\tilde{\rho}_a^{B_R}\sqrt{\rho^{B_R\otimes n}}^{-1}\right), \quad \gamma_{a^n} \triangleq \frac{1-\varepsilon}{1+\eta}\frac{1}{|\mathcal{M}||\mathcal{K}|}|\{(m,k) : A^n(m,k) = a^n\}|, \quad (5.23)$$

and $\eta \in (0,1)$ is a parameter that determines the probability of not obtaining a sub-POVM. Note that in the definition of $A_{a^n}^{B_R^n}$ the right hand side operates on $\mathcal{H}_{B_R^n}$, however, we define A_{a^n} belonging to $\mathcal{L}(\mathcal{H}_{A^n})$. To obtain this, we transform $A_{a^n}^{B_R^n}$ as

$$A_{a^n} = \sum_{b^n,\bar{b}^n} \langle b^n|A_{a^n}^{B_R^n}|\bar{b}^n\rangle_{B_R}|\bar{b}^n\rangle \langle b^n|_B .$$

Then construct a sub-POVM $\Gamma^{(n)}$ as

$$\Gamma^{(n)} \triangleq \{A_{a^n} : a^n \in \mathcal{T}_\delta^{(n)}(A)\}. \quad (5.24)$$

Let $\mathbb{1}_{\{\mathrm{sP}\}}$ denote the indicator random variable corresponding to the event that $\Gamma^{(n)}$ forms a sub-POVM. We have the following result.

Proposition V.17. *For any $\epsilon \in (0,1)$, any $\eta \in (0,1)$, any $\delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}\left[\mathbb{1}_{\{\mathrm{sP}\}}\right] > 1-\epsilon$, if $\frac{1}{n}(\log M + \log K) > \chi(\lambda_a^{B_R}, \hat{\rho}_a^{B_R})$.*

Proof. The result follows from *Winter* (2004). □

Define the code dependent random variables E_1 and E_2 as

$$E_1 \triangleq \sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}}(|\mathcal{M}||\mathcal{K}|)^{-1}\operatorname{Tr}\left\{\tilde{\rho}_{m,k}^{B_R}\right\}, \quad \text{and} \quad E_2 \triangleq \sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}}(|\mathcal{M}||\mathcal{K}|)^{-1}\left\|\tilde{\rho}_{m,k}^{B_R} - \hat{\rho}_{m,k}^{B_R}\right\|_1,$$

where $\hat{\rho}_{m,k}^{B_R}$, and $\tilde{\rho}_{m,k}^{B_R}$ are used as shorthand notations to denote $\tilde{\rho}_{a^n(m,k)}^{B_R}$ and $\tilde{\rho}_{a^n(m,k)}^{B_R}$, respectively. Further, using the results (*Wilde et al.*, 2012, Eq. (28) and Eq. (35)), for all $\epsilon \in (0,1)$, we have $\mathbb{E}[E_1] \geq 1-\epsilon$, and $\mathbb{E}[E_2] \leq \epsilon$, for all sufficiently large n and all sufficiently small $\delta > 0$.

Now, considering the ensemble $\{\lambda_a^E, \tau_a^E\}$, we construct the operators $\{\tau_{a^n(m,k)}^E\}$ using the codebook $\mathcal{C}$ and the distribution defined in (5.20), where $\tau_{a^n}^E \triangleq \bigotimes_i \tau_{a_i}^E$. For this ensemble, we construct a collection of n-letter POVMs, one for each $m \in \mathcal{M}$, capable of decoding the message $k \in \mathcal{K}$. In

particular, we employ the Holevo POVMs *Holevo* (2019) defined as

$$\xi_k^{(m)} \triangleq \pi^\tau \pi_k^{(m)} \pi^\tau \quad \text{and} \quad \Xi_k^{(m)} \triangleq \left(\sum_{k' \in \mathcal{K}} \xi_{k'}^{(m)}\right)^{-1/2} \xi_k^{(m)} \left(\sum_{k' \in \mathcal{K}} \xi_{k'}^{(m)}\right)^{-1/2}, \tag{5.25}$$

where π^τ is the $\delta-$typical projector (as in (*Wilde*, 2011, Def. 15.1.3)) defined for the density operator $\tau \triangleq \sum_{a \in \mathcal{A}} \lambda_a^E \tau_a^E$, and $\pi_k^{(m)}$ denotes the strong conditional typical projectors (as in (*Wilde*, 2011, Def. 15.2.4)) for the operators $\tau_{a^n(m,k)}$. For these POVMs, we know the average probability of error can be made arbitrarily small. More formally, we have the following.

Proposition V.18. *Given the ensemble $\{\lambda_a^E, \tau_a^E\}$ and the collection of POVMs $\{\Xi_k^{(m)}\}_k$, for any $\epsilon \in (0,1)$,*

$$\mathbb{E}\left[\frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} \mathrm{Tr}\left\{\Xi_k^{(m)} \tau_k^{(m)}\right\}\right] \geq 1 - \epsilon, \tag{5.26}$$

for sufficiently small $\delta > 0$ and for all sufficiently large n, and for all $m \in \mathcal{M}$, if $\frac{1}{n} \log K < \chi(\{\lambda_a^E, \tau_a^E\})$, where $\tau_k^{(m)}$ is used as a shorthand for $\tau_{a^n(m,k)}$.

Proof. The proof follows from the result of classical communication over quantum channels *Holevo* (2019) or the packing lemma of (*Wilde*, 2011, Lemma 16.3.1) while making the following identification. For each $m \in \mathcal{M}$, identify $\mathcal{M}$ with $\mathcal{K}$, X with $\mathcal{T}_\delta^{(n)}(\mathcal{A})$, $\{\sigma_{C_m}\}_m$ with $\{\tau_k^{(m)}\}_k$, Π with π^τ, Π_x with $\pi_k^{(m)}$, d with $2^{n(S(E|A)_{\bar{\tau}} + \bar{\delta})}$, D with $2^{n(S(E)_{\bar{\tau}} - \bar{\delta})}$, and Λ_m with $\Xi_k^{(m)}$, where $\bar{\tau}^{AE} \triangleq \sum_a \lambda_a^E |a\rangle\langle a|_A \otimes \tau_a^E$ and $\bar{\delta}(\delta) \searrow 0$ as $\delta \searrow 0$. □

The above result also implies a weaker average result which suffices here. This can be stated as $\mathbb{E}[E_3] \geq 1 - \epsilon$, for sufficiently small $\delta > 0$ and for all sufficiently large n, if $\frac{1}{n} \log K < \chi(\{\lambda_a^E, \tau_a^E\})$, where

$$E_3 \triangleq \frac{1}{MK} \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}} \mathrm{Tr}\left\{\Xi_k^{(m)} \tau_k^{(m)}\right\}. \tag{5.27}$$

Finally, toward finding a good code, we need one last property which is that all its codewords are distinct. In the dual, the quantum channel communication problem *Devetak* (2005b), Devetak used the double exponential decay of the covering error to argue the existence of an expurgated code for

exponentially many covering constraints. However, in the current problem, we have exponentially many packing constraints, with each having only an exponential decay in the error. To resolve this issue, we develop a proof that only requires the average of the packing constraints. However, in such a case, it becomes unclear as to what should be the expurgation strategy. For this, we introduce another event that captures the non-distinctness of the codebook, and expurgate with respect to this event. Precisely, we define a codeword $A^n(m,k)$ is bad if there exists $(m',k') \neq (m,k)$ such that $A^n(m,k) = A^n(m',k')$. Let

$$E_4 \triangleq \frac{1}{MK} \sum_{m\in\mathcal{M}} \sum_{k\in\mathcal{K}} \mathbb{1}_{\{A^n(m,k) \text{ is bad}\}}.$$

Computing its expectation, we get

$$\begin{aligned}\mathbb{E}[E_4] &= \mathbb{E}\left[\frac{1}{MK} \sum_{m\in\mathcal{M}} \sum_{k\in\mathcal{K}} \mathbb{1}_{\{\exists (m',k')\neq(m,k) \text{ such that } A^n(m,k)=A^n(m',k')\}}\right] \\ &\overset{a}{\leq} \frac{1}{MK} \sum_{\substack{m,m'\in\mathcal{M}\\ k,k'\in\mathcal{K}\\ (m,k)\neq(m',k')}} \sum_{a^n\in\mathcal{T}_\delta^{(n)}(A)} \mathbb{E}\left[\mathbb{1}_{\{A^n(m,k)=a^n\}}\right] \mathbb{E}\left[\mathbb{1}_{\{A^n(m',k')=a^n\}}\right] \overset{b}{\leq} MK 2^{-n(S(\lambda_a^A)-\delta_1)} \leq \epsilon, \quad (5.28)\end{aligned}$$

for all sufficiently large n and sufficiently small $\delta > 0$ if $\frac{1}{n}(\log M + \log K) < S(\lambda_a^A)$, where (a) uses the mutual independence of the codewords, and (b) define δ_1 as $\delta_1(\delta,\varepsilon) \searrow 0$ as $\delta, \varepsilon \searrow 0$. Using the Markov inequality and the union bound, we have

$$\mathbb{P}\Big(\{\mathbb{1}_{\{\mathrm{sP}\}} = 1\} \cap \{E_1 \geq 1-\sqrt{\epsilon}\} \cap \{E_2 \leq \sqrt{\epsilon}\} \cap \{E_3 \geq 1-\sqrt{\epsilon}\} \cap \{E_4 \leq \sqrt{\epsilon}\}\Big) \geq 1 - 5\sqrt{\epsilon}.$$

Therefore, for all $\epsilon \in (0, 1/25)$, and for all sufficiently small $\delta > 0$, for all sufficiently large n there exists a code $\mathcal{C}$ that satisfies the conditions $\{\mathbb{1}_{\{\mathrm{sP}\}} = 1\}$, $\{E_1 \geq 1-\sqrt{\epsilon}\}$, $\{E_2 \leq \sqrt{\epsilon}\}$, $\{E_3 \geq 1-\sqrt{\epsilon}\}$, and $\{E_4 \leq \epsilon\}$, simultaneously if

$$\frac{1}{n}(\log M + \log K) > \chi(\lambda_a^{B_R}, \hat{\rho}_a^{B_R}), \quad \frac{1}{n}\log K < \chi(\{\lambda_a^E, \tau_a^E\}), \quad \frac{1}{n}(\log M + \log K) < S(\lambda_a^A). \quad \text{(S-0)}$$

At this point, we choose one such code $\mathcal{C}$ satisfying all the above conditions, and fix it for the rest of the analysis.

Toward showing that this chosen code achieves the asymptotic performance stated in the theorem statement, we expurgate the code $\mathcal{C}$ with respect to the random variable E_4, ensuring that the code has all distinct codewords. The assumption of codebook being distinct becomes crucial at multiple places in the proof and will be highlighted as necessary. Since $\{E_4 \leq \sqrt{\epsilon}\}$ ensures at most $\sqrt{\epsilon}MK$ codewords in $\mathcal{C}$ are not distinct, we remove $\sqrt{\epsilon}MK$ codewords from $\mathcal{C}$. This is performed by first removing all the non-distinct codewords, and then further removing some more from the distinct ones arbitrarily (if needed) until we remain with a total of $(1-\sqrt{\epsilon})MK$ codewords. Let the expurgated set (the remainder of the codewords) be denoted by $\mathcal{C}_{\mathcal{E}}$, and define the sets $\mathcal{C}_{\mathcal{E}}^{(m)}$ as $\mathcal{C}_{\mathcal{E}}^{(m)} \triangleq \mathcal{C}_{\mathcal{E}} \cap \mathcal{C}^{(m)}$. Observe that, all the codewords in $\mathcal{C}_{\mathcal{E}}$ are distinct. However, as opposed to $\mathcal{C}$ which was consistent with regards to the size of $\mathcal{C}^{(m)}$ (equal to K for all $m \in \mathcal{M}$), $\mathcal{C}_{\mathcal{E}}$ has varying sizes. Therefore, we define K'_m to denote the size of $\mathcal{C}_{\mathcal{E}}^{(m)}$ and M' to denote number of non-empty sets in the collection $\{\mathcal{C}_{\mathcal{E}}^{(m)}\}_{m\in\mathcal{M}}$. Note that for some $m \in \mathcal{M}$, K'_m may be zero. Let $\mathcal{M}'$ denote the subset of $\mathcal{M}$ for which $K'_m > 0$, and let $\mathcal{H}'_M$ denote the corresponding Hilbert space with $\dim(\mathcal{H}'_M) = M' + 1$. As is evident, $\sum_{m\in\mathcal{M}'} K'_m = (1-\sqrt{\epsilon})MK$. In addition, define the set of indices corresponding to the expurgated codebook as $\mathcal{I}_E^{(m)} \triangleq \{k : a^n(m,k) \in \mathcal{C}_{\mathcal{E}}^{(m)}\}$ and $\mathcal{I}_E \triangleq \{(m,k) : a^n(m,k) \in \mathcal{C}_{\mathcal{E}}\}$. Further, for the expurgated code, we have

$$E'_1 \triangleq \frac{1}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|} \sum_{m\in\mathcal{M}'} \sum_{k\in\mathcal{I}_E^{(m)}} \mathrm{Tr}\Big\{\tilde{\rho}_{m,k}^{B_R}\Big\} \geq 1 - 2\sqrt{\epsilon}, \tag{5.29}$$

$$E'_2 \triangleq \frac{1}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|} \sum_{m\in\mathcal{M}'} \sum_{k\in\mathcal{I}_E^{(m)}} \left\|\tilde{\rho}_{m,k}^{B_R} - \hat{\rho}_{m,k}^{B_R}\right\|_1 \leq \frac{\sqrt{\epsilon}}{1-\sqrt{\epsilon}} \leq 2\sqrt{\epsilon}, \tag{5.30}$$

$$E'_3 \triangleq \frac{1}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|} \sum_{m\in\mathcal{M}'} \sum_{k\in\mathcal{I}_E^{(m)}} \mathrm{Tr}\Big\{\Xi_k^{(m)}\tau_k^{(m)}\Big\} \geq 1 - 2\sqrt{\epsilon}, \tag{5.31}$$

where the inequalities above uses the fact that codebook $\mathcal{C}$ satisfies $\{E_1 \geq 1-\sqrt{\epsilon}\}, \{E_2 \leq \sqrt{\epsilon}\}$ and $\{E_3 \geq 1-\sqrt{\epsilon}\}$ and that only $\sqrt{\epsilon}$ fraction of the code is expurgated. Observe that the event $\{\mathbb{1}_{\{\mathrm{sP}\}} = 1\}$ remains true for the expurgated $\mathcal{C}_{\mathcal{E}}$. Define the collection

$$\Gamma_{\mathcal{E}}^{(n)} \triangleq \{A_{a^n(m,k)}\}_{m\in\mathcal{M}', k\in\mathcal{I}_{\mathcal{E}}^{(m)}}.$$

The collection $\Gamma_{\mathcal{E}}^{(n)}$ is completed using the operator $I - \sum_{m\in\mathcal{M}'}\sum_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}} A_{a^n(m,k)}$, and the operator

is associated with sequence a_0^n chosen arbitrarily from $\mathcal{A}^n \backslash \mathcal{T}_\delta^{(n)}(A)$, i.e.,

$$A_{a_0^n} \triangleq I - \sum_{m \in \mathcal{M}'} \sum_{k \in \mathcal{I}_{\mathcal{E}}^{(m)}} A_{a^n(m,k)}.$$

Corresponding to this expurgated code, we now construct our encoding and decoding operations.

5.6.3 Encoding and Decoding Isometries

The encoding isometry $U_{\mathcal{E}}$ is constructed by concatenating three isometries: (i) the covering isometry $U_{\mathcal{M}} : \mathcal{H}_{B^n} \to \mathcal{H}_{B^n} \otimes \mathcal{H}'_M \otimes \mathcal{H}_K$, (ii) the rotation isometry $U_{\mathcal{R}} : \mathcal{H}_{B^n} \otimes \mathcal{H}'_M \otimes \mathcal{H}_K \to \mathcal{H}_{E^n} \otimes \mathcal{H}'_M \otimes \mathcal{H}_K$, and (iii) the packing isometry $U_{\mathcal{P}} : \mathcal{H}_{E^n} \otimes \mathcal{H}_{\bar{E}} \otimes \mathcal{H}'_M \otimes \mathcal{H}_K \to \mathcal{H}_{E^n} \otimes \mathcal{H}_{\bar{E}} \otimes \mathcal{H}'_M \otimes \mathcal{H}_K$, where $\mathcal{H}'_M, \mathcal{H}_K$ and $\mathcal{H}_{\bar{E}}$ are auxiliary Hilbert spaces with dimensions $M'+1, K+1$, and $K+1$, respectively.

Step 1.1: Covering Isometry

To define the covering isometry $U_{\mathcal{M}}$, we use the completion $[\Gamma_{\mathcal{E}}^{(n)}]$ as

$$U_{\mathcal{M}} \triangleq \sum_{m \in \mathcal{M}'} \sum_{k \in \mathcal{I}_{\mathcal{E}}^{(m)}} \sqrt{A_{a^n(m,k)}} \otimes |m\rangle \otimes |k\rangle + \sqrt{A_{a_0^n}} \otimes \left|M'\right\rangle_M \otimes |K\rangle_K . \tag{5.32}$$

Note that, for the chosen code, the event $\{\mathbb{1}_{\{\mathrm{sP}\}} = 1\}$ makes $U_{\mathcal{M}}$ a valid isometry. From now on, for the ease of notation, we use $M_{m,k}, \lambda_{m,k}^{B_R}$, and $A_{m,k}$ to denote the corresponding $n-$letter objects constructed for the codewords $A^n(m,k)$.

Step 1.2: Rotation Isometry

Although the above covering unitary aims to cover the source, it only does so for the reference system. To be able to apply the next step of packing, we wish to use the post-measured state as side information. This could be possible if the post-measured state also looks close to being product. For this, we employ a rotation unitary. A similar operation is discussed in (*Anshu et al.*, 2017b, Fact 6) with regards to classical-quantum states obtained post measurement. The construction below generalizes this to a coherent application of a measurement. More formally, for the expurgated

code $\mathcal{C}_{\mathcal{E}}$, we construct the states

$$|\hat{\sigma}\rangle^{B_R^n E^n MK} \triangleq \sum_{m\in\mathcal{M}'} \sum_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}} \frac{1}{\sqrt{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}} \frac{(I\otimes M_{m,k})}{\sqrt{\lambda_{m,k}}} \left|\psi_\rho^{\otimes n}\right\rangle^{B_R^n B^n} \otimes |m,k\rangle \text{ and}$$

$$|\tilde{\sigma}\rangle^{B_R^n B^n MK} \triangleq \sum_{m\in\mathcal{M}'} \sum_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}} \frac{1}{\sqrt{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}} \frac{(I\otimes \sqrt{A_{m,k}})}{\sqrt{\delta_{m,k}}} \left|\psi_\rho^{\otimes n}\right\rangle^{B_R^n B^n} \otimes |m,k\rangle\,, \tag{5.33}$$

where $\delta_{m,k} \triangleq \mathrm{Tr}\left\{A_{m,k}\rho^{B^{\otimes n}}\right\} = \gamma\,\mathrm{Tr}\left\{\tilde{\rho}_{m,k}^{B_R}\right\}$, and $\gamma \triangleq \frac{1-\varepsilon}{1+\eta}\frac{1}{|\mathcal{M}||\mathcal{K}|}$. For brevity in notation, we skip the sets in the summations over m or k when summations are performed over the codewords belonging to the set $\mathcal{I}_{\mathcal{E}}$ corresponding to the expurgated codebook $\mathcal{C}_{\mathcal{E}}$. Clearly, $|\hat{\sigma}\rangle^{B_R^n E^n MK}$ and $|\tilde{\sigma}\rangle^{B_R^n B^n MK}$ are valid states. Now to construct $U_{\mathcal{R}}$, consider the following lemma which upper bounds the fidelity.

Lemma V.19. *For any $\epsilon,\eta \in (0,1)$, there exists a collection of isometries $\{U_r(m,k) : \mathcal{H}_{B^n} \to \mathcal{H}_{E^n}\}$ and a collection of phases $\theta_{m,k}$ such that $F(|\hat{\sigma}\rangle^{B_R^n E^n MK}, (I_{B_R}\otimes U_{\mathcal{R}})\,|\tilde{\sigma}\rangle^{B_R^n B^n MK}) \geq 1-4\sqrt{\epsilon}$, for all sufficiently large n and all sufficiently small $\delta > 0$, where*

$$U_{\mathcal{R}} \triangleq \sum_{m\in\mathcal{M}'\cup\{M'\}} \sum_{k\in\mathcal{K}\cup\{K\}} e^{-i\theta_{m,k}} U_r(m,k)\otimes |m\rangle\langle m|_M \otimes |k\rangle\langle k|_K\,. \tag{5.34}$$

and $U_r(m,k) = I$ and $\theta_{m,k} = 0$ for all $(m,k) \in (\mathcal{M}\times\mathcal{K})$ such that $a^n(m,k) \notin \mathcal{C}_{\mathcal{E}}$.

Proof. The proof of the lemma follows using (5.29), (5.30) and from the Lemma D.1. For completeness, we detail the proof in Appendix D.2. □

Step 1.3: Packing Isometry

Observe that by coherently performing the covering and rotation operations, one can show that the source ρ_B can be successfully recovered by the quantum registers $\{|m,k\rangle\}$. This implies that the quantum states $\{|m,k\rangle\}$ can be used by the decoder to faithfully reconstruct the source as per Definition V.7. As a result, we would require a rate of $\frac{1}{n}\log M + \frac{1}{n}\log K$ which has to be greater than the Holevo information $\chi(\lambda_a^{B_R}, \hat{\rho}_a^{B_R})$, as constrained by Proposition V.17. However, we intend to further reduce the rate from this Holevo information to the coherent information provided in the statement of the theorem.

One can perhaps argue why we cannot simply release the information in the $\mathcal{H}_K$ system into

the environment (partial tracing)? But as expected for a purely quantum setup, this would lead to the protocol becoming incoherent. More precisely, the subsystem $\mathcal{H}_{E^n}$ that the encoder has in its possession is entangled with the subsystem $\mathcal{H}_K$, and tracing out the latter without decoupling the two systems would render the former in a mixed state. Once this entanglement is lost, the decoder would not be able to faithfully reconstruct the source by using such a (mixed) state.

Therefore, a major task here is to successfully decouple the system $\mathcal{H}_K$ before releasing it to the environment. To achieve this, we introduce the notion of coherent packing or coherent binning. This notion is built on the idea that the post-measured system present in $\mathcal{H}_{E^n}$ contains information about the quantum state $|k\rangle_K$, and hence, conditioned on the state $|m\rangle$, a copy of the state $|k\rangle$ can be recovered from the state present in subsystem $\mathcal{H}_{E^n}$, albeit with a small probability of error. Using this copy, we intend to decouple the existing copy of $|k\rangle$ from the latter. However, this new copy can erase (decouple) the original, but will itself still remain. Therefore, as will become evident in the sequel, we perform the process of erasing the information in $\mathcal{H}_K$ intrinsically without producing any additional copies.

Toward this, we employ the packing code consisting of the sub-POVMs $\{\Xi_k^{(m)}\}_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}}$, generated for the ensemble $\{\lambda_a^E, \tau_a^E\}$. We complete this sub-POVM for each $m \in \mathcal{M}'$ as

$$\Xi_K^{(m)} \triangleq I - \sum_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}} \Xi_k^{(m)}.$$

In addition, we also make use of Naimark's extension theorem (also provided in Lemma V.5). This lemma gives us a collection of orthogonal projectors $\{\Pi_k^{(m)}\}$ each acting on $\mathcal{H}_{E^n}\otimes\mathcal{H}_{\bar{E}}$, corresponding to the collection $\{\Xi_k^{(m)}\}_k$, such that

$$\mathrm{Tr}\Big\{\Pi_k^{(m)}(\tau_k^{(m)}\otimes|0\rangle\langle 0|_{\bar{E}})\Big\} = \mathrm{Tr}\Big\{\Xi_k^{(m)}\tau_k^{(m)}\Big\}, \tag{5.35}$$

for all $m \in \mathcal{M}'$ and $k \in \mathcal{I}_{\mathcal{E}}^{(m)} \cup \{K\}$, and $\dim \mathcal{H}_{\bar{E}} = K+1$. Finally, we define the packing unitary $U_{\mathcal{P}}$ as

$$U_{\mathcal{P}} \triangleq \sum_{m\in\mathcal{M}'} \left[\sum_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}\cup\{K\}} \Pi_k^{(m)} \otimes \left(\sum_{k'\in\mathcal{K}\cup\{K\}} e^{i\alpha_{k'}^{(m)}} \left|(k-k') \bmod (K+1)\right\rangle\left\langle k'\right|\right)\right] \otimes |m\rangle\langle m|$$

$$+ I_{E^n\bar{E}K} \otimes \left|M'\right\rangle\!\left\langle M'\right|, \quad (5.36)$$

where the phases $\{\alpha_k^{(m)}\}$ are introduced for later convenience, and will be specified in the sequel[1]. Note that by using $\Pi_k^{(m)}$ in the above definition, instead of $\Xi_k^{(m)}$ ensures that $U_{\mathcal{P}}U_{\mathcal{P}}^\dagger = I$, implying $U_{\mathcal{P}}$ is a valid unitary.

As a result, we can express the encoding CPTP map $\mathcal{N}_{\mathcal{E}}^{(n)}$ as

$$\left(I_{B_R^n} \otimes \mathcal{N}_{\mathcal{E}}^{(n)}\right)\left(\left|\psi_\rho^{\otimes n}\right\rangle\!\left\langle\psi_\rho^{\otimes n}\right|^{B_R^n B^n}\right)$$
$$\triangleq \mathrm{Tr}_{\bar{E}E^nK}\left((I \otimes U_{\mathcal{P}}U_{\mathcal{R}}U_{\mathcal{M}})\left|\psi_\rho^{\otimes n}\right\rangle\!\left\langle\psi_\rho^{\otimes n}\right|^{B_R^n B^n} \otimes |0\rangle\!\langle 0|_{\bar{E}} (I \otimes U_{\mathcal{P}}U_{\mathcal{R}}U_{\mathcal{M}})^\dagger\right). \quad (5.37)$$

The quantum state in $\mathcal{H}'_M$ is now sent to the decoder.

Step 1.4: Decoding Isometry:

The following decoding isometry is applied on the state in $\mathcal{H}'_M$:

$$U_{\mathcal{D}} \triangleq \sum_{m\in\mathcal{M}'}\left(\frac{1}{\sqrt{K'_m}}\sum_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}} e^{-i\beta_k^{(m)}}\left|a^n(m,k)\right\rangle\right)\langle m| + |a_0^n\rangle\langle M'|, \quad (5.38)$$

where the phases $\{\beta_k^{(m)}\}$ will be identified in the continuation. Observe that, to argue $U_{\mathcal{D}}$ is a valid isometric operation, we need the vectors $\{|a^n(m,k)\rangle\}$ to be distinct. By expurgating the codebook to generate $\mathcal{C}_{\mathcal{E}}$, and only using the codewords from $\mathcal{C}_{\mathcal{E}}$ ensures this distinctness. With the definitions of encoder and decoder, we move on to bounding the error incurred by the protocol (as defined in Definition 5.2).

5.6.4 Trace Distance

We begin by defining the following terms

$$|\omega\rangle^{B_R^n E^n \bar{E} A^n K} \triangleq (I \otimes U_{\mathcal{D}})(I \otimes U_{\mathcal{P}})(I \otimes U_{\mathcal{R}}U_{\mathcal{M}})\left|\psi_\rho^{\otimes n}\right\rangle^{B_R^n B^n}|0\rangle_{\bar{E}},$$
$$|\zeta\rangle^{B_R^n E^n A^n} \triangleq (W^{\otimes n} \otimes I)\,|\psi_\omega\rangle^{A_R^n A^n}, \quad (5.39)$$

[1]Moving forward, we implicitly assume the modulus operation rather than explicitly mentioning it for the purpose of brevity.

where[2] $|\psi_\omega\rangle^{A_R^n A^n}$ is the canonical purification of ω^{A^n}. Let

$$G \triangleq \|\omega^{B_R^n A^n} - \zeta^{B_R^n A^n}\|_1.$$

Following Definition 5.2, our objective now is to show G can be made arbitrarily small for all sufficiently large n for the code $\mathcal{C}_\mathcal{E}$.

Step 2.1: Closeness of $|\omega\rangle$ and $(I \otimes U_\mathcal{D})(I \otimes U_\mathcal{P})|\hat{\sigma}\rangle$:

Recall the definitions of $|\hat{\sigma}\rangle$ and $|\tilde{\sigma}\rangle$ from (5.33), and let $|\omega_1\rangle \triangleq (I \otimes U_\mathcal{R} U_\mathcal{M})|\psi_\rho\rangle^{B_R^n B^n}$ and $\varepsilon_1 \triangleq (1-\varepsilon)/(1+\eta)$. Consider

$$\begin{aligned}
&\sqrt{F\left(|\omega_1\rangle^{B_R^n E^n MK}, (I \otimes U_\mathcal{R})|\tilde{\sigma}\rangle^{B_R^n E^n MK}\right)} \\
&= \sum_{m,k} \frac{1}{\sqrt{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}} \frac{\langle\psi_\rho|(I \otimes U_r(m,k)\sqrt{A_{m,k}})^\dagger((I \otimes U_r(m,k)\sqrt{A_{m,k}})|\psi_\rho\rangle}{\sqrt{\delta_{m,k}}} \\
&= \frac{1}{\sqrt{1-\sqrt{\epsilon}}} \sum_{m,k} \frac{\sqrt{\varepsilon_1}}{|\mathcal{M}||\mathcal{K}|} \sqrt{\mathrm{Tr}\{\tilde{\rho}_{m,k}\}} \geq \frac{1}{\sqrt{1-\sqrt{\epsilon}}} \sum_{m,k} \frac{\sqrt{\varepsilon_1}}{|\mathcal{M}||\mathcal{K}|} \mathrm{Tr}\{\tilde{\rho}_{m,k}\} \geq \sqrt{\varepsilon_1}\sqrt{1-\sqrt{\epsilon}}(1-2\sqrt{\epsilon}),
\end{aligned} \tag{5.40}$$

where we note that there is no overlap between the term corresponding to $\sqrt{A_{a_0^n}} \otimes |M'\rangle_M \otimes |K\rangle_K$ of $|\omega_1\rangle^{B_R^n E^n MK}$ and the state $(U_\mathcal{R} \otimes I)|\tilde{\sigma}\rangle^{B_R^n E^n MK}$, and the last inequality follows from (5.29). Using Lemma V.3, and the inequality (5.40) , we get [3]

$$\begin{aligned}
&\left\|{\omega_1}^{B_R^n E^n MK} - (I \otimes U_\mathcal{R})\tilde{\sigma}^{B_R^n E^n MK}(I \otimes U_\mathcal{R})^\dagger\right\|_1 \\
&\qquad \leq 2\sqrt{1-(1-\sqrt{\epsilon})(1-2\sqrt{\epsilon})^2\left(1-\frac{\eta+\varepsilon}{1+\eta}\right)} \leq 2\sqrt{\frac{\eta+\varepsilon}{1+\eta}+5\sqrt{\epsilon}} \leq 6\sqrt[4]{\epsilon},
\end{aligned} \tag{5.41}$$

for all sufficiently large n and sufficiently small $\eta, \delta > 0$. Further, using the unitary invariance of trace distance, we get the closeness of the states:

$$\begin{aligned}
&\Big\|(I \otimes U_\mathcal{D}U_\mathcal{P}){\omega_1}^{B_R^n E^n MK} \otimes |0\rangle\langle 0|_{\bar{E}} (I \otimes U_\mathcal{D}U_\mathcal{P})^\dagger \\
&\qquad -(I \otimes U_\mathcal{D}U_\mathcal{P}U_\mathcal{R})\tilde{\sigma}^{B_R^n E^n MK} \otimes |0\rangle\langle 0|_{\bar{E}} (I \otimes U_\mathcal{D}U_\mathcal{P}U_\mathcal{R})^\dagger\Big\|_1 \leq 6\sqrt[4]{\epsilon}.
\end{aligned} \tag{5.42}$$

[2] For conciseness, we drop the $\otimes n$ from $|\psi_\rho^{\otimes n}\rangle^{B_R^n B^n}$ when understood from context.

[3] At times, the subspace notation is omitted when it is clear from the context.

Using Lemma V.19 and the fact that trace norm is invariant under isometric transformations, we have

$$\begin{aligned}
&\Big\|(I\otimes U_{\mathcal{D}}U_{\mathcal{P}})\hat{\sigma}^{B_R^nE^nMK}\otimes|0\rangle\langle 0|_{\bar{E}}\,(I\otimes U_{\mathcal{D}}U_{\mathcal{P}})^\dagger \\
&\qquad -(I\otimes U_{\mathcal{D}}U_{\mathcal{P}}U_{\mathcal{R}})\tilde{\sigma}^{B_R^nE^nMK}\otimes|0\rangle\langle 0|_{\bar{E}}\,(I\otimes U_{\mathcal{D}}U_{\mathcal{P}}U_{\mathcal{R}})^\dagger\Big\|_1 \\
&= \Big\|\hat{\sigma}-(I\otimes U_{\mathcal{R}})\tilde{\sigma}(I\otimes U_{\mathcal{R}})^\dagger\Big\|_1 \leq 2\sqrt{1-F\left(|\hat{\sigma}\rangle,(I\otimes U_{\mathcal{R}})\,|\tilde{\sigma}\rangle\right)} \leq 4\sqrt[4]{\epsilon}. \qquad (5.43)
\end{aligned}$$

Using triangle inequality and inequalities (5.42) and (5.43), we obtain

$$\Big\|\omega^{B_R^nE^n\bar{E}A^nK}-(I\otimes U_{\mathcal{D}}U_{\mathcal{P}})(\hat{\sigma}^{B_R^nE^nMK}\otimes|0\rangle\langle 0|_{\bar{E}})(I\otimes U_{\mathcal{D}}U_{\mathcal{P}})^\dagger\Big\|_1 \leq 10\sqrt[4]{\epsilon}, \qquad \text{(S-1)}$$

for all sufficiently large n and sufficiently small $\eta,\delta>0$, which concludes Step 2.1.

For the next step, define $|\hat{\zeta}\rangle^{B_R^nE^n\bar{E}MK}$ as

$$|\hat{\zeta}\rangle^{B_R^nE^n\bar{E}MK}\triangleq\sum_{m\in\mathcal{M}'}\sum_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}}\frac{1}{\sqrt{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}}e^{i\beta_k^{(m)}}\frac{(I\otimes M_{m,k})\,|\psi_\rho\rangle}{\sqrt{\lambda_{m,k}}}\otimes|m\rangle_M\otimes|0\rangle_K\otimes|0\rangle_{\bar{E}}\,, \qquad (5.44)$$

where the phases $\beta_k^{(m)}$ will be specified shortly. Observe that $|\hat{\zeta}\rangle^{B_R^nE^n\bar{E}MK}$ is a valid pure state due to (i) the distinctness of codewords in $\mathcal{C}_{\mathcal{E}}$ and (ii) the identity (5.17). Furthermore, in its definition, the information in the subsystem $\mathcal{H}_K$ is decoupled from the remaining subsystems. Since $U_{\mathcal{P}}$ acts on $\mathcal{H}_{E^n}\otimes\mathcal{H}_{\bar{E}}$, an additional pure ancilla is attached for appropriate comparisons. We aim to show that this state is close to the action of $(I\otimes U_{\mathcal{P}})$ on the state $|\hat{\sigma}\rangle\,|0\rangle_{\bar{E}}$.

Step 2.2: Closeness of $(I\otimes U_{\mathcal{P}})\,|\hat{\sigma}\rangle\,|0\rangle_{\bar{E}}$ and $|\hat{\zeta}\rangle$:

We begin by simplifying $(I\otimes U_{\mathcal{P}})\,|\hat{\sigma}\rangle\,|0\rangle_{\bar{E}}$ as

$$(I\otimes U_{\mathcal{P}})\,|\hat{\sigma}\rangle\,|0\rangle_{\bar{E}}=\sum_{m\in\mathcal{M}'}\sum_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}}\frac{1}{\sqrt{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}}e^{i\alpha_k^{(m)}}|\phi_k^{(m)}\rangle\otimes|m\rangle_M\,,$$

where

$$|\phi_k^{(m)}\rangle \triangleq \sum_{k' \in \mathcal{I}_{\mathcal{E}}^{(m)} \cup \{K\}} \frac{(I \otimes \Pi_{k'}^{(m)} M_{m,k}) |\psi_\rho\rangle |0\rangle_{\bar{E}}}{\sqrt{\lambda_{m,k}}} \otimes \left|k' - k\right\rangle_K, \text{ for all } k \in \mathcal{I}_{\mathcal{E}}^{(m)} \text{ and } m \in \mathcal{M}'. \quad (5.45)$$

Similarly, let

$$|\hat{\zeta}\rangle = \sum_{m,k} \frac{1}{\sqrt{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}} e^{i\beta_k^{(m)}} |\chi_k^{(m)}\rangle \otimes |m\rangle_M, \quad |\chi_k^{(m)}\rangle \triangleq \frac{(I \otimes M_{m,k}) |\psi_\rho\rangle |0\rangle_{\bar{E}}}{\sqrt{\lambda_{m,k}}} \otimes |0\rangle_K, \quad (5.46)$$

for all $m \in \mathcal{M}'$ and $k \in \mathcal{I}_{\mathcal{E}}^{(m)}$, and the phases $\{\beta_k^{(m)}\}$ are the same phases incorporated in the construction of the decoding isometry $U_{\mathcal{D}}$. Further, from (5.35), we know for all $m \in \mathcal{M}'$,

$$\frac{1}{K'_m} \sum_{k \in \mathcal{I}_{\mathcal{E}}^{(m)}} \langle \phi_k^{(m)} | \chi_k^{(m)} \rangle = \frac{1}{K'_m} \sum_{k \in \mathcal{I}_{\mathcal{E}}^{(m)}} \mathrm{Tr}\left\{ \Xi_k^{(m)} \tau_k^{(m)} \right\}. \quad (5.47)$$

Now the fidelity between $|\hat{\zeta}\rangle$ and $(I \otimes U_{\mathcal{P}}) |\hat{\sigma}\rangle |0\rangle_{\bar{E}}$ can be written as

$$\sqrt{F\left((I \otimes U_{\mathcal{P}}) |\hat{\sigma}\rangle |0\rangle_{\bar{E}}, |\hat{\zeta}\rangle \right)} = \frac{1}{M'} \left| \sum_{m \in \mathcal{M}'} \langle \phi_m | \chi_m \rangle \right|, \quad (5.48)$$

where, for all $m \in \mathcal{M}'$,

$$|\phi_m\rangle \triangleq c \sum_{k \in \mathcal{I}_{\mathcal{E}}^{(m)}} e^{i\alpha_k^{(m)}} |\phi_k^{(m)}\rangle \quad \text{and} \quad |\chi_m\rangle \triangleq c \sum_{k \in \mathcal{I}_{\mathcal{E}}^{(m)}} e^{i\beta_k^{(m)}} |\chi_k^{(m)}\rangle,$$

and $c \triangleq \sqrt{\frac{M'}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}}$. Toward a lower bound on the fidelity, we provide the following proposition.

Proposition V.20. *For any $\epsilon \in (0,1)$, there exists phases $\{\alpha_k^{(m)}\}$, and $\{\beta_k^{(m)}\}$ such that*

$$\left| \frac{1}{M'} \sum_{m \in \mathcal{M}'} \langle \phi_m | \chi_m \rangle \right| \geq 1 - 2\sqrt{\epsilon}, \quad (5.49)$$

for all sufficiently small $\delta > 0$ and all sufficiently large n.

Proof. The proof is provided in Appendix D.4. □

Observe that, using the relation in Lemma V.3, and the result of Proposition V.20 and (5.48), we

get

$$\left\|(I\otimes U_{\mathcal{P}})\hat{\sigma}\otimes|0\rangle\langle 0|_{\bar{E}}\,(I\otimes U_{\mathcal{P}})^{\dagger}-\hat{\zeta}\right\|_1\leq 2\sqrt{1-F\left((U_{\mathcal{P}}\otimes I)\,|\hat{\sigma}\rangle\,|0\rangle_{\bar{E}}\,,|\hat{\zeta}\rangle\right)}\leq 4\sqrt[4]{\epsilon}, \tag{S-2}$$

for all sufficiently large n, and sufficiently small $\eta,\delta>0$. Observe that $|\hat{\zeta}\rangle^{B_R^nE^nMK}=|\hat{\zeta}\rangle^{B_R^nE^nM}\otimes|0\rangle_{K\bar{E}}$, and hence $|\hat{\zeta}\rangle^{B_R^nE^nM}$ remains pure after partial tracing over the subsystem $\mathcal{H}_K\otimes\mathcal{H}_{\bar{E}}$. Finally, we are left with showing the closeness of the state $(I\otimes U_{\mathcal{D}})|\hat{\zeta}\rangle^{B_R^nE^nM}$ with the state $|\zeta\rangle^{B_R^nE^nA^n}$.

Step 2.3: Closeness of $(I\otimes U_{\mathcal{D}})|\hat{\zeta}\rangle^{B_R^nE^nM}$ and $|\zeta\rangle^{B_R^nE^nA^n}$:

We begin by defining σ^{A^n} as

$$\sigma^{A^n}\triangleq\mathrm{Tr}_{B_R^nE^n}\{(I\otimes U_{\mathcal{D}})\hat{\zeta}^{B_R^nE^nM}(I\otimes U_{\mathcal{D}})^{\dagger}\},$$

and perform the simplification

$$\begin{aligned}\sigma^{A^n}&=U_{\mathcal{D}}\Bigg(\sum_{m,m'}\sum_{k,k'}\frac{1}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}e^{-i(\beta_k^{(m)}-\beta_{k'}^{(m')})}\frac{\mathrm{Tr}\left(M_{mk}\rho^BM_{m'k'}^{\dagger}\right)}{\sqrt{\lambda_{mk}\lambda_{m'k'}}}|m\rangle\langle m'|\Bigg)U_{\mathcal{D}}^{\dagger}\\&=U_{\mathcal{D}}\left(\sum_m\frac{K'_m}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}\,|m\rangle\langle m|\right)U_{\mathcal{D}}^{\dagger}=\sum_m\frac{K'_m}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}\,|b^n(m)\rangle\langle b^n(m)|^{A^n}\,,\end{aligned} \tag{5.50}$$

where the second equality uses (5.17) and the crucial condition that the codebook $\mathcal{C}_{\mathcal{E}}$ obtained after expurgation is distinct, and last equality defines $|b^n(m)\rangle^{A^n}$ as

$$|b^n(m)\rangle^{A^n}\triangleq\frac{1}{\sqrt{K'_m}}\sum_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}}e^{-i\beta_k^{(m)}}\,|a^n(m,k)\rangle^{A^n}\,, \tag{5.51}$$

for all $m\in\mathcal{M}'$. This implies, we can write the canonical purification of σ^{A^n} as

$$\begin{aligned}|\psi_\sigma\rangle^{A_R^nA^n}\triangleq(I_{A_R^n}\otimes\sqrt{\sigma^{A^n}})\,|\Gamma^{\otimes n}\rangle^{A_R^nA^n}&=\sum_m\sqrt{\frac{K'_m}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}}\,|b^n(m)\rangle^{A_R^n}\otimes|b^n(m)\rangle^{A^n}\\&=(I\otimes U_{\mathcal{D}}^{A^n})\sum_m\sqrt{\frac{K'_m}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}}\,|b^n(m)\rangle^{A_R^n}\otimes|m\rangle^M\,,\end{aligned} \tag{5.52}$$

where the first equality follows by defining $|b^n(m)\rangle^{A_R^n}\triangleq(I_{A_R^n}\otimes\langle b^n(m)|^{A^n})\,|\Gamma^{\otimes n}\rangle^{A_R^nA^n}$. Using the

relation from (5.18) and definition (5.51), we can write

$$W^{\otimes n}|b^n(m)\rangle^{A_R^n} = \frac{1}{\sqrt{K'_m}} \sum_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}} e^{i\beta_k^{(m)}} W^{\otimes n}|a^n(m,k)\rangle^{A_R^n} = \frac{1}{\sqrt{K'_m}} \sum_{k\in\mathcal{I}_{\mathcal{E}}^{(m)}} e^{i\beta_k^{(m)}} \frac{(I_{B_R}\otimes M_{m,k})\left|\psi_\rho^{\otimes n}\right\rangle^{B_R^n B^n}}{\sqrt{\lambda_{m,k}^A}},$$

for all $m \in \mathcal{M}'$, which gives

$$\begin{aligned}
(W^{\otimes n}\otimes I_{A^n})\,|\psi_\sigma\rangle^{A_R^n A^n} &= (I\otimes U_{\mathcal{D}}^{A^n}) \sum_m \sqrt{\frac{K'_m}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}} W^{\otimes n}\,|b^n(m)\rangle^{A_R^n} \otimes |m\rangle^M \\
&= (I\otimes U_{\mathcal{D}}^{A^n}) \sum_{m,k} \frac{1}{\sqrt{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}} e^{i\beta_k^{(m)}} \frac{(I_{B_R}\otimes M_{m,k})\,\left|\psi_\rho^{\otimes n}\right\rangle^{B_R^n B^n}}{\sqrt{\lambda_{m,k}^A}} \otimes |m\rangle^M \\
&= (I\otimes U_{\mathcal{D}}^{A^n})|\hat{\zeta}\rangle^{B_R^n E^n M}, \qquad (5.53)
\end{aligned}$$

where the last equality follows from the definition of $|\hat{\zeta}\rangle^{B_R^n E^n M}$ in (5.44). Observing that

$$|\zeta\rangle^{B_R^n E^n A^n} = (W^{\otimes n}\otimes I_{A^n})\,|\psi_\omega\rangle^{A_R^n A^n},$$

we are left with showing the closeness of $|\psi_\sigma\rangle^{A_R^n A^n}$ and $|\psi_\omega\rangle^{A_R^n A^n}$. Using (S-1) and (S-2), the triangle inequality, the monotonicity of trace distance, and identification of appropriate purifications, we obtain for all sufficiently large n and sufficiently small $\eta, \delta > 0$,

$$\|\omega^{A^n} - \sigma^{A^n}\|_1 \leq \|\omega^{B_R^n E^n A^n} - (I\otimes U_{\mathcal{D}})\hat{\zeta}^{B_R^n E^n M}(I\otimes U_{\mathcal{D}})^\dagger\|_1 \leq 14\sqrt[4]{\epsilon}.$$

This implies,

$$\begin{aligned}
\left\|(I\otimes U_{\mathcal{D}}^{A^n})\hat{\zeta}^{B_R^n E^n M}(I\otimes U_{\mathcal{D}}^{A^n})^\dagger - \zeta^{B_R^n E^n A^n}\right\|_1 &\overset{a}{=} \left\|{\psi_\sigma}^{A^n A_R^n} - {\psi_\omega}^{A^n A_R^n}\right\|_1 \\
&\overset{b}{\leq} 2\sqrt{1 - F(|\psi_\sigma\rangle^{A^n A_R^n}, |\psi_\omega\rangle^{A^n A_R^n})} \\
&\overset{c}{\leq} 2\sqrt{\|\omega^{A^n} - \sigma^{A^n}\|_1} \leq 2\sqrt{14\sqrt[4]{\epsilon}} \leq 8\sqrt[8]{\epsilon}, \qquad \text{(S-3)}
\end{aligned}$$

for all sufficiently large n and sufficiently small $\eta, \delta > 0$, where (a) follows from the isometric invariance of trace distance, (b) uses Lemma V.3, and (c) uses Lemma V.4, which concludes this step.

In summary, combining results of (S-0), (S-1), (S-2) and (S-3), we have showed that there exist a code $\mathcal{C}$ satisfying $G \leq 14\sqrt[4]{\epsilon} + 8\sqrt[8]{\epsilon}$ with the following rate constraints:

$$S(\lambda_a^A) > \frac{1}{n}(\log M + \log K) > \chi(\lambda_a^{B_R}, \hat{\rho}_a^{B_R}), \quad \frac{1}{n}\log K < \chi(\{\lambda_a^E, \tau_a^E\}), \quad \frac{1}{n}\log M \geq 0, \quad \frac{1}{n}\log K \geq 0,$$

for all sufficiently large n and sufficiently small $\eta, \delta > 0$, where we have also included the necessary non-negativity constraints. Eliminating $\frac{1}{n}\log K$ using Fourier-Motzkin elimination *Ziegler* (2012) gives

$$\frac{1}{n}\log M > \chi(\lambda_a^{B_R}, \hat{\rho}_a^{B_R}) - \chi(\{\lambda_a^E, \tau_a^E\}), \quad \text{and} \quad \frac{1}{n}\log M \geq 0,$$

where we remove the redundant constraints. This completes the proof.

Remark V.21 (Zero performance rate). The coherent information $I_c(\mathcal{N}_W, \rho^{A_R})$ is negative when the Holevo information quantities are such that $\chi(\lambda_a^{B_R}, \hat{\rho}_a^{B_R}) < \chi(\{\lambda_a^E, \tau_a^E\})$. The asymptotic performance limit for these situations is zero, according to the statement of the theorem. We must therefore demonstrate that a rate of zero is feasible. To put it another way, we must construct a protocol (see Definition (V.7)) that satisfies (5.2) with a rate that can be made arbitrarily close to zero. The constraints imposed by the preceding proof are still met if we select $M = 1$ and $\frac{1}{n}\log K = \chi(\lambda_a^{B_R}, \hat{\rho}_a^{B_R}) + \delta_0 < \chi(\{\lambda_a^E, \tau_a^E\})$, while achieving a rate of $log(2)/n$, for a sufficiently small δ_0. This rate $(1/n)$ can be made arbitrarily close to zero (i.e., smaller than the provided ϵ) for any given ϵ, for all sufficiently large n and sufficiently small $\eta, \delta > 0$. Similarly, when the coherent information $I_c(\mathcal{N}_W, \rho^{A_R})$ is exactly zero, i.e., $\chi(\lambda_a^{B_R}, \hat{\rho}_a^{B_R}) = \chi(\{\lambda_a^E, \tau_a^E\})$, we choose M and K such that $\frac{1}{n}\log M = 2\delta_0$, and $\frac{1}{n}\log K = \chi(\{\lambda_a^E, \tau_a^E\}) - \delta_0$. This gives a rate of $2\delta_0$ which can be again made arbitrarily close to zero. Therefore, even though the coherent information is not necessarily positive, the rate in the theorem can still be achieved.

5.7 Proof of Converse

Let R be an achievable rate. Then from Definition V.8, given a triple $(\rho_B, \mathcal{H}_A, \mathcal{N}_W)$, for all $\epsilon > 0$, and all sufficiently large n, there exists (n, Θ) lossy compression protocol with an encoding

CPTP map $\mathcal{N}_{\mathcal{E}}^{(n)}$ and a decoding CPTP map $\mathcal{N}_{\mathcal{D}}^{(n)}$ that satisfies the following constraints:

$$c_0 : \frac{1}{n}\log\Theta \leq R+\epsilon, \quad \text{and} \quad c_1 : \|\omega^{B_R^n A^n} - \upsilon^{B_R^n A^n}\|_1 \leq \epsilon, \tag{5.54}$$

where $\omega^{B_R^n A^n} \triangleq (I\otimes\mathcal{N}_{\mathcal{D}}^{(n)})(I\otimes\mathcal{N}_{\mathcal{E}}^{(n)})(|\psi_\rho\rangle^{B_R^n B^n})$,

$$\upsilon^{B_R^n A^n} = \mathrm{Tr}_{E^n}\left\{\upsilon^{B_R^n A^n E^n}\right\} \triangleq \mathrm{Tr}_{E^n}\left\{(W^{\otimes n}\otimes I)\Psi_\omega^{A_R^n A^n}(W^{\otimes n}\otimes I)^\dagger\right\},$$

and $|\psi_\omega\rangle^{A^n A_R{}^n}$ is the canonical purification of ω^{A^n}, and W is the Stinespring's dilation of the CPTP map $\mathcal{N}_W$. Let $\omega^{A_R^n} \triangleq \mathrm{Tr}_{A^n}(\Psi_\omega^{A^n A_R^n})$.

Step 1: Quantum Data Processing Inequality: Let M denote the quantum state at the output of the encoder. Let $V_{\mathcal{E}}^{(n)} : \mathcal{H}_{B^n} \to \mathcal{H}_M \otimes \mathcal{H}_{\tilde{E}_1}$ and $V_{\mathcal{D}}^{(n)} : \mathcal{H}_M \to \mathcal{H}_{A^n}\otimes\mathcal{H}_{\tilde{E}_2}$ be Stinesping dilations of encoding and decoding maps $\mathcal{N}_{\mathcal{E}}^{(n)}$ and $\mathcal{N}_{\mathcal{D}}^{(n)}$, respectively, such that $\dim(\mathcal{H}_{\tilde{E}_1}) \geq \dim(\mathcal{H}_M)$ and $\dim(\mathcal{H}_{\tilde{E}_2}) \geq \dim(\mathcal{H}_{A^n})$, as shown in Figure 5.3(a). Let

$$\begin{aligned}
\omega_1{}^{B_R^n M\tilde{E}_1} &\triangleq (I_{B_R^n}\otimes V_{\mathcal{E}}^{(n)})(\Psi_\rho^{B_R^n B^n})(I_{B_R^n}\otimes V_{\mathcal{E}}^{(n)})^\dagger \\
\omega^{B_R^n \tilde{E}_1\tilde{E}_2 A^n} &\triangleq (I_{B_R^n\tilde{E}_1}\otimes V_{\mathcal{D}}^{(n)})(\omega_1{}^{B_R^n\tilde{E}_1 M})(I_{B_R^n\tilde{E}_1}\otimes V_{\mathcal{D}}^{(n)})^\dagger.
\end{aligned} \tag{5.55}$$

Let $|\psi_{\omega_1}\rangle^{M_R M}$ denote the canonical purification of the quantum state ω_1^M. Let $W_{\mathcal{E}}^{(n)} : \mathcal{H}_{M_R} \to \mathcal{H}_{B_R^n}\otimes\mathcal{H}_{\tilde{E}_1}$ denote the posterior reference isometry (see Definition V.1) of $V_{\mathcal{E}}^{(n)}$ with respect to ω_1^M, as shown in Figure 5.3(b). Moreover, let $W_{\mathcal{D}}^{(n)} : \mathcal{H}_{A_R^n} \to \mathcal{H}_{M_R}\otimes\mathcal{H}_{\tilde{E}_2}$ denote the posterior reference isometry of $V_{\mathcal{D}}^{(n)}$ with respect to w^{A^n}, as shown in Figure 5.3(c). Let $\mathcal{N}_{W_{\mathcal{E}}}(\cdot) = \mathrm{Tr}_{\tilde{E}_1}(W_{\mathcal{E}}^{(n)}\cdot(W_{\mathcal{E}}^{(n)})^\dagger)$ and $\mathcal{N}_{W_{\mathcal{D}}}(\cdot) = \mathrm{Tr}_{\tilde{E}_2}(W_{\mathcal{D}}^{(n)}\cdot(W_{\mathcal{D}}^{(n)})^\dagger)$ be the induced CPTP maps. Let

$$\tilde{\omega}_1^{M_R\tilde{E}_2 A^n} \triangleq (W_{\mathcal{D}}^{(n)}\otimes I_{A^n})(\Psi_\omega^{A_R^n A^n})(W_{\mathcal{D}}^{(n)}\otimes I_{A^n})^\dagger.$$

Using the quantum data processing inequality for coherent information (*Wilde*, 2011, Theorem 11.3.2), we obtain

$$I_c(\mathcal{N}_{W_{\mathcal{D}}}, \omega^{A_R^n}) \geq I_c(\mathcal{N}_{W_{\mathcal{E}}}\circ\mathcal{N}_{W_{\mathcal{D}}}, \omega^{A_R^n}).$$

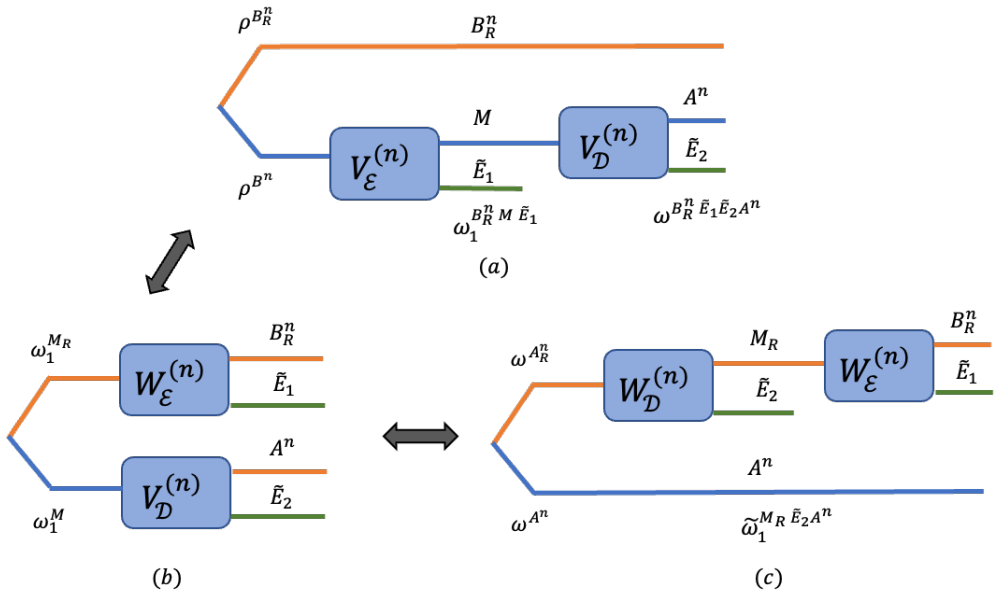

Figure 5.3: Lossy quantum source coding protocol and the associated CPTP maps and their Stinespring dilations.

Expanding the coherent information in terms of Von Neuman entropy, we get

$$S(M_R)_{\tilde{\omega}_1} - S(\tilde{E}_2)_{\tilde{\omega}_1} \geq S(B_R^n)_\omega - S(\tilde{E}_1\tilde{E}_2)_\omega,$$

which implies that

$$S(M)_{\omega_1} \geq S(B_R^n)_\omega - S(B_R^n A^n)_\omega. \tag{5.56}$$

Step 2: Implication of the constraints c_0 and c_1: Consider the following sequence of inequalities:

$$\begin{aligned}
nR &\geq \log\Theta - n\epsilon \geq S(M)_{\omega_1} - n\epsilon && (5.57)\\
&\overset{a}{\geq} S(B_R^n)_\omega - S(B_R^n, A^n)_\omega - n\epsilon && (5.58)\\
&\overset{b}{\geq} S(B_R^n)_\omega - S(B_R^n, A^n)_\upsilon - n\epsilon - n\tilde{\epsilon}_1 && (5.59)\\
&= S(B_R^n)_\omega - S(E^n)_\upsilon - n\epsilon - n\tilde{\epsilon}_1 && (5.60)\\
&\overset{c}{\geq} S(B_R^n)_\omega - \sum_{i=1}^{n} S(E_i)_\upsilon - n\epsilon - n\tilde{\epsilon}_1 && (5.61)
\end{aligned}$$

$$\stackrel{d}{=} \sum_{i=1}^{n} S(B_{Ri})_{\omega} - \sum_{i=1}^{n} S(E_i)_{\upsilon} - n\epsilon - n\tilde{\epsilon}_1 \tag{5.62}$$

$$\stackrel{e}{\geq} \sum_{i=1}^{n} S(B_{Ri})_{\upsilon} - \sum_{i=1}^{n} S(E_i)_{\upsilon} - n\epsilon - n\tilde{\epsilon}_1 - n\tilde{\epsilon}_2 \tag{5.63}$$

$$\stackrel{f}{=} \sum_{i=1}^{n} I_c(\mathcal{N}_W, \omega^{A_{Ri}}) - n\epsilon - n\tilde{\epsilon}_1 - n\tilde{\epsilon}_2 \tag{5.64}$$

$$\stackrel{g}{\geq} n \min_{\rho^{A_R} \in \mathcal{D}(\mathcal{H}_{A_R}) : \|\rho^{B_R} - \mathcal{N}_W(\rho^{A_R})\|_1 \leq \epsilon} I_c(\mathcal{N}_W, \rho^{A_R}) - n\epsilon - n\tilde{\epsilon}_1 - n\tilde{\epsilon}_2, \tag{5.65}$$

where the inequalities are argued as follows. (a) follows from (5.56). (b) follows from the condition c_1 and the Fannes-Audenaert inequality (*Wilde*, 2011, Theorem 11.10.1) by defining $\tilde{\epsilon}_1 \triangleq \epsilon \log |\mathcal{H}_A||\mathcal{H}_B| + h_b(\epsilon)$. (c) follows from the subadditivity of entropy. (d) follows from the memoylessness of the quantum source. (e) follows from condition c_2 and the Fannes-Audenaert inequality (*Wilde*, 2011, Theorem 11.10.1), where condition

$$c_2 : \|\omega^{B_{Ri}} - \upsilon^{B_{Ri}}\|_1 \leq \epsilon, \quad \forall\, 1 \leq i \leq n,$$

is implied by c_1 using the monotonicity of trace distance with respect to partial trace. $\tilde{\epsilon}_2$ is defined as $\tilde{\epsilon}_2 \triangleq \epsilon \log |\mathcal{H}_B| + h_b(\epsilon)$. (f) follows from the fact that $\upsilon^{B_{Ri}E_i} = W\omega^{A_{Ri}}W^{\dagger}$. (g) follows from condition c_2 which can also be stated as

$$c_2 : \|\rho^{B_R} - \mathcal{N}_W(\omega^{A_{Ri}})\|_1 \leq \epsilon, \quad \forall\, 1 \leq i \leq n,$$

and the fact that coherent information is continuous, and the constraint set is closed and bounded. The continuity follows from the following arguments: for the fixed CPTP map $\mathcal{N}_W$, let a function $f : \mathcal{D}(\mathcal{H}_{A_R}) \to \mathbb{R}$ be defined as $f(\rho^{A_R}) = I_c(\mathcal{N}_W, \rho^{A_R})$. One can establish the continuity of f for a fixed $\mathcal{N}_W$ by writing $I_c(\mathcal{N}_W, \rho^{A_R}) = S(B_R)_{W\rho^{A_R}W^{\dagger}} - S(E)_{W\rho^{A_R}W^{\dagger}}$, and using the Fannes–Audenaert Inequality (*Wilde*, 2011, Theorem 11.10.2), where W is the Stinespring's extension of the given CPTP map $\mathcal{N}_W$.

Step 3: Continuity Argument: We have shown that

$$R \in \bigcap_{\epsilon > 0} \mathcal{I}_{\epsilon},$$

where we have defined for all $\epsilon \geq 0$,

$$\mathcal{I}_\epsilon \triangleq \left\{ R : \exists\, \rho^{A_R} \in \mathcal{S}_\epsilon(\rho^B, \mathcal{N}_W) \text{ such that } R \geq I_c(\mathcal{N}_W, \rho^{A_R}) - g(\epsilon) \right\}, \tag{5.66}$$

and

$$\mathcal{S}_\epsilon(\rho^B, \mathcal{N}_W) \triangleq \left\{ \rho^{A_R} \in \mathcal{D}(\mathcal{H}_{A_R}) : \|\mathcal{N}_W(\rho^{A_R}) - \rho^{B_R}\|_1 \leq \epsilon \right\}, \tag{5.67}$$

$g(\epsilon) \triangleq \epsilon + \tilde{\epsilon}_1 + \tilde{\epsilon}_2$. Condition c_2 ensures that the set $\mathcal{S}_\epsilon$ is non-empty for all $\epsilon > 0$. Now, by arguing continuity of $\mathcal{I}_\epsilon$ at $\epsilon = 0$, we obtain the desired result.

Lemma V.22. *For the above definitions of $\mathcal{S}_\epsilon$ and $\mathcal{I}_\epsilon$, we have $\mathcal{S}_0(\rho^B, \mathcal{N}_W)$ non-empty, and*

$$\mathcal{I}_0 = \bigcap_{\epsilon > 0} \mathcal{I}_\epsilon.$$

Proof. This is a standard argument used in the literature *Dueck* (1981); *Ahlswede and Cai* (2006); *Cuff* (2013). A proof is provided in Appendix D.5 for completeness. □

This completes the proof.

5.8 Conclusion

In this work, we explored a new formulation of the lossy quantum source coding problem. The two ingredients that make our formulation different from the standard rate-distortion problem are (i) the usage of a global error criterion to measure the quality of reconstruction, and (ii) the notion of a posterior reference channel defined as a CPTP map acting on the reference of the reconstruction to produce the reference of the source. Instead of a single-letter distortion function, a global error criterion measures the error incurred by using the given single-letter posterior channel. The given channel characterizes the nature of the loss incurred in the encoding and decoding operations.

As our main result, we provide a single-letter characterization of the asymptotic performance limit of this source coding problem using the minimal coherent information of the posterior reference map, where the minimization is over all reconstructions. Even though the formulation uses a global error criterion, it sheds light on an "optimistic" perspective of the lossy source coding theory. In

this regard, our results provide the missing duality pair of the quantum channel coding problem, and also broadens the framework of performing lossy quantum source compression. Investigation of this formulation to other variants of lossy source coding problem can be an interesting research avenue to pursue. Similarly, it would be interesting to explore other techniques of establishing the achievability and converse of this limit.

Part II

Classical Communication Over Quantum Channels

CHAPTER VI

Computation over a Classical-Quantum Multiple Access Channel

6.1 Introduction

Early research in quantum state discrimination led to the investigation of the information carrying capacity of quantum states. Suppose Alice - a sender - can prepare any one of the states in the collection $\{\rho_x \in \mathcal{D}(\mathcal{H}_Y) : x \in \mathcal{X}\}$ and Bob - the receiver - has to rely on a measurement to infer the label x of the state, then what is the largest sub-collection $\mathcal{C} \subseteq \mathcal{X}$ of states that Bob can distinguish perfectly? Studying this question in a Shannon-theoretic sense, Schumacher, Westmoreland *Schumacher and Westmoreland* (1997) and Holevo *Holevo* (1998) characterized the exponential growth of this sub-collection, thereby characterizing the capacity of a classical-quantum (CQ) point-to-point (PTP) channel. In the following years, generalizations of this question with multiple senders and/or receivers have been studied with an aim of characterizing the corresponding information carrying capacity of quantum states in network scenarios *Winter* (2001).

In this chapter, we consider the problem of computing functions of information sources over a CQ multiple access channel (MAC). Let $(\rho_{x_1x_2} \in \mathcal{D}(\mathcal{H}_Y) : (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2)$ model a CQ-MAC. Sender j - the party having access to the choice of label $x_j \in \mathcal{X}_j$ - observes a classical information stream $S_{jt} \in \mathcal{S}_j : t \geq 1$. The pairs $(S_{1t}, S_{2t}) : t \geq 1$ are independent and identically distributed (IID) with a single-letter joint distribution $\mathbb{W}_{S_1S_2}$. The receiver, who is provided with the prepared quantum state, intends to reconstruct a specific function $f(S_1, S_2)$ of the information observed by the senders. The question of interest is under what conditions, specified in terms of the CQ-MAC, $\mathbb{W}_{S_1S_2}$ and f, can the receiver reconstruct the desired function losslessly?

The conventional approach to characterizing sufficient conditions for this problem relies on

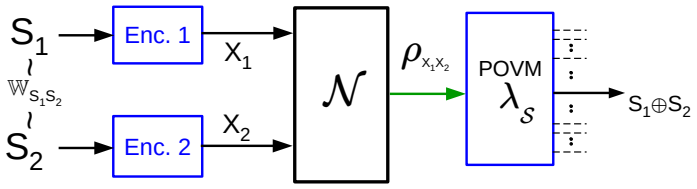

Figure 6.1: CQ-MAC used for computing sum of classical sources.

enabling the receiver reconstruct the pair of classical source sequences. Since the receiver is only interested in recovering the bivariate function f, and not the pair, this approach can be strictly sub-optimal. Can we exploit this and design a more efficient communication strategy, thereby weakening the set of sufficient conditions? In this work, we present one such communication strategy for a general CQ-MAC that is more efficient than the conventional approach. This strategy is based on asymptotically good random nested coset codes. We analyze its performance and derive new sufficient conditions for a general problem instance and identify examples for which the derived conditions are strictly weaker.

Our findings here are built on the ideas developed in the classical setting. Focusing on a source coding formulation, i.e. a noiseless MAC, Körner and Marton *Korner and Marton* (1979) devised an ingenious coding technique that enabled the receiver recover the sum of the sources without recovering either source. In *Nazer and Gastpar* (2007), the linearity of the Körner-Marton (KM) source coding map was further exploited to enable the receiver recover the sum of the sources using *only* the *sum* of the KM indices, not even requiring the pair. Leveraging this observation and focusing on the subclass of additive MACs, specific MAC channel coding techniques are devised in *Nazer and Gastpar* (2007) that enabled the receiver recover the sum of two channel coding message indices.

The techniques of *Korner and Marton* (1979), *Nazer and Gastpar* (2007) are instances of a broader framework of coding strategies. Decoding functions of sources or channel inputs efficiently require codes endowed with algebraic closure properties. To emphasize, the conventional approach of deriving inner bounds/achievable rate region by analyzing expected performance of IID random codes is incapable of yielding performance limits - capacity or rate-distortion regions as the case may be- in network communication scenarios. To improve upon this, it is necessary to analyze the expected performance of random codes endowed with algebraic closure properties. In a series of

works *Pradhan et al.* (2021), an information theoretic study of the latter codes has been carried out yielding new inner bounds for multiple network communication scenarios.

In this work, we embark on developing these ideas in the CQ setup. After having provided the problem statement in Sec. 6.2, we focus on a simplified CQ MAC and illustrate the core idea of our coding scheme. The latter relies on developing a *nested coset code* (NCC) based communication scheme for a CQ PTP channel and analyzing its performance (Sec. 6.4). Leveraging this building block, we design and analyze the performance of an NCC-based coding scheme for computing sum over a general CQ-MAC (Sec. 6.6). Going further we generalize this idea for computing arbitrary functions over a general CQ-MAC.

6.2 Preliminaries and Problem Statement

We supplement the notation in *Wilde* (2013a) with the following. For positive integer n, $[n] \triangleq \{1, \cdots, n\}$. For a Hilbert space $\mathcal{H}$, $\mathcal{L}(\mathcal{H}), \mathcal{P}(\mathcal{H})$ and $\mathcal{D}(\mathcal{H})$ denote the collection of linear, positive and density operators acting on $\mathcal{H}$, respectively. The von Neumann entropy of a density operator $\rho \in \mathcal{D}(\mathcal{H})$ is denoted by $S(\rho)$. Given any ensemble $\{p_i, \rho_i\}_{i\in[1,m]}$, the Holevo information *Holevo* (2012) is denoted as $\chi\big(\{p_i; \rho_i\}\big)$. A POVM acting on $\mathcal{H}$ is a collection $\lambda_{\mathcal{X}} \triangleq \{\lambda_x\}_{x\in\mathcal{X}}$ of positive operators that form a resolution of the identity: $\sum_{x\in\mathcal{X}} \lambda_x = I$, where $\mathcal{X}$ is a finite set. We employ an underline notation to aggregate objects of similar type. For example, $\underline{s}$ denotes (s_1, s_2), $\underline{x}^n$ denotes (x_1^n, x_2^n), $\underline{\mathcal{S}}$ denotes the Cartesian product $\mathcal{S}_1 \times \mathcal{S}_2$.

Consider a (generic) *CQ-MAC* $(\rho_{x_1x_2} \in \mathcal{D}(\mathcal{H}_Y) : (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2)$ specified through (i) finite sets $\mathcal{X}_j : j \in [2]$, (ii) Hilbert space $\mathcal{H}_Y$, and (iii) a collection $(\rho_{x_1,x_2} \in \mathcal{D}(\mathcal{H}_Y) : (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2)$ of density operators. This CQ-MAC is employed to enable the receiver reconstruct a bivariate function of the classical information streams observed by the senders. Let $\mathcal{S}_1, \mathcal{S}_2$ be finite sets and $(S_1, S_2) \in \mathcal{S}_1 \times \mathcal{S}_2$ distributed with PMF $\mathbb{W}_{S_1S_2}$ models the pair of information sources observed at the encoders. Specifically, sender j observes the sequence $S_{jt} \in \mathcal{S}_j : t \geq 1$ and the sequence $(S_{1t}, S_{2t}) : t \geq 1$ are IID with single-letter PMF $\mathbb{W}_{S_1S_2}$. The receiver aims to recover the sequence $f(S_{1t}, S_{2t}) : t \geq 1$ losslessly, where $f : \mathcal{S}_1 \times \mathcal{S}_2 \to \mathcal{R}$ is a specified function.

A CQ-MAC code $c_f = (n, e_1, e_2, \lambda_{\mathcal{R}^n})$ of block-length n for recovering f consists of two encoders maps $e_j : \mathcal{S}^n \to \mathcal{X}_j^n : j \in [2]$, and a POVM $\lambda_{\mathcal{R}^n} = \{\lambda_{r^n} \in \mathcal{P}(\mathcal{H}_Y^{\otimes n}) : r^n \in \mathcal{R}^n\}$. The average error

probability of the CQ-MAC code c_f is

$$\overline{\xi}(c_f) \quad = \quad 1 - \sum_{\underline{s}^n : f(\underline{s}^n) = r^n} \mathbb{W}^n_{S_1 S_2}(s_1^n, s_2^n) \operatorname{Tr}\left(\lambda_{r^n} \rho_{c,\underline{s}^n}^{\otimes n}\right)$$

where $\rho_{c,\underline{s}^n}^{\otimes n} \triangleq \otimes_{i=1}^n \rho_{x_{1i}(s_1^n) x_{2i}(s_2^n)}$, where $e_j(s_j^n) = x_{j1}(s_j^n), x_{j2}(s_j^n), \cdots, x_{jn}(s_j^n)$ for $j \in [2]$.

A function f of the sources $\mathbb{W}_{S_1 S_2}$ is said to be reconstructible over a CQ-MAC if for $\epsilon > 0$, $\exists$ a sequence $c_f^{(n)} = (n, e_1^{(n)}, e_2^{(n)}, \lambda_{\mathcal{R}^n}) : n \geq 1$ such that $\lim_{n \to \infty} \overline{\xi}(c_f^{(n)}) = 0$.

In this article, we are concerned with the problem of characterizing sufficient conditions under which a function f of the sources $\mathbb{W}_{S_1 S_2}$ is reconstructible over a generic MAC $(\rho_{x_1 x_2} \in \mathcal{D}(\mathcal{H}_Y) : (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2)$. One of our findings - Proposition VI.10 - provides a characterization of sufficient conditions in terms of a computable function of the associated objects- density operators that characterize the CQ-MAC, function f and the source distribution $\mathbb{W}_{S_1 S_2}$.

As we shall see, the specific problem of computing sum of sources will play an important role in our work. In this case, $\mathcal{S} = \mathcal{S}_1 = \mathcal{S}_2 = \mathcal{F}_q$ is a finite field with q elements and the receiver aims to reconstruct $f(S_1, S_2) = S_1 \oplus_q S_2$ where $\oplus_q$ denotes addition in $\mathcal{F}_q$. A CQ-MAC code $c_\oplus = (n, e_1, e_2, \lambda_{\mathcal{S}^n})$ of block-length n for recovering the sum consists of two encoders maps $e_j : \mathcal{S}^n \to \mathcal{X}_j^n : j \in [2]$, and a POVM $\lambda_{\mathcal{S}^n} = \{\lambda_{s^n} \in \mathcal{P}(\mathcal{H}_Y^{\otimes n}) : s^n \in \mathcal{S}^n\}$.

Restricting f to a sum, we say the sum of sources $\mathbb{W}_{S_1 S_2}$ over field $\mathcal{F}_q$ is reconstructible over a CQ-MAC if $\mathcal{S}_1 = \mathcal{S}_2 = \mathcal{F}_q$ and the function $f(S_1, S_2) = S_1 \oplus_q S_2$ is reconstructible over the CQ-MAC. The problem of characterizing sufficient conditions under which a sum of sources is reconstructible over a CQ-MAC plays an important role in this work. One of our findings - Theorem VI.9 - provides a computable characterization of a set of sufficient conditions under which a sum of sources is reconstructible over a CQ-MAC. As the reader will note, this encapsulates the central element of our characterization in Proposition VI.10.

We also formalize the notions of a CQ-PTP and CQ-MAC codes for communicating uniform messages. A CQ-MAC code $c_{\underline{m}} = (n, \mathcal{I}_1, \mathcal{I}_2, e_1, e_2, \lambda_{\underline{\mathcal{I}}})$ for a CQ-MAC $(\rho_{\underline{x}} \in \mathcal{D}(\mathcal{H}_Y) : \underline{x} \in \underline{\mathcal{X}})$ consists of (i) index sets $\mathcal{I}_j : j \in [2]$, (ii) encoder maps $e_j : \mathcal{I}_j \to \mathcal{X}_j^n : j \in [2]$ and a decoding POVM $\lambda_{\underline{\mathcal{I}}} = \{\lambda_{\underline{m}} \in \mathcal{P}(\mathcal{H}_Y^{\otimes n}) : \underline{m} \in \mathcal{I}_1 \times \mathcal{I}_2\}$. For $\underline{m} \in \mathcal{I}_1 \times \mathcal{I}_2$, we let $\rho_{c,\underline{m}}^{\otimes n} \triangleq \otimes_{i=1}^n \rho_{x_{1i}, x_{2i}}$ where $e_j(m_j) = x_{j1} \cdots x_{jn}$ for $j \in [2]$.

A CQ-PTP code $c_m = (n, \mathcal{I}, e, \lambda_{\mathcal{I}})$ for a CQ-PTP $(\rho_x \in \mathcal{D}(\mathcal{H}_Y) : x \in \mathcal{X})$ consists of (i) an index

set $\mathcal{I}$, (ii) and encoder map $e : \mathcal{I} \to \mathcal{X}^n$ and a decoding POVM $\lambda_{\mathcal{I}} = \{\lambda_m \in \mathcal{P}(\mathcal{H}_Y^{\otimes n}) : m \in \mathcal{I}\}$. For $m \in \mathcal{I}$, we let $\rho_{c,m}^{\otimes n} \triangleq \otimes_{i=1}^n \rho_{x_i}$ where $e(m) = x_1 \cdots x_n$.

6.3 The Central Idea

Let us consider the specific problem of reconstructing the sum of sources each taking values in $\mathcal{S} = \mathcal{F}_q$. We begin by reviewing the KM coding scheme for the case of a noiseless classical MAC. It was shown in *Korner and Marton* (1979) the existence of linear code with a parity matrix $H \in \mathcal{S}^{l\times n}$ and decoder map $d : \mathcal{F}_q^l \to \mathcal{S}^n$ such that $\sum_{\underline{s}^n \in \underline{\mathcal{S}}^n} \mathbb{W}_{\underline{S}}^n(\underline{s}^n) \mathbb{1}_{\{d(Hs_1^n \oplus_q Hs_2^n) \neq s_1^n \oplus_q s_2^n\}} \leq \epsilon$, for any $\epsilon > 0$, and sufficiently large n, so long as $\frac{l \log_2 q}{n} > H(S_1 \oplus_q S_2)$. This implies that a receiver equipped with the decoding map d can recover the sum if it possesses the sum $M_1^l \oplus_q M_2^l$ of the Körner-Marton indices $M_j^l = HS_j^l : j \in [2]$.

We are therefore led to building an efficient CQ-MAC coding scheme that enables the receiver only reconstruct the sum of the two message indices. Indeed, if the two senders send the KM indices to such a CQ-MAC channel code and the receiver employs the above source decoder d on the decoded sum of the KM indices, it can recover the sum of sources. To illustrate the design of the desired CQ-MAC channel code, let us consider a CQ-MAC $(\rho_{x_1x_2} \in \mathcal{D}(\mathcal{H}_Y) : (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2)$ wherein $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{F}_q$ and the collection $\rho_{\underline{x}} : \underline{x} \in \underline{\mathcal{X}}$ satisfies $\rho_{x_1x_2} = \rho_{\hat{x}_1\hat{x}_2}$ whenever $x_1 \oplus_q x_2 = \hat{x}_1 \oplus_q \hat{x}_2$. Consider a CQ-PTP $(\mathcal{X} = \mathcal{F}_q, \sigma_u : u \in \mathcal{X})$ where $\sigma_u = \rho_{x_1 \oplus x_2}$ for any (x_1, x_2) satisfying $x_1 \oplus_q x_2 = u$. Suppose we are able to communicate over this CQ-PTP via a *linear CQ-PTP code* $\mathcal{C} \subseteq \mathcal{X}^n$. Specifically, suppose there exists a generator matrix $G \in \mathcal{X}^{l\times n}$ and a POVM $\{\lambda_{m^l} : m^l \in \mathcal{F}_q^l\}$ so that $1 - q^{-l} \sum_{m^l} \mathrm{Tr}\big(\lambda_{m^l} \sigma_{m^lG}^{\otimes n}\big) \leq \epsilon$. for any $\epsilon > 0$ and sufficiently large n, where $\sigma_{m^lG}^{\otimes} = \sigma_{x_1} \otimes \cdots \otimes \sigma_{x_n}$ where $m^lG = x^n$. We can then use this linear CQ-PTP code as our desired CQ-MAC channel code. Indeed, observe that, suppose both senders employ this same linear CQ-PTP code, then sender j maps its KM index $m_j^l = Hs_j^n$ to the channel codeword $x_j^n = m_j^lG$. Observe that the structure of the CQ-MAC implies $\rho_{x_1^n,x_2^n}^{\otimes n} = \sigma_{x_1^n \oplus x_2^n}^{\otimes n} = \sigma_{(m_1^l \oplus_q m_2^l)G}^{\otimes}$. If the receiver employs the POVM $\{\lambda_{m^l} : m^l \in \mathcal{F}_q^l\}$ designed for the CQ-PTP, it ends up decoding the sum of the KM indices $m_1^l \oplus_q m_2^l$, and consequently, recover the sum of the sources.

A careful analysis of the above idea reveals that two MAC channel codes employed by the encoders do not 'blow up' when added, is crucial to the efficiency of the above scheme. A linear

code being algebraically closed enables this. However, the codewords of a random linear code are uniformly distributed and cannot achieve the capacity of an arbitrary classical PTP channel, let alone a CQ-PTP channel. We are therefore forced to enlarge a linear code to identify sufficiently many codewords of the desired empirical distribution. We are thus led to a *nested coset code* (NCC)*Padakandla and Pradhan*. A NCC comprises of cosets of a coarse linear code within a fine code. Within each coset, we can identify a codeword of the desired empirical distribution. We choose as many cosets as the number of messages. Analogous to our illustration above where we chose a linear code that achieves the capacity of the CQ-PTP ($\mathcal{X} = \mathcal{F}_q, \sigma_u : u \in \mathcal{X}$), our first step (Sec. 6.4) is to design a NCC with its POVM that can achieve capacity of an arbitrary CQ PTP. Our second step is to endow both senders with this same NCC and analyze decoding the sum of the messages. This gets us to our next challenge - How do we analyze decoding their message sum, for a general CQ-MAC $\rho_{\underline{x}} : \underline{x} \in \underline{\mathcal{X}}$ for which $x_1 \oplus_q x_2 = \hat{x}_1 \oplus_q \hat{x}_2$ does *not* necessarily imply $\rho_{x_1 x_2} = \rho_{\hat{x}_1 \hat{x}_2}$. In Sec. 6.5, we address this challenge, leverage our findings in Sec. 6.4 and generalize the idea for a general CQ-MAC.

6.4 Nested Coset Codes Achieve Capacity of CQ-PTP

We begin by formalizing the structure of an NCC.

Definition VI.1. An $(n, k, l, g_I, g_{O/I}, b^n)$ NCC built over a finite field $\mathcal{V} = \mathcal{F}_q$ comprises of (i) generator matrices $g_I \in \mathcal{V}^{k\times n}$, $g_{O/I} \in \mathcal{V}^{l\times n}$ (ii) a bias vector b^n, an encoder map $e : \mathcal{V}^l \to \mathcal{V}^k$. We let $v^n(a, m) = ag_I \oplus_q mg_{O/I} \oplus_q b^n : (a, m) \in \mathcal{V}^k \times \mathcal{V}^l$ denote elements in the range space of the generator matrix $[g_I^t \; g_{O/I}^t]^t$.

Definition VI.2. A CQ-PTP code $(n, \mathcal{I} = \mathcal{F}_q^l, e, \lambda_{\mathcal{I}})$ is an NCC CQ-PTP if there exists an $(n, k, g_I, g_{O/I}, b^n)$ NCC such that $e(m) \in \{u^n(a, m) : a \in \mathcal{F}_q^k\}$ for all $m \in \mathcal{F}_q^l$.

Theorem VI.3. *Given a CQ-PTP $(\rho_v \in \mathcal{D}(\mathcal{H}_Y) : v \in \mathcal{F}_q)$ and a PMF p_V on $\mathcal{F}_q$, $\epsilon > 0$ there exists a CQ-PTP code $c = (n, \mathcal{I} = \mathcal{F}_q^l, e, \lambda_{\mathcal{I}})$ such that (i) $q^{-l} \sum_{m\in[\mathcal{I}]} \sum_{\hat{m}\neq[\mathcal{I}]\setminus\{m\}} \mathrm{Tr}\big(\lambda_{\hat{m}} \rho_{c,m}^{\otimes n}\big) \leq \epsilon$, (ii) $c = (n, \mathcal{I} = \mathcal{F}_q^l, e, \lambda_{\mathcal{I}})$ is a NCC CQ-PTP, (iii) $\frac{k \log_2 q}{n} > \log_2 q - H(V)$ and $\frac{(k+l)\log_2 q}{n} < \log_2 q - H(V) + \chi(\{p_v, \rho_v\})$ for all n sufficiently large.*

Proof. In order to achieve a rate $R = \chi(\{p_v, \rho_v\})$, the standard approach is to pick 2^{nR} codewords uniformly and independently from $T_\delta^n(p_V)$. However, the resulting code is not algebraically closed.

On the other hand, if we pick a random generator matrix $G \in \mathcal{F}_q^{l\times n}$, with $l = \frac{nR}{\log_2 q}$, whose entries from $\mathcal{F}_q$ are IID uniform, then its range space - the resulting collection of 2^{nR} codewords - are uniformly distributed and pairwise independent but not p_V-typical.

To satisfy the dual requirements of algebraically closure and p_V-typicality, we observe the following. If a collection of q^k codewords are uniformly distributed in $\mathcal{F}_q^n$ and pairwise independent, as we found the range space of G to be, then the expected number of codewords that are p_V-typical is $\frac{q^k}{q^n}|T_\delta^n(p_V)| = \exp\{n\log_2 q\left(\frac{k}{n} - \left[1 - \frac{H(V)}{\log_2 q}\right]\right)\}$. This indicates that if we pick a generator matrix $G_I \in \mathcal{F}^{k\times n}$ with entries uniformly distributed and IID, such that $\frac{k}{n} > 1 - \frac{H(V)}{\log_2 q}$, then its range space will contain codewords that are p_V-typical. The latter codewords can be used for communication.

Each coset of $G_I \in \mathcal{F}^{k\times n}$ where $\frac{k}{n} > 1 - \frac{H(V)}{\log_2 q}$ will play an analogous role as a single codeword in a conventional IID random code. Just as we pick 2^{nR} of the latter, we consider 2^{nR} cosets of G_I within a larger linear code with generator matrix $G = \begin{bmatrix} G_I \\ G_{O/I} \end{bmatrix} \in \mathcal{F}_q^{(k+l)\times n}$ with $l = \frac{nR}{\log_2 q}$. The messages index the 2^{nR} cosets of G_I. A predetermined element in each coset that is p_V-typical is the assigned codeword for the message and chosen for communication.[1] A formal proof we provide below has two parts - error probability analysis for a generic fixed code followed by an upper bound on the latter via code randomization.

Upper bound on Error Prob. for a generic fixed code : Consider a generic NCC $(n, k, l, g_I, g_{O/I}, b^n)$ with its range space $v^n(a, m) = ag_I \oplus_q mg_{O/I} \oplus_q b^n : (a, m) \in \mathcal{V}^k \times \mathcal{V}^l$. We shall use this and define a CQ-PTP code $(n, \mathcal{I} = \mathcal{F}_q^l, e, \lambda_{\mathcal{I}})$ that is an NCC CQ-PTP. Towards that end, let $\theta(m) \triangleq \sum_{a\in\mathcal{V}^k} \mathbb{1}_{\{v^n(a,m)\in T_\delta^n(p_V)\}}$ and

$$s(m) \triangleq \begin{cases} \{a \in \mathcal{V}^K : v^n(a, m) \in T_\delta^n(p_V)\} & \text{if } \theta(m) \geq 1 \\ \{0^k\} & \text{if } \theta(m) = 0 \end{cases}$$

for each $m \in \mathcal{V}^l$. For $m \in \mathcal{V}^l$, a predetermined element $a_m \in s(m)$ is chosen. On receiving message $m \in \mathcal{V}^l$, the encoder prepares the quantum state $\rho_m^{\otimes n} \triangleq \rho_{v^n(a_m,m)}^{\otimes n} \triangleq \otimes_{i=1}^n \rho_{v_i(a_m,m)}$ and is communicated. The encoding map e is therefore determined via the collection $(a_m \in s(m) : m \in \mathcal{V}^l)$.

[1] The reader is encouraged to relate to the bounds stated in theorem statement and induced bounds on the rate of communication $\frac{l\log_2 q}{n}$.

Towards specifying the decoding POVM let $\rho_v = \sum_{y\in\mathcal{Y}} p_{Y|V}(y|v) \left|e_{y|v}\right\rangle \left\langle e_{y|v}\right|$ be a spectral decomposition for $v \in \mathcal{V}$. We let $p_{VY} \triangleq p_V p_{Y|V}$. For any $v^n \in \mathcal{V}^n$, let π_{v^n} be the conditional typical projector as in (*Wilde*, 2013a, Defn. 15.2.4) with respect to the ensemble $\{\rho_v : v \in \mathcal{V}\}$ and distribution p_V. Similarly, let π_ρ be the (unconditional) typical projector of the state $\rho \triangleq \sum_{v\in\mathcal{V}} p_V(v)\rho_v$ as defined in (*Wilde*, 2013a, Defn. 15.1.3). For $(a, m) \in \mathcal{V}^k \times \mathcal{V}^l$, we let $\pi_{a,m} \triangleq \pi_{v^n(a,m)} \mathbb{1}_{\{v^n(a,m)\in T^n_\delta(p_V)\}}$. We let $\lambda_{\mathcal{I}} \triangleq \{\sum_{a\in\mathcal{V}^k} \lambda_{a,m} : m \in \mathcal{I} = \mathcal{V}^l, \lambda_{-1}\}$, where

$$\lambda_{a,m} \triangleq \Big(\sum_{\hat{a}\in\mathcal{V}^k}\sum_{\hat{m}\in\mathcal{V}^l} \gamma_{\hat{a},\hat{m}}\Big)^{-1/2} \gamma_{a,m} \Big(\sum_{\tilde{a}\in\mathcal{V}^k}\sum_{\tilde{m}\in\mathcal{V}^l} \gamma_{\tilde{a},\tilde{m}}\Big)^{-1/2} \tag{6.1}$$

$\lambda_{-1} \triangleq I - \sum_{m\in\mathcal{V}^l}\sum_{a\in\mathcal{V}^k} \lambda_{a,m}$ and $\gamma_{a,m} \triangleq \pi_\rho \pi_{a,m} \pi_\rho$. Since $0 \le \gamma_{a,m} \le I$, we have $0 \le \lambda_{a,m} \le I$. The latter lower bound implies $\lambda_{\mathcal{I}} \subseteq \mathcal{P}(\mathcal{H})$. The same lower bound coupled with the definition of the generalized inverse implies $I \ge \sum_{a\in\mathcal{V}^k}\sum_{m\in\mathcal{V}^l} \lambda_{a,m} \ge 0$. We thus have $0 \le \lambda_{-1} \le I$. It can now be verified that $\lambda_{\mathcal{I}}$ is a POVM. In essence, the elements of this POVM is identical to the standard POVMs except the POVM elements corresponding to a coset have been added together. Indeed, since each coset corresponds to one message, there is no need to disambiguate within the coset.

We have thus associated an NCC $(n, k, l, g_I, g_{O/I}, b^n)$ and a collection $(a_m \in s(m) : m \in \mathcal{V}^l)$ with a CQ-PTP code. The error probability of this code is

$$q^{-l}\sum_{m\in\mathcal{I}} \mathrm{tr}((I - \sum_{a\in\mathcal{V}^k} \lambda_{a,m})\rho_m^{\otimes n}) \le q^{-l}\sum_{m\in\mathcal{I}} \mathrm{tr}((I - \lambda_{a_m,m})\rho_m^{\otimes n}). \tag{6.2}$$

Denoting event $\mathscr{E} = \{\theta(m) < 1\}$, its complement $\mathscr{E}^c$ and the associated indicator functions $\mathbb{1}_{\mathscr{E}}, \mathbb{1}_{\mathscr{E}^c}$ respectively, a generic term in the RHS of the above sum satisfies

$$\begin{aligned}
&\mathrm{tr}((I - \lambda_{a_m,m})\rho_m^{\otimes n})\mathbb{1}_{\mathscr{E}^c} + \mathrm{tr}((I - \lambda_{a_m,m})\rho_m^{\otimes n})\mathbb{1}_{\mathscr{E}} \\
&\le \quad \mathbb{1}_{\mathscr{E}^c} + \sum_{i=1}^{3} T_{2i}, \text{ where } T_{21} = 2\mathrm{tr}((I - \gamma_{a_m,m})\rho_m^{\otimes n})\mathbb{1}_{\mathscr{E}}, \\
T_{22} &= 4\sum_{\hat{a}\neq a_m} \mathrm{tr}(\gamma_{\hat{a},m}\rho_m^{\otimes n})\mathbb{1}_{\mathscr{E}}, \; T_{23} = 4\sum_{\hat{m}\neq m}\sum_{\tilde{a}} \mathrm{tr}(\gamma_{\tilde{a},\hat{m}}\rho_m^{\otimes n})\mathbb{1}_{\mathscr{E}}
\end{aligned}$$

where we have used Hayashi-Nagaoka inequality *Hayashi and Nagaoka* (2003).

Distribution of the Random Code : The objects $g_I \in \mathcal{V}^{k\times n}, g_{O/I} \in \mathcal{V}^{l\times n}, b^n \in \mathcal{V}^n$ and the collection $(a_m \in s(m) : m \in \mathcal{V}^l)$ specify an NCC CQ-PTP code unambiguously. A distribution

for a random code is therefore specified through a distribution of these objects. We let upper case letters denote the associated random objects, and obtain

$$\mathcal{P}\begin{pmatrix} G_I = g_I, G_{O/I} = g_{O/I} \\ B^n = b^n, A_m = a_m \colon m \in S(m) \end{pmatrix} = q^{-(k+l+1)n} \prod_{m \in \mathcal{V}^l} \frac{1}{\Theta(m)},$$

and analyze the expectation of $\mathscr{E}$ and the terms $T_{2i}; i \in [1,3]$ in regards to the above random code. We begin by $\mathbb{E}_{\mathcal{P}}[\mathscr{E}] = \mathcal{P}(\sum_{a \in \mathcal{V}^k} \mathbb{1}_{\{V^n(a,m) \in T_\delta^n(p_V)\}} < 1)$. For this, we provide the following proposition.

Proposition VI.4. *There exist $\epsilon_{T_1}(\delta), \delta_{T_1}(\delta)$, such that for all sufficiently small δ and sufficiently large n, we have $\mathbb{E}_{\mathcal{P}}[\mathscr{E}] \leq \epsilon_{T_1}(\delta)$, if $\frac{k}{n} \geq \log q - H(V) + \delta_S$, where $\epsilon_S, \delta_S \searrow 0$ as $\delta \searrow 0$.*

Proof. The proof follows from Appendix B of *Padakandla and Pradhan* (2017) with the identification of $\mathcal{S} = \phi$. □

We now consider T_{21}. An upper bound on T_{21} is derived by obtaining a lower bound on $\mathrm{Tr}(\lambda_{a_m,m} \rho_m^{\otimes n})$. This follows by an argument that is colloquially referred to as 'pinching'. Lemma E.1 in Appendix E.1 proves the existence of $\lambda > 0$ such that $\mathbb{E}_{\mathcal{P}}\{T_{21}\} \leq \exp\{-n\lambda\delta^2\}$ for sufficiently large n. We now analyze $\mathbb{E}_{\mathcal{P}}[T_{22}]$. Denoting the event

$$\mathcal{J} \triangleq \begin{Bmatrix} \Theta(m) \geq 1, V^n(\hat{a}, \hat{m}) = \hat{x}^n \\ A_m = d, V^n(d,m) = x^n \end{Bmatrix} \subseteq \mathcal{K} \triangleq \begin{Bmatrix} V^n(\hat{a}, \hat{m}) = \hat{x}^n \\ V^n(d,m) = x^n \end{Bmatrix} \tag{6.3}$$

we perform the following steps.

$$\begin{aligned} \mathbb{E}_{\mathcal{P}}[T_{22}] &= \sum_{\hat{a} \in \mathcal{V}^k} \mathbb{E}_{\mathcal{P}}[\mathrm{tr}(\Gamma_{\hat{a},m} \rho_m^{\otimes n}) \mathbb{1}_{\{\theta(m) \geq 1\}} \mathbb{1}_{\{\hat{a} \neq A_m\}}] \\ &= \sum_{d \in \mathcal{V}^k} \sum_{\hat{a} \in \mathcal{V}^k} \sum_{x^n \in T_\delta^n(p_V)} \sum_{\hat{x}^n \in \mathcal{V}^n} \mathbb{E}\left[\mathrm{tr}(\Gamma_{\hat{a},m} \rho_m^{\otimes n}) \mathbb{1}_{\hat{a} \neq d} \mathbb{1}_{\mathcal{J}}\right] \\ &= \sum_{d \in \mathcal{V}^k} \sum_{\hat{a} \neq d} \sum_{x^n \in T_\delta^n(p_V)} \sum_{\hat{x}^n \in \mathcal{V}^n} \mathbb{E}\left[\mathrm{tr}(\Gamma_{\hat{a},m} \rho_m^{\otimes n}) \mathbb{1}_{\mathcal{J}}\right] \end{aligned}$$

where the restriction of the summation x^n to $T_\delta^n(p_V)$ is valid since $S(m) > 1$ forces the choice

$A_m \in S(m)$ such that $V^n(A_m, m) \in T_\delta^n(p_V)$. Going further, we have

$$\begin{aligned}
\mathbb{E}_{\mathcal{P}}[T_{22}] &= \sum_{\substack{d,\hat{a}\in\mathcal{V}^k \\ \hat{a}\neq d}} \sum_{x^n\in T_\delta^n(p_V)} \sum_{\hat{x}^n\in T_\delta^n(p_V)} \mathbb{E}\left[\operatorname{tr}(\pi_\rho \pi_{\hat{x}^n} \pi_\rho \rho_{x^n}^{\otimes n}) \mathbb{1}_{\mathcal{J}}\right] \\
&= \sum_{d,\hat{a}:\hat{a}\neq d} \sum_{x^n\in T_\delta^n(p_V)} \sum_{\hat{x}^n\in T_\delta^n(p_V)} \operatorname{tr}(\pi_\rho \pi_{\hat{x}^n} \pi_\rho \rho_{x^n}^{\otimes n}) \mathcal{P}(\mathcal{J}) \\
&\overset{(a)}{\leq} \sum_{d,\hat{a}:\hat{a}\neq d} \sum_{\hat{x}^n\in T_\delta^n(p_V)} \operatorname{tr}(\pi_{\hat{x}^n} \pi_\rho) \mathcal{P}(\mathcal{J}) 2^{-n[S(\rho)-H(p_V)+\epsilon_V]} \\
&\overset{(b)}{\leq} \sum_{d,\hat{a}:\hat{a}\neq d} \sum_{\hat{x}^n\in T_\delta^n(p_V)} \operatorname{tr}(\pi_{\hat{x}^n} \pi_\rho) \mathcal{P}(\mathcal{K}) 2^{-n[S(\rho)-H(p_V)+\epsilon_V]} \\
&\overset{(c)}{=} \sum_{d,\hat{a}:\hat{a}\neq d} \sum_{\hat{x}^n\in T_\delta^n(p_V)} \operatorname{tr}(\pi_{\hat{x}^n} \pi_\rho) \frac{1}{q^{2n}} 2^{-n[S(\rho)-H(p_V)+\epsilon_V]} \\
&\overset{(d)}{\leq} 2^{-n\left[\chi(\{p_V;\rho_v\})+\epsilon_V-2H(p_V)-\frac{2k}{n}\log q+2\log q\right]}
\end{aligned} \tag{6.4}$$

where the restriction of the summation $\hat{x}^n$ to $T_\delta^n(p_V)$ follows from the fact that $\pi_{\hat{x}^n}$ is the zero projector if $\hat{x}^n \notin T_\delta^n(p_V)$, (a) follows from the operator inequality

$$\sum_{x^n\in T_\delta(p_V)} \pi_\rho \rho_{x^n} \pi_\rho \leq 2^{n(H(p_V)+\epsilon_V(\delta))} \pi_\rho \rho^{\otimes n} \pi_\rho \leq 2^{n(H(p_V)+\epsilon_V(\delta)-S(\rho))} \pi_\rho$$

found in (*Wilde*, 2017, Eqn. 20.34, 15.20), (b) follows from Def. 6.3, (c) follows from pairwise independence of the distinct codewords, and (d) follows from $\pi_\rho \leq I$ and (*Wilde*, 2013a, Eqn. 15.77) and $\epsilon_V(\delta) \searrow 0$ as $\delta \searrow 0$. We now derive an upper bound on $\mathbb{E}_{\mathcal{P}}[T_{23}]$. We have

$$\begin{aligned}
\mathbb{E}_{\mathcal{P}}[T_{23}] &= \sum_{d,\hat{a}\in\mathcal{V}^k} \sum_{\hat{m}\neq m} \sum_{\substack{x^n,\hat{x}^n\in \\ T_\delta^n(p_V)}} \mathbb{E}\Big[\operatorname{tr}(\pi_\rho \Pi_{\hat{a},\hat{m}} \pi_\rho \rho_{A_m,m}^{\otimes n}) \mathbb{1}_{\mathcal{J}}\Big] \\
&= \sum_{d,\hat{a}\in\mathcal{V}^k} \sum_{\hat{m}\neq m} \sum_{x^n,\hat{x}^n\in T_\delta^n(p_V)} \operatorname{tr}(\pi_{\hat{x}^n} \pi_\rho \rho_{x^n}^{\otimes n} \pi_\rho) \mathcal{P}(\mathcal{J}) \\
&\leq \sum_{d,\hat{a}\in\mathcal{V}^k} \sum_{\hat{m}\neq m} \sum_{\hat{x}^n\in T_\delta^n(p_V)} \operatorname{tr}(\pi_{\hat{x}^n} \pi_\rho) \mathcal{P}(\mathcal{J}) 2^{-n[S(\rho)-H(p_V)+\epsilon_V]} \\
&\leq \sum_{d,\hat{a}\in\mathcal{V}^k} \sum_{\hat{m}\neq m} \sum_{\hat{x}^n\in T_\delta^n(p_V)} \operatorname{tr}(\pi_{\hat{x}^n} \pi_\rho) \mathcal{P}(\mathcal{K}) 2^{-n[S(\rho)-H(p_V)+\epsilon_V]} \\
&= \sum_{d,\hat{a}\in\mathcal{V}^k} \sum_{\hat{m}\neq m} \sum_{\hat{x}^n\in T_\delta^n(p_V)} \operatorname{tr}(\pi_{\hat{x}^n} \pi_\rho) \frac{1}{q^{2n}} 2^{-n[S(\rho)-H(p_V)+\epsilon_V]}
\end{aligned}$$

$$\leq\ 2^{-n\left[\chi(\{p_V;\rho_v\})+2\log_2 q-2H(p_V)-\frac{2k+l}{n}\log_2 q+\epsilon_V\right]}$$

where the inequalities above uses similar reasoning as in (6.4).

We have therefore obtained three bounds $\frac{k}{n} > 1-\frac{H(p_V)}{\log_2 q}$, $\frac{2k}{n} < 2+\frac{\chi(\{p_V;\rho_v\})-2H(p_V)}{\log_2 q}$, $\frac{2k+l}{n} < 2+\frac{\chi(\{p_V;\rho_v\})-2H(p_V)}{\log_2 q}$. A rate of $\chi(\{p_V;\rho_v\})-\epsilon$ is achievable by choosing $\frac{k}{n} = 1-\frac{H(p_V)}{\log_2 q}+\frac{\epsilon}{2}$, $\frac{l}{n} = \frac{\chi(\{p_V;\rho_v\})-\epsilon\log_2\sqrt{q}}{\log_2 q}$ thus completing the proof. □

6.5 Decoding Sum over CQ-MAC

Throughout this section, the source alphabets $\mathcal{S} \triangleq \mathcal{S}_1 = \mathcal{S}_2 = \mathcal{F}_q$ is a finite field with q elements and the receiver intends to reconstruct the sum $f(S_1,S_2) = S_1 \oplus_q S_2$ of the sources. As discussed in Sec. 6.3, we propose a 'separation based' coding scheme consisting of a Körner Marton (KM) source code followed by a CQ MAC channel code designed to communicate the sum of the message indices input at the channel code encoders. The focus of this section is to design, analyze and thereby characterize performance of the latter CQ MAC channel code tasked to communicate the sum of messages. Towards that end, we begin with a definition.

Definition VI.5. Let $\mathcal{V} = \mathcal{F}_q$ be a finite field and $(\rho_{x_1x_2} \in \mathcal{D}(\mathcal{H}_Y) : (x_1,x_2) \in \mathcal{X}_1\times\mathcal{X}_2)$ be a CQ-MAC. A CQ-MAC code $c_{m\oplus} = (n, \mathcal{I}_1 = \mathcal{I}_2 = \mathcal{F}_q^l, e_1, e_2, \lambda_{[q^l]})$ of block-length n for recovering $\mathcal{F}_q-$sum of messages consists of two encoders maps $e_j : \mathcal{V}^l \to \mathcal{X}_j^n : j \in [2]$, and a POVM $\lambda_{q^l} = \{\lambda_m \in \mathcal{P}(\mathcal{H}_Y^{\otimes n}) : m \in \mathcal{V}^l\}$.

An $\mathcal{F}_q-$*message-sum* rate $R>0$ is achievable if given any sequence $l(n) \in \mathbb{N} : n \in \mathbb{N}$ such that $\limsup_{n\to\infty}\frac{l(n)\log q}{n} < R$, any sequence $p^{(n)}_{M_1M_2}$ of PMFs on $\mathcal{F}_q^{l(n)}\times\mathcal{F}_q^{l(n)}$, there exists a CQ-MAC code $c^{(n)}_{m\oplus} = (n, \mathcal{I} = \mathcal{I}_1 = \mathcal{I}_2 = \mathcal{F}_q^{l(n)}, e_1^{(n)}, e_2^{(n)}, \lambda_{\mathcal{I}})$ of block-length n for recovering $\mathcal{F}_q^{l(n)}-$sum of messages such that for every $\delta>0$, have

$$\lim_{n\to\infty}\bar{\xi}(c^{(n)}_{m\oplus}) \quad = \quad \lim_{n\to\infty} 1-\sum_{\substack{(m_1,m_2)\\ \in\mathcal{I}_1\times\mathcal{I}_2}} p_{M_1}p_{M_2}(m_1,m_2)\,\mathrm{Tr}\big(\lambda_{m_1\oplus m_2}\rho^{\otimes n}_{c,\underline{m}}\big) = 0$$

where $\rho^{\otimes n}_{c,\underline{m}} \triangleq \otimes_{i=1}^n \rho_{x_{1i}(m_1)x_{2i}(m_2)}$, where $e_j(m_j) = x_{j1}(m_j), x_{j2}(m_j), \cdots, x_{jn}(m_j)$ for $j\in[2]$. The closure of the set of all achievable $\mathcal{F}_q-$message-sum rates is the message-sum capacity of the CQ-MAC.

From our discussion in Sec. 6.3 and the above definition, a road map for characterizing sufficient conditions for computing the sum over a CQ-MAC must be evident. Referring back to Sec. 6.3, we note that is joint PMF $\mathbb{W}_{S_1S_2}$ of the sources is such that $H(S_1 \oplus_q S_2)$ is dominated by the message-sum capacity of the CQ-MAC, then the corresponding sum of sources can be reconstructed over the CQ-MAC. Therefore, if $R > 0$ is an achievable message-sum rate over a CQ-MAC, then $H(S_1 \oplus_q S_2) < R$ is a sufficient condition. We now state the main contribution of this section - a lower bound on the message-sum capacity of a CQ-MAC. Following its proof, we leverage the above argument in Thm. VI.9 to characterize sufficient conditions for reconstructing sum of sources over an arbitrary CQ-MAC.

Definition VI.6. Given a CQ-MAC $\rho_{\underline{\mathcal{X}}} \triangleq (\rho_{x_1x_2} \in \mathcal{D}(\mathcal{H}_Y) : (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2)$ and a prime power q, let $\mathscr{P}(\rho_{\underline{\mathcal{X}}}, q)$

$$\triangleq \left\{ (p_{V_1V_2U}, \rho_u : u \in \mathcal{V}) \ : \begin{array}{l} p_{V_1X_1}p_{V_2X_2} \text{ is a PMF on } \mathcal{V} \times \mathcal{X}_1 \times \mathcal{V} \times \mathcal{X}_2, \mathcal{V} = \mathcal{F}_q, \\ p_{V_1V_2U}(v_1, v_2, u) = \sum\limits_{x_1,x_2 \in \underline{\mathcal{X}}} p_{V_1X_1}(v_1, x_1)p_{V_2X_2}(v_2, x_2)\mathbb{1}_{\{u=v_1 \oplus_q v_2\}} \\ \rho_u \triangleq \sum\limits_{v_1 \in \mathcal{F}_q} \sum\limits_{v_2 \in \mathcal{F}_q} p_{V_1V_2|U}(v_1, v_2|u)\rho_{v_1v_2}\mathbb{1}_{\{v_1 \oplus_q v_2 = u\}}, \\ \rho_{v_1v_2} \triangleq \sum\limits_{x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2} p_{X_1|V_1}(x_1|v_1)p_{X_2|V_2}(x_2|v_2)\rho_{x_1x_2} \end{array} \right\}$$

For $(p_{V_1V_2U}, \rho_u : u \in \mathcal{F}_q) \in \mathscr{P}(\rho_{\underline{\mathcal{X}}}, q)$, let

$$\begin{aligned} \mathscr{R}(p_{V_1V_2U}, \rho_{\mathcal{V}}) &\triangleq \min\{H(V_1), H(V_2)\} - H(U) + \chi(\{p_U; \rho_u\}) \quad \text{and} \\ \mathscr{R}(\rho_{\underline{\mathcal{X}}}, q) &\triangleq \sup_{p_{V_1V_2U}, \rho_{\mathcal{V}} \in \mathscr{P}(\rho_{\underline{\mathcal{X}}}, q)} \mathscr{R}(p_{V_1V_2U}, \rho_{\mathcal{V}}) \end{aligned} \tag{6.5}$$

Lemma VI.7. *$\mathcal{F}_q-$message-sum rate $\mathscr{R}(\rho_{\underline{\mathcal{X}}}, q)$ is achievable over a CQ-MAC $\rho_{\underline{\mathcal{X}}} = (\rho_{x_1x_2} \in \mathcal{D}(\mathcal{H}_Y) : (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2)$.*

Proof. Let $(p_{V_1V_2U}, \rho_u : u \in \mathcal{V}) \in \mathscr{P}(\rho_{\underline{\mathcal{X}}}, q)$ with associated collection $(\rho_{v_1v_2} : (v_1, v_2) \in \mathcal{V}_1 \times \mathcal{V}_2)$ of density operators and PMF $p_{V_1X_1}p_{V_2X_2}$ on $\mathcal{V}_1 \times \mathcal{X}_1 \times \mathcal{V}_2 \times \mathcal{X}_2$ where $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{F}_q$.

We now describe the coding scheme in terms of a specific code. It is instructive to revisit Sec. 6.3, wherein we specified the import of both encoders employing cosets of the the *same* linear code. In order to choose codewords of a desired empirical distribution p_{V_j}, we employ NCCs (as

was done for the same reason in Sec. 6.4). Following the same notation as in proof of Thm. VI.3, we now specify the random coding scheme.

Let $G_I \in \mathcal{F}_q^{k\times n}, G_{O/I} \in \mathcal{F}_q^{l\times n}, B_j \in \mathcal{F}_q^n : j \in [2]$ be mutually independent and uniformly distributed on their respective range spaces. Through out this proof, we let $\oplus = \oplus_q$. Let $V_j^n(a, m_j) \triangleq aG_I \oplus m_j G_{O/I} \oplus B_j^n : (a, m_j) \in \mathcal{F}_q^{k+l}$ for $j \in [2]$ and $U^n(a, m) \triangleq aG_I \oplus mG_{O/I} \oplus B_1^n \oplus B_2^n : (a, m) \in \mathcal{F}_q^{k+l}$. For $j \in [2]$, let

$$S_j(m_j) \triangleq \begin{cases} \{a \in \mathcal{V}^k : V_j^n(a, m_j) \in T_\delta^n(p_{V_j})\} & \text{if } \sum_{a\in\mathcal{V}^k} \mathbb{1}_{\left\{V_j^n(a,m_j)\in T_\delta^n(p_{V_j})\right\}} \geq 1 \\ \{0^k\} & \text{otherwise, i.e } \sum_{a\in\mathcal{V}^k} \mathbb{1}_{\left\{V_j^n(a,m_j)\in T_\delta^n(p_{V_j})\right\}} = 0 \end{cases}$$

for each $m_j \in \mathcal{V}^l$. For $m_j \in \mathcal{V}^l$, a predetermined element $A_{j,m_j} \in S_j(m_j)$ is chosen. We let $\Theta_j(m_j) \triangleq |S_j(m_j)|$. For $m_j \in \mathcal{V}^l$, a predetermined $X_j^n(m_j) \in \mathcal{X}_j^n$ is chosen. As we shall see later, the choice of $X_j^n(m_j)$ is based on $V_j^n(A_{j,m_j}, m_j)$. We are thus led to the encoding rule.

Encoding Rule: On receiving message $(m_1, m_2) \in \mathcal{V}^l \times \mathcal{V}^l$, the quantum state

$$\rho_{m_1 m_2} \triangleq \rho_{X_1^n(m_1)X_2^n(m_2)} \triangleq \otimes_{t=1}^n \rho_{X_{1t}(m_1)X_{2t}(m_2)}$$

is (distributively) prepared.

Distribution of the Random Code: The distribution of the random code is completely specified through the distribution $\mathcal{P}(\cdot)$ of $G_I, G_{O/I}, B_1^n, B_2^n, (A_{1,m_1} : m_1 \in \mathcal{V}^l), (A_{2,m_2} : m_2 \in \mathcal{V}^l)$ and $(X_j^n(m_j) : m_j \in \mathcal{V}^l)$. We let

$$\begin{aligned} &\mathcal{P}\left(\begin{array}{c}(A_{1,m_1} = a_{1,m_1} : m_1 \in \mathcal{V}^l), (A_{2,m_2} = a_{2,m_2} : m_2 \in \mathcal{V}^l), \\ B_j^n = b_j^n : j \in [2], (X_1(m_1) = x_1^n(m_1) : m_1 \in \mathcal{V}^l), \\ G_I = g_I, G_{O/I} = g_{O/I}, (X_2(m_2) = x_2^n(m_2) : m_2 \in \mathcal{V}^l)\end{array}\right) \\ &\quad = \left[\prod_{m_1} \frac{\mathbb{1}_{\{a_{1,m_1}\in s_1(m_1)\}}}{\Theta(m_1)} p_{X_1|V_1}^n(x_1^n(m_1)|v_1^n(a_{1,m_1}, m_1))\right] \\ &\qquad \times \left[\prod_{m_2} \frac{\mathbb{1}_{\{a_{2,m_2}\in s_2(m_2)\}}}{\Theta(m_2)} p_{X_2|V_2}^n(x_2^n(m_2)|v_2^n(a_{2,m_2}, m_2)\right] \times \frac{1}{q^{kn+ln+2n}} \end{aligned} \tag{6.6}$$

Towards specifying a decoding POVM, we state the associated density operators modeling the quantum systems, their spectral decompositions and projectors. Let

$$\rho = \sum_{y\in\mathcal{Y}} s_Y(y)\left|h_y\right\rangle\left\langle h_y\right|,\ \rho_{x_1x_2} = \sum_{y\in\mathcal{Y}} p_{Y|X_1X_2}(y|x_1,x_2)\left|e_{y|x_1x_2}\right\rangle\left\langle e_{y|x_1x_2}\right| : (x_1,x_2)\in\underline{\mathcal{X}}$$

$$\rho_{v_1v_2} = \sum_{y\in\mathcal{Y}} q_{Y|V_1V_2}(y|v_1,v_2)\left|f_{y|v_1v_2}\right\rangle\left\langle f_{y|v_1v_2}\right| : (v_1,v_2)\in\underline{\mathcal{V}},\ \rho_u = \sum_{y\in\mathcal{Y}} r_{Y|U}(y|u)\left|g_{y|u}\right\rangle\left\langle g_{y|u}\right| : u\in\mathcal{U},$$

Decoding POVM: Unlike a generic CQ-MAC decoder *Winter* (2001), which aims at decoding both the classical messages from the quantum state received, the decoder here is designed to decode only the sum of messages transmitted. For this, the decoder employs the nested coset code $(n,k,l,G_I,G_{O/I},B^n)$, where $B^n = B_1^n \oplus B_2^n$. We define $U^n(a,m) \triangleq aG_I + mG_{O/I} + B^n$ to represent a generic codeword. We let $\Pi_{a,m} \triangleq \pi_{U^n(a,m)}\mathbb{1}_{\{U^n(a,m)\in\mathcal{T}_\delta^{(n)}(p_U)\}}$, where p_U is as defined in the theorem statement. The decoder is provided with a sub-POVM $\Lambda_{\mathcal{I}} \triangleq \{\Lambda_m \triangleq \sum_{a\in\mathcal{F}_q^k}\Lambda_{a,m} : m\in\mathcal{F}_q^l\}$ where

$$\Lambda_{a,m} \triangleq \Big(\sum_{\hat{a}\in\mathcal{F}_q^k}\sum_{\hat{m}\in\mathcal{F}_q^l}\Gamma_{\hat{a},\hat{m}}\Big)^{-1/2}\Gamma_{a,m}\Big(\sum_{\hat{a}\in\mathcal{F}_q^k}\sum_{\hat{m}\in\mathcal{F}_q^l}\Gamma_{\hat{a},\hat{m}}\Big)^{-1/2},$$

$\Lambda_{-1} \triangleq I - \sum_{a\in\mathcal{F}_q^k}\sum_{m\in\mathcal{F}_q^l}\Lambda_{a,m}$ and $\Gamma_{a,m} \triangleq \pi_\rho\Pi_{(a,m)}\pi_\rho$. We note that

$$\pi_\rho \triangleq \sum_{y^n\in T_\delta^n(s_Y)}\bigotimes_{t=1}^n\left|h_{y_t}\right\rangle\left\langle h_{y_t}\right| \text{ and } \pi_{u^n} \triangleq \sum_{y^n:(u^n,y^n)\in T_\delta^n(p_Ur_{Y|U})}\bigotimes_{t=1}^n\left|g_{y_t|u_t}\right\rangle\left\langle g_{y_t|u_t}\right|$$

denote the typical and conditional typical projectors (as stated in Definition 15.2.4 *Wilde* (2013a)) with respect to $\rho \triangleq \sum_{u\in\mathcal{F}_q} p_U(u)\rho_u$ and $(\rho_u : u\in\mathcal{U})$, respectively.

Error Analysis: We derive upper bounds on $\mathbb{E}_{\mathcal{P}}\{\bar{\xi}(c_{m\oplus})\}$. Our derivation will be similar to those adopted in proof of Thm. VI.3. Let us define event

$$\mathscr{E} \triangleq \left\{\begin{pmatrix} V_1^n(A_{1.m_1},m_1), X_1^n(m_1), \\ V_2^n(A_{2.m_2},m_2), X_2^n(m_2), \\ V_1^n(A_{1.m_1},m_1)\oplus V_2^n(A_{2.m_2},m_2)\end{pmatrix} \in T_{8\delta}(p_{V_1X_1V_2X_2U})\right\}. \tag{6.7}$$

We have

$$\begin{aligned}
&\mathbb{E}_{\mathcal{P}}\left\{\sum_{m_1}\sum_{m_2} p_{M_1M_2}(m_1,m_2)\,\mathrm{Tr}([I-\Lambda_{m_1\oplus m_2}])\rho_{m_1m_2}^{\otimes n}\right\}\\
&\quad\leq \underbrace{\mathbb{E}_{\mathcal{P}}\left\{\sum_{m_1}\sum_{m_2} p_{M_1M_2}(m_1,m_2)\,\mathrm{Tr}([I-\Lambda_{m_1\oplus m_2}])\rho_{m_1m_2}^{\otimes n}\mathbb{1}_{\mathscr{E}^c}\right\}}_{T_1}\\
&\quad+\underbrace{\mathbb{E}_{\mathcal{P}}\left\{\sum_{m_1}\sum_{m_2} p_{M_1M_2}(m_1,m_2)\,\mathrm{Tr}([I-\Lambda_{m_1\oplus m_2}])\rho_{m_1m_2}^{\otimes n}\mathbb{1}_{\mathscr{E}}\right\}}_{T_2}.
\end{aligned}$$

In regards to T_1, the sub-POVM nature of $\Lambda_{\mathcal{I}}$ and the fact that $\rho_{m_1,m_2}^{\otimes n}$ is a density operator enables us conclude $T_1 \leq \mathbb{E}_{\mathcal{P}}\{\mathbb{1}_{\mathscr{E}^c}\}$. Furthermore, observe that $X_j(m_j)$ is distributed with PMF $p_{X_j|V_j}^n$ conditionally on $V_j^n(A_{j,m_j,m_j})$. (See (6.6). In addition, $p_{V_1X_1V_2X_2} = p_{V_1X_1}p_{V_2X_2}$ implies that, standard conditional typicality arguments yields

$$\mathbb{E}_{\mathcal{P}}\{\mathbb{1}_{\mathscr{E}^c}\} \leq \mathbb{E}_{\mathcal{P}}\left\{\sum_{m_1} p_{M_1}(m_1)\mathbb{1}_{\{\Theta_1(m_1)=0\}} + \sum_{m_2} p_{M_2}(m_2)\mathbb{1}_{\{\Theta_1(m_2)=0\}}\right\} + \exp\{-n\delta\}, \tag{6.8}$$

where δ is chosen appropriately. In the above inequality, the second term on the RHS is an upper bound on the probability of the event $(X_1^n(m_1, X_2^n(m_2)) \notin T_\delta^n(p_{V_1X_1V_2X_2U}|v_1^n, v_2^n, v_1^n \oplus v_2^n)$ conditioned on $(V_1^n(A_{1.m_1}, m_1), V_2^n(A_{2.m_2}, m_2), V_1^n(A_{1.m_1}, m_1) \oplus V_2^n(A_{2.m_2}, m_2)) = (v_1^n, v_2^n, v_1^n \oplus v_2^n) \in T_\delta^n(p_{V_1V_2U})$ and the first term provides an upper bound on the complement of the latter event. An upper bound on T_1 therefore reduces to deriving an upper bound on the first term on the RHS of (6.8). This task - deriving an upper bound on the first term on the RHS of (6.8) - being a classical analysis, has been detailed in several earlier works *Padakandla* (2014); *Padakandla et al.* (2016); *Padakandla and Pradhan* (2017, 2018) and in particular (*Pradhan et al.*, 2021, Proof of Thm. 2.5). Following this, we have

$$\mathbb{E}_{\mathcal{P}}\left\{\sum_{m_j} p_{M_j}(m_j)\mathbb{1}_{\{\Theta_j(m_j)=0\}}\right\} \;\leq\; \exp\left\{-n\left(\frac{k\log q}{n} - [\log q - H(V_j)]\right)\right\}$$

thereby ensuring $T_1 \leq 2\exp\{-n\delta\}$ if

$$\frac{k\log q}{n} \geq \max\{\log q - H(V_1), \log q - H(V_2)\} = \log q - \min\{H(V_1), H(V_2)\}. \tag{6.9}$$

We now analyze T_2. Applying the Hayashi-Nagaoka inequality, we have

$$T_2 \leq \mathbb{E}_{\mathcal{P}}\big[T_{21} + T_{22} + T_{23}\big], \tag{6.10}$$

where

$$\begin{aligned}
T_{21} &\triangleq 2\sum_{m_1}\sum_{m_2} p_{M_1M_2}(m_1,m_2)\,\mathrm{Tr}\Big([I - \Gamma_{A^{\oplus}_{\underline{m}},\underline{m}^{\oplus}}]\rho^{\otimes n}_{m_1m_2}]\Big)\mathbb{1}_{\mathscr{E}},\\
T_{22} &\triangleq 4\sum_{m_1}\sum_{m_2}\sum_{\hat{a}\neq A^{\oplus}_{\underline{m}}} p_{M_1M_2}(m_1,m_2)\,\mathrm{Tr}\big(\Gamma_{\hat{a},\underline{m}^{\oplus}}\rho^{\otimes n}_{m_1m_2}\big)\mathbb{1}_{\mathscr{E}},\\
T_{23} &\triangleq 4\sum_{m_1}\sum_{m_2}\sum_{\hat{a}\neq A^{\oplus}_{\underline{m}}}\sum_{\hat{m}\neq \underline{m}^{\oplus}} p_{M_1M_2}(m_1,m_2)\,\mathrm{Tr}\big(\Gamma_{\hat{a},\hat{m}}\rho^{\otimes n}_{m_1m_2}\big)\mathbb{1}_{\mathscr{E}},
\end{aligned}$$

and $A^{\oplus}_{\underline{m}} \triangleq A_{1,m_1} \oplus A_{2,m_2} \in \mathcal{V}^k, \underline{m}^{\oplus} \triangleq m_1 \oplus m_2 \in \mathcal{V}^l$. We note that (6.10) follows from an argument analogous to the one in (6.2). We now analyze T_{21}, T_{22} and T_{23}.

We begin with T_{21}. For each m_1 and m_2, denote the events

$$\begin{aligned}
&\mathcal{J} \triangleq \left\{\left(V_1^n(A_{1.m_1},m_1), X_1^n(m_1), V_2^n(A_{2.m_2},m_2), X_2^n(m_2)\right) = (v_1^n,x_1^n,v_2,x_2) \in T_\delta(p_{V_1X_1V_2X_2})\right\},\\
&\mathcal{V} \triangleq \{V_j^n(a_j,m_j) = v_j^n \colon j\in[2]\}, \quad \hat{\mathcal{V}} \triangleq \{V^n(\underline{a}^{\oplus}, m_1\oplus m_2) = \underline{v}^n_{\oplus}\}, \quad \mathcal{A} \triangleq \{A_{j,m_j} = a_j : j\in[2]\},
\end{aligned}$$

abbreviating $\underline{v}^n_{\oplus} = v_1^n \oplus v_2^n$, $\underline{a}^{\oplus} = a_1 \oplus a_2$. We have

$$\mathbb{E}_{\mathcal{P}}[T_{21}] = 1 - \sum_{\underline{m}}\sum_{a_1,a_2}\sum_{\substack{(\underline{v}^n,\underline{x})\in\\ T_\delta(p_{\underline{VX}})}} p_{\underline{M}}(\underline{m})\,\mathrm{Tr}\Big(\pi_{v^n}\pi_\rho\rho^{\otimes n}_{x_1^nx_2^n}\pi_\rho\Big)\mathbb{E}_{\mathcal{P}}\left[\mathbb{1}_{\mathcal{J}}\mathbb{1}_{\mathcal{A}}\mathbb{1}_{\mathcal{V}}\mathbb{1}_{\hat{\mathcal{V}}}\right] \leq \epsilon$$

for all sufficiently large n and sufficiently small δ, where the last inequality follows from the pinching argument, also provided in Lemma E.1 (see Appendix E.1). Set $\mathcal{A} = \mathcal{V} = \mathcal{F}_q$, $\mathcal{B} = \mathcal{X}$, $p_{AB} = p_{V_1\oplus V_2,X}$ and the density operators correspondingly.

The proposition below bounds the terms T_{22} and T_{23} as follows.

Proposition VI.8. *For any $\epsilon \in (0,1)$, and for all sufficiently small $\delta > 0$ and sufficiently large n, we have $\mathbb{E}_{\mathcal{P}}[T_{22} + T_{23}] \leq \epsilon$ if the following inequalities hold:*

$$\begin{aligned} \frac{3k}{n} \log q &\leq 3 \log q + I(V;Z)_\sigma - H(V_1, V_2) - H(V) - \epsilon \\ \frac{(3k+l)}{n} \log q &\leq 3 \log q + I(V;Z)_\sigma - H(V_1, V_2) - H(V) - \epsilon. \end{aligned}$$

Proof. The proof is provided in Appendix E.2. □

This completes the proof of the claimed statement. □

We conclude this section with our final result of this chapter in regards to decoding sum of sources. The proof of the following theorem follows from the discussion provided just prior to Definition VI.6. We therefore omit a detailed proof but just state the encoding and decoding techniques for completeness.

Theorem VI.9. *The sum of a pair of sources distributed with PMF $\mathbb{W}_{S_1 S_2}$ can be reconstructed on a CQ-MAC $\rho_{\underline{\mathcal{X}}} = (\rho_{x_1 x_2} \in \mathcal{D}(\mathcal{H}_{\underline{Y}}) : (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2)$ if $H(S_1 \oplus_q S_2) < \mathscr{R}(\rho_{\underline{\mathcal{X}}}, q)$.*

Proof. We begin with an outline of our coding scheme. As stated in Sec. 6.3, we propose a 'separation based approach' with two modules - source and channel. The source coding module employs a (distributed) Körner Marton (KM) source code. Specifically, *Korner and Marton* (1979) guarantees the existence of a parity check matrix $h \in \mathcal{F}_q^{l \times n} = \mathcal{S}^{l \times n}$ and a decoder map $d : \mathcal{F}_q^l \to \mathcal{S}^n$ such that $\sum_{\underline{s}^n \in \underline{\mathcal{S}}^n} \mathbb{W}^n_{\underline{S}}(\underline{s}^n) \mathbb{1}_{\{d(hs_1^n \oplus_q hs_2^n) \neq s_1^n \oplus_q s_2^n\}} \leq \epsilon$, for any $\epsilon > 0$, and sufficiently large n, so long as $\frac{l \log_2 q}{n} > H(S_1 \oplus_q S_2)$.

Both encoders of this KM source coding module employ one such parity check matrix $h \in \mathcal{F}_q^{l \times n}$. The decoder of the KM source code employs the corresponding decoder map d. KM Source encoder j outputs $M_j^l = h(S_j^n)$. If the KM source decoder is provided $M_1^l \oplus_q M_2^l$, then it can reconstruct $S_1^n \oplus_q S_2^n$ with reliability at least $1 - \epsilon$. The task of the CQ-MAC channel coding module is to make $M_1^l \oplus_q M_2^l$ available to the KM source decoder. We are thus confronted with the problem of designing a CQ-MAC channel coding module that can reliably communicate the sum of the messages indices that are input at the encoders.

Specifically, this channel coding module must communicate $M_1^l \oplus_q M_2^l \in \mathcal{F}_q^l$ within n channel uses. If we can prove that there exists a MAC channel coding module for sufficiently large n so

long as

$$\frac{l \log_2 q}{n} < \min\{H(V_1), H(V_2)\} - H(U) + \chi(\{p_U; \rho_u\})$$

for any choice of auxiliary

The source module employs the KM code. The corresponding KM decoder at the receiver only requires the sum of the message indices output by the KM code. The CQ-MAC channel coding module needs to communicate only the *sum* of the two message indices input by the two encoders. Given ϵ_c, we seek to identify a CQ-MAC code $c = (n, e_1, e_2, \mathcal{M})$ such that $\overline{\xi}(c) \leq \epsilon_c$.

Encoding: The process of mapping source sequences to the CQ-MAC channel inputs is divided into two stages. In the first stage, a distributed source code proposed in *Korner and Marton* (1979) is employed which maps the $n-$length source sequences to message indices taking values over $\mathcal{F}_q^l$. For the second stage we develop functions mapping these message indices to channel input codewords. We begin by defining the first stage of encoding which relies on Lemma 1 of *Padakandla and Pradhan*. This lemma guarantees the existence of a parity check matrix $h \in \mathcal{F}_q^{l(n) \times n}$ and a map $d : \mathcal{F}_q^l(n) \to \mathcal{F}_q^n$, for a sufficiently large n, such that (i) $\frac{l(n)}{n} \leq H(S_1 \oplus_q S_2) + \frac{\epsilon_c}{2}$ and (ii) $\mathbb{P}(d(hS_1^n \oplus hS_2^n) \neq S_1^n \oplus S_2^n) \leq \frac{\epsilon_c}{2}$. We use one such parity matrix which satisfies the above conditions and define $M_j \triangleq hS_j$, for $j = 1, 2$. This forms our first stage encoder.

Moving on to the second stage encoder, let us denote the maps of the two encoders as $\kappa_j : \mathcal{F}_q^l \to \mathcal{X}_j^n : j = 1, 2..$ For this stage, we use the NCC encoding developed in Section 6.4 for a CQ-PTP. Consider two NCCs with parameters $(n, k, l, g_I, g_{O/I}, b_j^n) : j \in \{1, 2\}$ with range spaces as $v_j^n(a, m_j) \triangleq ag_I \oplus m_j g_{O/I} \oplus b_j^n : j \in \{1, 2\}$, respectively. Note that the two NCCs share the common g_I and $g_{O/I}$, but not necessarily the bias vector b_j^n. Encoder j then constructs its NCC CQ-PTP code $(n, \mathcal{I} = \mathcal{F}_q^l, e_j, \lambda_{\mathcal{I}}^j)$ using the corresponding NCC $(n, k, l, g_I, g_{O/I}, b_j^n)$ as described in Definition 6.1. This defines the second stage encoding. Integrating the two stages, we obtain the following. To transmit the source sequence pair (s_1^n, s_2^n) the sequence pair $(e_1(hs_1^n), e_2(hs_2^n))$ is send over the CQ-MAC channel which produces the quantum state $\rho_{e_1(hs_1^n), e_2(hs_2^n)}$ as the output.

After performing the measurement and decoding the message $\hat{m}$, the decoder then employs the KM decoder $d(.)$ to obtain the sum of sources $d(\hat{m})$. An analysis of this coding scheme is provided in the Proof of Lemma VI.7.

□

6.6 Decoding arbitrary functions over CQ-MAC

Leveraging the technique developed in Theorem VI.9, in this section we provide the following proposition to reconstruct an arbitrary function of the sources

Proposition VI.10. *The function $f : \underline{\mathcal{S}} \to \mathcal{S}$ of sources $\mathbb{W}_{S_1S_2}$ can be reconstructed on a CQ-MAC $(\rho_{x_1x_2} \in \mathcal{D}(\mathcal{H}_{\underline{Y}}) : (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2)$ if there exists functions $h_j : \mathcal{S}_j \to \mathcal{F}_q$ for $j : 1, 2$, a function $g : \mathcal{F}_q \to \mathcal{S}$, and a PMF $p_{V_1V_2X_1X_2} = p_{V_1X_1}p_{V_2X_2}$ on $\mathcal{V}_1 \times \mathcal{X}_1 \times \mathcal{V}_2 \times \mathcal{X}_2$, where $\mathcal{V}_1 = \mathcal{V}_1 = \mathcal{F}_q$,such that $f(s_1, s_2) = g(h_1(s_1) \oplus h_2(s_2))$ and $H(h_1(S_1) \oplus_q h_2(S_2)) \leq \min\{H(V_1), H(V_2)\} - H(U) + \chi(\{p_U; \rho_u\})$, where p_U and ρ_u are as defined in Theorem VI.9.*

Proof. The proof follows from the proof of Theorem VI.9. □

Example VI.11. Let $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{S}_1 = \mathcal{S}_2 = \mathcal{X} = \{0, 1\}$, $\mathcal{H} = \mathbb{C}^2$, and $\rho_{x_1,x_2} = (1 - q)\sigma_{(x_1 \vee x_2)} + q\sigma_{(\bar{x}_1 \wedge \bar{x}_2)}$, where $\sigma_0, \sigma_1 \in \mathcal{D}(\mathcal{H})$ be arbitrary. Let $\rho(q) \triangleq (1 - q)\sigma_0 + q\sigma_1$. Consider correlated symmetric individually uniform sources with $\mathbb{W}_{S_1|S_2}(1|0) = \mathbb{W}_{S_1|S_2}(0|1) = p$ for $p \in (0, 1)$. Let $f(S_1, S_2) = S_1 \vee S_2$. Consider the sufficient conditions given by the unstructured coding scheme: $H(S_1, S_2) < \chi(\{P_{X_1,X_2}, \rho_{x_1,x_2}\})$, with X_1 and X_2 being independent, which can be simplified as $1 + h_b(p) < S(\rho(0.5)) - S(\rho(q))$. This implies that the f is not reconstructible using the unstructured codes. We embed f in the ternary field. In other words, the encoders and decoder work toward reconstructing $S_1 \oplus_3 S_2$. The sufficient condition given by the algebraic coding scheme turns out to be

$$H(S_1 \oplus_3 S_2) < H(X_1) - H(X_1 \oplus_3 X_2) + \chi(\{p_{X_1 \oplus_3 X_2}, \rho_{x_1 \oplus_3 x_2}\}),$$

for some p_{X_1,X_2}, which can be simplified as

$$\begin{aligned} &h_b(2p - p^2) + (2p - p^2)h_b(p/(2 - p)) < \\ &\max_{\theta}[h_b(\theta) - h_b(2\theta - \theta^2) - (2\theta - \theta^2)h_b(\theta/(2 - \theta)) \\ &\qquad + S(\rho((2\theta - \theta^2) * q)) - S(\rho(q))]. \end{aligned}$$

One can show that there exists choices for p, q, σ_0 and σ_1 such that this condition is satisfied.

CHAPTER VII

Structured Codes for 3−User Classical-Quantum Interference Channel

7.1 Introduction

We consider the scenario of communicating over a 3−user classical-quantum interference channel (3−CQIC) (Fig. 7.1). We undertake a Shannon-theoretic study for characterizing an inner bound to its capacity region. The current known coding schemes for CQICs *Sen* (2012); *Savov* (2012); *Sen* (2018a); *Hirche et al.* (2016) are based on unstructured codes. In this chapter, we propose a new coding scheme for a 3−CQIC based on *nested coset codes* (NCCs) - codes possessing algebraic structure. Analyzing its performance, we derive a new inner bound (Sec. 7.4) to the capacity region of 3to1−CQIC - a sub-class of 3−CQICs. The inner bound is proven to subsume any current known inner bounds based on unstructured codes. Furthermore, we identify examples of 3to1−CQICs for which the derived inner bound is strictly larger. These findings are a first step towards characterizing a new inner bound to the capacity region of a general 3−CQIC.

The current approach of characterizing the performance limits of CQ channels is based on unstructured codes, which remained for several decades the de facto ensemble of codes for information-theoretic study of any classical channels. Inspired by the work in *Korner and Marton* (1979) and followed by findings in a multitude of network communication scenarios *Krithivasan and Pradhan* (2011); *Nazer and Gastpar* (2007); *Philosof and Zamir* (2009); *Jafarian and Vishwanath* (2012); *Padakandla et al.* (2016); *Padakandla and Pradhan* (2013), it has been analytically proven that coding schemes designed using codes endowed with algebraic closure properties can strictly outperform all known unstructured coding schemes.

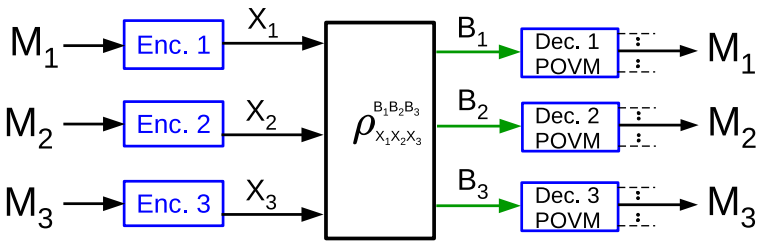

Figure 7.1: Communication over 3−CQIC.

The goal of this chapter is to build on this and enhance the current known coding schemes in the context of CQ channels. Our experience with classical channels suggests that a first step toward this is to design and analyze coding schemes for basic building block channels. Indeed, the ensemble of NCCs studied in the simple context of point-to-point (PTP) channels form an important element of this work *Pradhan et al.* (2021). On the other hand, the mathematical complexity of analyzing CQ channels makes it challenging to generalize even well known coding schemes to the CQ setting. In the light of this, our work maybe viewed as a first step in designing new coding schemes for network CQ channels based on coset codes.

In the context of CQICs, the focus of current research is on 2−user. There has been considerable effort in *Fawzi et al.* (2012); *Sen* (2012); *Savov* (2012); *Hirche et al.* (2016); *Sen* (2018b,a) at proving the achievability of the Han-Kobayashi rate-region (CHK) *Han and Kobayashi* (1981) for 2−user ICs. Analogous to these, one can leverage all known coding techniques - message splitting, superposition coding, Marton's binning - and derive an achievable rate region for a 3−CQIC. See discussion in (*Padakandla et al.*, 2016, Sec. III). This rate region, henceforth referred to as the $\mathscr{USB}$−region contains the largest current known inner bound for any 3−CQIC. In this work, we focus on 3to1−CQICs (Defn. VII.3) - a subclass of 3−IC in which only one receiver (Rx) experiences interference. We propose a coding scheme based on NCCs and derive an inner bound for this subclass that subsumes the $\mathscr{USB}$−region in general, and strictly larger for identified examples (see Ex. VII.6). Our analysis of this simultaneous decoder builds on the technique proposed in *Fawzi et al.* (2012). In the next step, we leverage these building blocks and employ a multi-terminal simultaneous decoder *Sen* (2018b) to derive a new achievable rate region for 3to1−CQICs.

7.2 Preliminaries and Problem Statement

We supplement our notation thus far with the following. For $n \in \mathbb{N}$, $[n] \triangleq \{1, \cdots, n\}$. We let an underline denote an appropriate aggregation of objects. For example, $\underline{\mathcal{X}} \triangleq \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3$, $\underline{x} \triangleq (x_1, x_2, x_3) \in \underline{\mathcal{X}}$ and in regards to Hilbert spaces $\mathcal{H}_{Y_i} : i \in [3]$, we let $\mathcal{H}_{\underline{Y}} \triangleq \otimes_{i=1}^{3} \mathcal{H}_{Y_i}$. We abbreviate the Positive Operator Valued Measure and Block-Length as POVM and B-L, respectively.

Consider a (generic) $3-CQIC$ $(\rho_{\underline{x}} \in \mathcal{D}(\mathcal{H}_{\underline{Y}}) : \underline{x} \in \underline{\mathcal{X}}, \kappa_j : j \in [3])$ specified through (i) three finite sets $\mathcal{X}_j : j \in [3]$, (ii) three Hilbert spaces $\mathcal{H}_{Y_j} : j \in [3]$, (iii) a collection of density operators $(\rho_{\underline{x}} \in \mathcal{D}(\mathcal{H}_{\underline{Y}}) : \underline{x} \in \underline{\mathcal{X}})$ and (iv) three cost functions $\kappa_j : \mathcal{X}_j \to [0, \infty) : j \in [3]$. The cost function is assumed to be additive, i.e., cost expended by encoder j in preparing the state $\otimes_{t=1}^{n} \rho_{x_{1t} x_{2t} x_{3t}}$ is $\overline{\kappa}_j^n \triangleq \frac{1}{n} \sum_{t=1}^{n} \kappa_j(x_{jt})$. Reliable communication on a $3-$CQIC entails identifying a code.

Definition VII.1. A $3-CQIC$ *code* $c = (n, \underline{\mathcal{M}}, \underline{e}, \underline{\lambda})$ of B-L n consists of three (i) message index sets $[\mathcal{M}_j] : j \in [3]$, (ii) encoder maps $e_j : [\mathcal{M}_j] \to \mathcal{X}_j^n : j \in [3]$ and (iii) POVMs $\lambda_j \triangleq \{\lambda_{j,m} : \mathcal{H}_j^{\otimes n} \to \mathcal{H}_j^{\otimes n} : m \in [\mathcal{M}_j]\} : j \in [3]$. The average probability of error of the $3-$CQIC code $(n, \underline{\mathcal{M}}, \underline{e}, \lambda^{[3]})$ is

$$\overline{\xi}(\underline{e}, \underline{\lambda}) \triangleq 1 - \frac{1}{\mathcal{M}_1 \mathcal{M}_2 \mathcal{M}_3} \sum_{\underline{m} \in \underline{\mathcal{M}}} \mathrm{Tr}\big(\lambda_{\underline{m}} \rho_{c,\underline{m}}^{\otimes n}\big).$$

where $\lambda_{\underline{m}} \triangleq \otimes_{j=1}^{3} \lambda_{j,m_j}$, $\rho_{c,\underline{m}}^{\otimes n} \triangleq \otimes_{t=1}^{n} \rho_{x_{1t} x_{2t} x_{3t}}$ where $(x_{jt} : 1 \leq t \leq n) = x_j^n(m_j) \triangleq e_j(m_j)$ for $j \in [3]$. Average cost per symbol of transmitting message $\underline{m} \in \underline{\mathcal{M}} \in \underline{\tau}(\underline{e}|\underline{m}) \triangleq \Big(\overline{\kappa}_j^n(e_j(m_j)) : j \in [3]\Big)$ and the average cost per symbol of $3-$CQIC code is $\underline{\tau}(\underline{e}) \triangleq \frac{1}{|\underline{\mathcal{M}}|} \sum_{\underline{m} \in \underline{\mathcal{M}}} \underline{\tau}(\underline{e}|\underline{m})$.

Definition VII.2. A rate-cost vector $(R_1, R_2, R_3, \tau_1, \tau_2, \tau_3) \in [0, \infty)^6$ is *achievable* if there exists a sequence of $3-$CQIC code $(n, \underline{\mathcal{M}}^{(n)}, \underline{e}^{(n)}, \underline{\lambda}^{(n)})$ for which $\lim_{n \to \infty} \overline{\xi}(\underline{e}^{(n)}, \underline{\lambda}^{(n)}) = 0$,

$$\lim_{n \to \infty} n^{-1} \log \mathcal{M}_j^{(n)} = R_j, \text{ and } \lim_{n \to \infty} \underline{\tau}(\underline{e})_j \leq \tau_j : j \in [3].$$

The capacity region $\mathcal{C}(\rho_{\underline{x}} : \underline{x} \in \mathcal{X})$ of the $3-$CQIC $(\rho_{\underline{x}} \in \mathcal{D}(\mathcal{H}_{\underline{Y}}) : \underline{x} \in \underline{\mathcal{X}})$ is the set of all achievable rate-cost vectors. We define below the sub-class of 3to1$-$CQICs.

Definition VII.3. A $3-$CQIC $(\rho_{\underline{x}} \in \mathcal{D}(\mathcal{H}_{\underline{Y}}) : \underline{x} \in \underline{\mathcal{X}})$ is a 3to1$-$CQIC if (i) for every $\Lambda \in \mathcal{P}(\mathcal{H}_{Y_2})$, $\Gamma \in \mathcal{P}(\mathcal{H}_{Y_3})$, $\mathrm{Tr}((I \otimes \Lambda \otimes I)\rho_{x_1 x_2 x_3}) = \mathrm{Tr}((I \otimes \Lambda \otimes I)\rho_{\hat{x}_1 \hat{x}_2 \hat{x}_3})$ for every $\underline{x}, \underline{\hat{x}} \in \underline{\mathcal{X}}$ satisfying $x_2 = \hat{x}_2$, and (ii) $\mathrm{Tr}((I \otimes I \otimes \Gamma)\rho_{x_1 x_2 x_3}) = \mathrm{Tr}((I \otimes I \otimes \Gamma)\rho_{\hat{x}_1 \hat{x}_2 \hat{x}_3})$ for every $\underline{x}, \underline{\hat{x}} \in \underline{\mathcal{X}}$ satisfying $x_3 = \hat{x}_3$.

7.3 Illustration of the Central Idea

The goal here is to demonstrate the utility of algebraic closure in coding schemes for 3−ICs. While, we state Ex. VII.4 in the context of 3to1−CQICs, we discus in the context of a classical 3to1−IC. The latter provides an exposition on the utility of algebraic closure in network scenarios.

Example VII.4. Let $\mathcal{X}_j = \mathcal{X} = \{0,1\}, \mathcal{H}_j = \mathcal{H}$, $\sigma_b^{(j)} \in \mathcal{D}(\mathcal{H})$ for $j \in [3]$ and $b \in \mathcal{X}$. For $\underline{x} \in \underline{\mathcal{X}}$, let $\rho_{\underline{x}} \triangleq \sigma^{(1)}_{x_1 \oplus x_2 \oplus x_3} \otimes \sigma^{(2)}_{x_2} \otimes \sigma^{(3)}_{x_3}$. For $x \in \{0,1\}$, we let $\kappa_1(x) = x$ and $\kappa_k(x) = 0$ for $k = 2,3$.

Let $\mathcal{H} = \mathbb{C}^2$, $\sigma_b(\eta) \triangleq (1-\eta)\,|b\rangle\langle b| + \eta\,|1-b\rangle\langle 1-b|$ for $b \in \mathcal{X}$, $\eta \in [0,1]$. Let $\sigma_b^{(1)} \triangleq \sigma_b(\delta_1)$ and $\sigma_b^{(2)} \triangleq \sigma_b^{(3)} \triangleq \sigma_b(\delta)$ for $b \in \mathcal{X}$ and some specified $\delta, \delta_1 \in (0,1)$. In addition, let $\tau \in (0, \frac{1}{2})$ specify a Hamming cost constraint on Tx 1's input. With this choice, one identifies the above example with a 3to1−IC $Y_1 = X_1 \oplus X_2 \oplus X_3 \oplus N_1$, $Y_k = X_k \oplus N_k : k = 2,3$ with $N_1 \sim \text{Ber}(\delta_1)$, $N_k \sim \text{Ber}(\delta)$ $k = 2,3$ being independent. Tx $k \in \{2,3\}$ splits its information into U_k, X_k. Rx 1 decodes U_2, U_3, X_1, while Rx $k \in \{2,3\}$ decodes U_k, X_k. So long as $H(U_k|X_k) > 0$ for either $k \in \{2,3\}$, it can be shown that $H(X_2 \oplus X_3|U_2, U_3) > 0$ implying Tx-Rx 1 cannot achieve $h_b(\delta_1 * \tau) - h_b(\delta_1)$ - its interference free cost constrained capacity. If $h_b(\delta_1 * \tau) - h_b(\delta_1) + 2(1 - h_b(\delta)) > 1 - h_b(\delta_1)$, it can be shown that $H(U_k|X_k) > 0$ for either $k \in \{2,3\}$ precluding Tx-Rx 1 achieving a rate $h_b(\delta_1 * \tau) - h_b(\delta_1)$ using unstructured coding. Suppose users $2,3$ employ codes of rate $1 - h_b(\delta)$ that are cosets of the *same linear code*, then the above condition does not preclude Tx-Rx 1 from achieving a rate $h_b(\delta_1 * \tau) - h_b(\delta_1)$, so long as $\tau * \delta < \delta$, even if $1 + h_b(\tau * \delta_1) > 2h_b(\delta)$. The reason is, user 2 and 3's codebooks when added is another coset of the same rate $1 - h_b(\delta)$. Rx 1 can just decode this interference if $h_b(\delta_1 * \tau) - h_b(\delta_1) + 1 - h_b(\delta) < 1 - h_b(\delta_1)$ which is equivalent to $\tau * \delta < \delta < \frac{1}{2}$. Hence, for this 3to1−IC, if $h_b(\delta_1 * \tau) - h_b(\delta_1) + 2(1 - h_b(\delta)) > 1 - h_b(\delta_1)$ and $\tau * \delta < \delta < \frac{1}{2}$ hold, then coset codes are strictly more efficient than unstructured codes.

7.4 Rate region using Coset Codes for 3to1−CQIC

In this section we consider the above described 3to1-CQIC and provide an achievable rate-region.

Theorem VII.5. *Given a 3to1-CQIC* $(\rho_{\underline{x}} \in \mathcal{D}(\mathcal{H}_{\underline{Y}}) : \underline{x} \in \underline{\mathcal{X}}, \kappa_j : j \in [3])$ *and a PMF* $p_{V_2V_3X_1X_2X_3} = p_{X_1}p_{V_2X_2}p_{V_3X_3}$ *on* $\mathcal{V}_2 \times \mathcal{V}_3 \times \mathcal{X}_2 \times \mathcal{X}_3$ *where* $\mathcal{V}_2 = \mathcal{V}_3 = \mathcal{F}_q$*, a rate-cost triple* $(R_1, R_2, R_2, \tau_1, \tau_2, \tau_3)$

is achievable if it satisfies the following

$$\begin{aligned} R_1 &\leq I(Y_1; X_1|U)_{\sigma_1}, \quad R_j \leq I(Y_j; V_j)_{\sigma_2}, \\ R_j &\leq \min\{H(V_2), H(V_3)\} - H(U) + I(Y_1; U|X_1)_{\sigma_1}, \\ R_1 + R_j &\leq \min\{H(V_2), H(V_3)\} - H(U) + I(Y_1; V_1 U)_{\sigma_1}, \end{aligned}$$

for $j = 2, 3$, *and* $\mathbb{E}[\kappa_j(X_j)] \leq \tau_j : j \in [3]$, *where*

$$\begin{aligned} \sigma_1^{\underline{Y}} &\triangleq \sum_{x_1 \in \mathcal{X}_1, u \in \mathcal{F}_q} p_{X_1}(x_1) p_U(u) \rho_{x_1,u}^{\underline{Y}} \otimes |x_1\rangle\langle x_1| \otimes |u\rangle\langle u|, \\ \rho_{x_1,u}^{\underline{Y}} &\triangleq \sum_{v_2,v_3} \sum_{x_2,x_3} p_{V_2,V_3,X_2,X_3|U}(v_2, v_3, x_2, x_3|u) \rho_{\underline{x}}^{\underline{Y}} \\ \sigma_2 &= \sum_{v_1,v_2,v_3} p_{\underline{X}V_2V_3}(\underline{x}, v_2, v_3) \rho_{\underline{x}}^{\underline{Y}} \otimes |v_2\rangle\langle v_2| \otimes |v_3\rangle\langle v_3|, \end{aligned}$$

for $U \triangleq V_2 \oplus V_3$, *and* $\{|v_2\rangle\}$, $\{|v_3\rangle\}$ *as some orthonormal basis on* $\mathcal{H}_Y$.

Example VII.6. Let $\mathcal{X}_j = \mathcal{X} = \{0, 1\}, \mathcal{H}_j = \mathbb{C}^2$ and let

$$\sigma_0 \triangleq \begin{bmatrix} 2/3 & 0 \\ 0 & 1/3 \end{bmatrix}, \text{ and } \sigma_1 \triangleq \begin{bmatrix} 1/2 & 1/6 \\ 1/6 & 1/2 \end{bmatrix}.$$

$$\rho_{\underline{x}} \triangleq [(1-\delta_1)\sigma_{x_1 \oplus x_2 \oplus x_3} + \delta_1 \sigma_{x_1 \oplus x_2 \oplus x_3 \oplus 1}] \otimes [(1-\delta)\sigma_{x_2} + \delta\sigma_{x_2 \oplus 1}] \otimes [(1-\delta)\sigma_{x_3} + \delta\sigma_{x_3 \oplus 1}],$$

for $\underline{x} \in \underline{\mathcal{X}}$, where N_1, N_2 and N_3 are mutually independent Bernoulli random variables with biases δ_1, δ and δ, respectively. We let $\delta_1, \delta \in (0, 0.5)$. For $x \in \{0, 1\}$, we let $\kappa_1(x) = x$ and $\kappa_k(x) = 0$ for $k = 2, 3$. Let $\rho(p) := p\sigma_0 + (1-p)\sigma_1$. Note that $\rho(p)$ and $\rho(1-p)$ do not commute except for $p = 0.5$. It can be checked that $S(\rho(p))$ is a symmetric concave function of $p \in (0, 1)$. Consider the case when $\tau * \delta_1 \leq \delta$. Using NCC, the three users can achieve their PTP capacities simultaneously: $S(\rho(\tau * \delta_1)) - S(\rho(\delta_1))$, $S(\rho(0.5)) - S(\rho(\delta))$, and $S(\rho(0.5)) - S(\rho(\delta))$, respectively. These correspond to the rates given by $I(X_1; B_1|X_2 \oplus X_3)$, $I(X_2; Y_2)$, and $I(X_3; Y_3)$. One can show that if $S(\rho(\tau * \delta_1)) - S(\rho(\delta_1)) + 2(S(\rho(0.5)) - S(\rho(\delta))) > S(\rho(0.5)) - S(\rho(\delta_1))$, then using

unstructured codes, all three users cannot achieve their respective capacities simultaneously. This condition is equivalent to the condition: $S(\rho(\tau * \delta_1)) + S(\rho(0.5)) > 2S(\rho(\delta))$. Hence by choosing $\tau * \delta_1 = \delta$, and $\delta < 0.5$, we see that NCC-based coding scheme enables all users achieve their respective capacities simultaneously, while this is not possible in unstructured coding scheme.

Proof. We divide the proof into three parts entailing the encoding, decoding and error analysis techniques.

7.4.1 Encoding Technique

Consider a PMF $p_{V_2V_3\underline{X}}$ on $\mathcal{V}_2 \times \mathcal{V}_3 \times \underline{\mathcal{X}}$ with $\mathcal{V}_2 = \mathcal{V}_3 = \mathcal{F}_q$, and choose n and $R_j : j = [3]$ as non-negative integers. For encoder 1, we use the random coding strategy and construct a codebook $\mathcal{C}_1 \triangleq \{x_1(m_1) : m_1 \in [2^{nR_1}]\}$ on $\mathcal{X}_1$ using the marginal PMF $p^n_{X_1}$. Let $e_1(m_1) \triangleq x_1(m_1) : m_1 \in [2^{nR_1}]$ denote this encoding map. However, to construct the codebooks for encoders 2 and 3, we employ a technique based on nested coset codes. Since, the structure and encoding rules are identical for the these two encoders, we describe it using a generic index $j \in \{2, 3\}$. Let $e_j : \mathcal{F}_q \to \mathcal{X}^n_j : j = 1, 2$ denote the encoding maps. We define an NCC as follows.

Definition VII.7. An $(n, k, l, g_I, g_{O/I}, b^n)$ NCC built over a finite field $\mathcal{V} = \mathcal{F}_q$ comprises of (i) generator matrices $g_I \in \mathcal{V}^{k\times n}$, $g_{O/I} \in \mathcal{V}^{l\times n}$ (ii) a dither/bias vector b^n, an encoder map $e : \mathcal{V}^l \to \mathcal{V}^k$. We let $v^n(a, m) = ag_I \oplus_q mg_{O/I} \oplus_q b^n : (a, m) \in \mathcal{V}^k \times \mathcal{V}^l$ denote elements in its range space.

Consider two NCCs with parameters $(n, k, l, g_I, g_{O/I}, Y^n_j) : j \in \{2, 3\}$ defined using the above definition, with their range spaces denoted by $v^n_j(a_j, m_j) : j \in \{2, 3\}$, respectively. Note that the choice of g_I and $g_{O/I}$ are identical for the two NCCs. Further, let

$$\theta_j(m_j) \triangleq \sum_{a_j \in \mathcal{F}^k_q} \mathbb{1}_{\{v^n_j(a_j,m_j) \in \mathcal{T}^{(n)}_\delta(p_{V_j})\}}.$$

For every message m_j the encoder j looks for a codeword in the coset $v^n_j(a_j, m_j) : a_j \in \mathcal{F}^k_q$ that is typical according to p_{V_j}. If it finds at least one such codeword, one of them, say $v^n_j(\alpha_j(m_j), m_j)$, is chosen randomly and uniformly. $e_j(m_j)$ is generated according to $p^n_{X_j|V_j}(\cdot|v^n_j(\alpha_j(m_j), m_j))$ and is transmitted on the CQIC. Otherwise, if it finds none in the coset that is typical according to p_{V_j},,

and error is declared. This specifies the encoding rule for the three encoders. Now we describe the decoding rule.

7.4.2 Decoding Description

We begin the describing first decoder. Unlike a generic 3 CCIC decoding technique of recovering the three messages, the decoder here is constructs its POVM to recover its own message and only a bi-variate function of the two interfering messages. Since, the POVMs here require joint typicality of two messages, we employ the POVM construction similar to *Fawzi et al.* (2012), while incorporating the bi-variate function being decoded. For this, we equip the decoder 1 with the NCC $(n, k, l, g_I, g_{O/I}, b^n)$, where $b^n = b_1^n \oplus Y_2^n$.. We define $u^n(a, l)$ as the range space of the above NCC. Toward specifying the decoding POVM, we let $\pi_{m_1} \triangleq \pi_{x_1^n(m_1)}$, $\pi_{a,l} \triangleq \pi_{u^n(a,l)} \mathbb{1}_{\{u^n(a,l) \in \mathcal{T}_\delta^{(n)}(p_U)\}}$, $\pi_{m_1}^{a,l} \triangleq \pi_{x_1^n(m_1),u^n(a,l)} \mathbb{1}_{\{(x_1^n(m_1),u^n(a,l)) \in \mathcal{T}_\delta^{(n)}(p_{X_1U})\}}$, denote the conditional typical projectors (as defined in (*Wilde*, 2013a, Def. 15.2.4)) with respect to the states $\rho_{x_1}^{Y_1} \triangleq \sum_u p_U(u) \rho_{x_1,u}^{Y_1}$, $\rho_u^{Y_1} \triangleq \sum_{x_1} p_{X_1}(x_1) \rho_{x_1,u}^{Y_1}$ and $\rho_{x_1,u}^{Y_1}$, respectively, where $\rho_{x_1,u}^{Y_1}$ is as defined in the theorem statement. In addition, let $\pi_\rho^{Y_1}$ denote the typical projector with respect to the state $\rho \triangleq \sum_{x_1,u} p_{X_1}(x_1) p_U(u) \rho_{x_1,u}^{Y_1}$. Using these projectors, we define the POVM $\lambda_{\mathcal{I}_1}^{Y_1} \triangleq \{\lambda_{m_1,a,l}^{Y_1}\}$, where

$$\lambda_{m_1,a,l}^{Y_1} \triangleq \Big(\sum_{\substack{\hat{m}_1 \in \\ [2^{nR_1}]}} \sum_{\substack{\hat{a} \in \mathcal{F}_q^k \\ \hat{l} \in \mathcal{F}_q^l}} \gamma_{\hat{m}_1}^{\hat{a},\hat{l}} \Big)^{-1/2} \gamma_{m_1}^{a,l} \Big(\sum_{\substack{\hat{m}_1 \in \\ [2^{nR_1}]}} \sum_{\substack{\hat{a} \in \mathcal{F}_q^k \\ \hat{l} \in \mathcal{F}_q^l}} \gamma_{\hat{m}_1}^{\hat{a},\hat{l}} \Big)^{-1/2},$$

$\lambda_{-1} \triangleq I - \sum_{m_1 \in [2^{nR_1}]} \sum_{a \in \mathcal{F}_q^k} \sum_{l \in \mathcal{F}_q^l} \lambda_{m_1,a,l}^{Y_1}$ and $\gamma_{m_1}^{a,l} \triangleq \pi_\rho \pi_{m_1} \pi_{m_1}^{a,l} \pi_{m_1} \pi_\rho$. Having described the first decoder, we move on to describing the other two. Since these two decoders are identical, we use a generic variable j to refer to each of these. We define π_ρ^j and π_{a_j,m_j}^j as the typical (*Wilde*, 2013a, Def. 15.1.3) and the conditional typical projectors (*Wilde*, 2013a, Def. 15.2.4) with respect to the states $\rho^{Y_j} \triangleq \sum_{v_j} p_{V_j}(v_j) \rho_{v_j}^{Y_j}$ and $\rho_{v_j}^{Y_j}$, respectively. Using this, we construct the POVM $\lambda_{\mathcal{I}_j}^{Y_j} \triangleq \{\lambda_{m_j,a_j}^{Y_j}\}$, for encoder j as

$$\lambda_{a_j,m_j}^{Y_j} \triangleq \Big(\sum_{\hat{a_j} \in \mathcal{F}_q^k} \sum_{\hat{m}_j \in \mathcal{F}_q^l} \zeta_{\hat{a_j},\hat{m_j}} \Big)^{-1/2} \zeta_{a_j,m_j} \Big(\sum_{\hat{a_j} \in \mathcal{F}_q^k} \sum_{\hat{m}_j \in \mathcal{F}_q^l} \zeta_{\hat{a_j},\hat{m_j}} \Big)^{-1/2},$$

$\lambda_{-1}^{Y_j} \triangleq I - \sum_{m\in\mathcal{V}^l}\sum_{a\in\mathcal{V}^k}\lambda_{a,m}^{Y_j}$ and $\zeta_{a_j,m_j} \triangleq \pi_\rho^j \pi_{a_j,m_j}^j \pi_\rho^j$. Lastly, we provide the distribution of the random NCC.

Distribution of the Random Coset Code : The objects $g_I \in \mathcal{V}^{k\times n}, g_{O/I} \in \mathcal{V}^{l\times n}, b^n \in \mathcal{V}^n$ and the collection $(a_m \in s(m) : m \in \mathcal{V}^l)$ specify a NCC CQ-PTP code unambiguously. A distribution for a random code is therefore specified through a distribution of these objects. We let upper case letters denote the associated random objects, and obtain

$$\mathcal{P}\left(\begin{array}{c} G_I = g_I, G_{O/I} = g_{O/I} \\ B_j^n = b_j^n, \alpha_j(m_j) = a_j : m_j \in \mathcal{F}_q^l \end{array}\right) = q^{-(k+l+1)n} \prod_{m\in\mathcal{F}_q^l} \frac{1}{\Theta_j(m_j)}.$$

7.4.3 Error Analysis

As in a general information theoretic setting, we derive upper bounds on probability of error $\overline{\xi}(\underline{e},\underline{\lambda})$ by averaging over the random code of the first user and the ensemble of nested coset codes used by the other two users. The error probability of this code is given by

$$\overline{\xi}(\underline{e},\underline{\lambda}) \triangleq 1 - \frac{1}{\mathcal{M}_1\mathcal{M}_2\mathcal{M}_3}\sum_{\underline{m}\in\underline{\mathcal{M}}} \mathrm{Tr}\big(\lambda_{\underline{m}}^{\underline{Y}}\rho_{c,\underline{m}}^{\otimes n}\big). \tag{7.1}$$

Using the inequality

$$(I - \lambda_{\underline{m}}^{\underline{Y}}) \leq (I - \lambda_{m_1}^{Y_1}) \otimes I^{Y_2Y_3} + (I - \lambda_{m_1}^{Y_2}) \otimes I^{Y_1Y_3} + (I - \lambda_{m_1}^{Y_3}) \otimes I^{Y_1Y_2}, \tag{7.2}$$

from *Abeyesinghe et al.* (2009), we get $\overline{\xi}(\underline{e},\underline{\lambda}) \leq S_1 + S_2 + S_3$, where

$$S_j \triangleq \frac{1}{\underline{\mathcal{M}}}\sum_{\underline{m}\in\underline{\mathcal{M}}} \mathrm{Tr}\Big(\Big((I - \lambda_{m_j}^{Y_j}) \otimes I^{\underline{Y}\setminus B_i}\Big)\rho_{c,\underline{m}}^{\otimes n}\Big) : j \in [3].$$

Using the definition of $3to1$-CQIC, we can further simplify S_2 and S_3 as

$$S_j = \frac{1}{\mathcal{M}_j}\sum_{m_j} \mathrm{Tr}\Big((I - \lambda_{m_j}^{Y_j})\rho_{e(m_j)}\Big) : j \in \{2,3\}.$$

We first consider the terms S_2, S_3 . Note that, due to the nature of the 3to1−CQIC problem

definition, the terms S_2 and S_3 are identical to a point-to-point (PTP) setup. Therefore, to bound these terms we construct a CQ-PTP problem setup in the sequel (see Sec. 6.4) and employ that as a module in bounding S_2, S_3. The following proposition formalizes this.

Proposition VII.8. *There exists $\epsilon_S(\delta), \delta_S(\delta)$, such that for all δ and sufficiently large n, we have $\mathbb{E}\left[S_2 + S_3\right] \leq \epsilon_S(\delta)$, if $R_j \leq I(Y_j; V_j)_{\sigma_2} + \delta_S : j = 2, 3$, where $\epsilon_S, \delta_S \searrow 0$ as $\delta \searrow 0$.*

Proof. The proof is provided in Section 6.4. □

Now, we move on to bounding the term S_1. Let $\mathscr{E} \triangleq \{\theta_1(m_1) = 0 \text{ or } \theta_2(m_2) = 0\}$. By noting that $S_1 \leq 1$, we obtain $S_1 \leq S_1' + \mathbb{1}_{\mathscr{E}}$, where $S_1' \triangleq S_1 \cdot \mathbb{1}_{\mathscr{E}^c}$. As a first step, we bound the indicator $\mathbb{1}_{\mathscr{E}}$ using the following proposition.

Proposition VII.9. *There exist $\epsilon_E(\delta), \delta_E(\delta)$, such that for all δ and sufficiently large n, we have $\mathbb{E}_{\mathcal{P}}\left[\mathscr{E}\right] \leq \epsilon_E(\delta)$, if $\frac{k}{n} \geq \log q - \min\{H(V_1), H(V_2)\} + \delta_E$, where $\epsilon_E, \delta_E \searrow 0$ as $\delta \searrow 0$.*

Proof. The proof follows from (*Padakandla and Pradhan*, 2017, App. B). □

Now considering the term S_1', and using the linearity of trace while ignoring some negative terms, we get

$$\begin{aligned} S_1' &\leq \frac{1}{\underline{\mathcal{M}}} \sum_{\underline{m} \in \underline{\mathcal{M}}} \mathrm{Tr}\Big(\Big((I - \lambda^{Y_1}_{m_1,a,l}) \otimes I^{Y_2 Y_3} \Big) \rho^{\otimes n}_{c,\underline{m}} \Big) \mathbb{1}_{\mathscr{E}^c} \\ &\leq \frac{1}{\underline{\mathcal{M}}} \sum_{\underline{m} \in \underline{\mathcal{M}}} \mathrm{Tr}\Big((I - \lambda^{Y_1}_{m_1,a,l}) \pi^{a,l}_{m_1} \rho^{Y_1}_{c,\underline{m}} \pi^{a,l}_{m_1} \Big) \mathbb{1}_{\mathscr{E}^c} + S_{11}, \end{aligned} \tag{7.3}$$

where the second inequality defines the following $S_{11} \triangleq \left\| \pi^{a,l}_{m_1} \rho^{Y_1}_{c,\underline{m}} \pi^{a,l}_{m_1} - \rho^{Y_1}_{c,\underline{m}} \right\|_1$, $\rho^{Y_1}_{c,\underline{m}} \triangleq \mathrm{Tr}_{Y_2 Y_3}(\rho^{\otimes n}_{c,\underline{m}})$, $a \triangleq \alpha_1(m_1) \oplus \alpha_2(m_2)$, and $l \triangleq m_1 \oplus m_2$ and uses the inequality $\mathrm{Tr}(\lambda\rho) \leq \mathrm{Tr}(\lambda\sigma) + \|\rho - \sigma\|_1$ which holds for all $0 \leq \rho, \sigma, \lambda \leq 1$. Before we begin the proof, we provide the following lemma based on the pinching for non-commutating operators *Wilde* (2013a); *Sutter* (2018).

Lemma VII.10. *For $\pi^{a,l}_{m_1}, \pi_{m_1}, \pi^a_l$ and π_ρ as defined above, we have*

$$tr(\pi^{a,l}_{m_1} \rho^{Y_1}_{c,\underline{m}}) \geq 1 - \epsilon_{p_1}(\delta), \quad tr(\pi_{m_1} \rho^{Y_1}_{c,\underline{m}}) \geq 1 - \epsilon_{p_2}(\delta), \quad tr(\pi^a_l \rho^{Y_1}_{c,\underline{m}}) \geq 1 - \epsilon_{p_3}(\delta), \quad tr(\pi_\rho \rho^{Y_1}_{c,\underline{m}}) \geq 1 - \epsilon_{p_4}(\delta),$$

where $\epsilon_{p_i}(\delta) : i \in [4] \searrow 0$ as $\delta \searrow 0$.

Proof. The proof follows from using the ideas of pinching as provided in Lemma E.1 provided in the Appendix E.1. □

Using the above lemma, we first bound the term corresponding to S_{11}. Applying the gentle operator lemma (*Wilde*, 2013a, Lem. 9.4.2) on S_{11}, we obtain

$$S_{11} \triangleq \left\|\pi_l^a \rho_{c,\underline{m}}^{Y_1} \pi_l^a - \rho_{c,\underline{m}}^{Y_1}\right\|_1 \leq 2\sqrt{1 - \mathrm{Tr}\left(\pi_l^a \rho_{c,\underline{m}}^{Y_1}\right)} \leq 2\sqrt{\epsilon_p(\delta)}. \tag{7.4}$$

Considering the first term in the right hand side of (7.3), let T denote a generic term within its summation, defined as

$$T \triangleq \mathrm{Tr}\left((I - \lambda_{m_1,a,l}^{Y_1})\pi_l^a \rho_{c,\underline{m}}^{Y_1} \pi_l^a\right) \mathbb{1}_{\mathscr{E}^c}.$$

This term can be bounded using the Hayashi-Nagaoka inequality *Wilde* (2013a) as $T \leq 2(1-T_1) + 4T_2$, where

$$T_1 \triangleq \mathrm{Tr}\left(\gamma_{m_1}^{a,l} \pi_l^a \rho_{c,\underline{m}}^{Y_1} \pi_l^a\right), \qquad T_2 \triangleq \sum_{(m_1',a',l') \neq (m_1,a,l)} \mathrm{Tr}\left(\gamma_{m_1'}^{a',l'} \pi_l^a \rho_{c,\underline{m}}^{Y_1} \pi_l^a\right).$$

Note the objective now is to prove T_1 is close to one and T_2 is close to zero. Consider the following proposition in regards to T_1.

Proposition VII.11. *There exist $\epsilon_{T_1}(\delta), \delta_{T_1}(\delta)$, such that for all sufficiently small δ and sufficiently large n, we have $\mathbb{E}[T_1] \geq 1 - \epsilon_{T_1}(\delta)$, where $\epsilon_{T_1}, \delta_{T_1} \searrow 0$ as $\delta \searrow 0$.*

Proof. Using $\mathrm{Tr}(\lambda\rho) \geq \mathrm{Tr}(\lambda\sigma) - \|\rho - \sigma\|_1$ for $0 \leq \rho, \sigma, \lambda \leq I$, we have

$$\begin{aligned} T_1 &\geq \mathrm{Tr}\left(\pi_{m_1}^{a,l} \rho_{c,\underline{m}}^{Y_1}\right) - \left\|\pi_\rho \rho_{c,\underline{m}}^{Y_1} \pi_\rho - \rho_{c,\underline{m}}^{Y_1}\right\| - \left\|\pi_l^a \rho_{c,\underline{m}}^{Y_1} \pi_l^a - \rho_{c,\underline{m}}^{Y_1}\right\| - \left\|\pi_{m_1} \rho_{c,\underline{m}}^{Y_1} \pi_{m_1} - \rho_{c,\underline{m}}^{Y_1}\right\| \\ &\geq \mathrm{Tr}\left(\pi_{m_1}^{a,l} \rho_{c,\underline{m}}^{Y_1}\right) - 2\sqrt{1 - \mathrm{Tr}\left(\pi_\rho \rho_{c,\underline{m}}^{Y_1}\right)} - 2\sqrt{1 - \mathrm{Tr}\left(\pi_l^a \rho_{c,\underline{m}}^{Y_1}\right)} - 2\sqrt{1 - \mathrm{Tr}\left(\pi_{m_1} \rho_{c,\underline{m}}^{Y_1}\right)} \\ &\geq 1 - \epsilon_{T_1}(\delta), \end{aligned} \tag{7.5}$$

where the second inequality follows from the gentle opertor lemma and the last inequality uses the above Lemma VII.10 by defining $\epsilon_{T_1} \triangleq \epsilon_{p_1} + 2(\sqrt{\epsilon_{p_2}} + \sqrt{\epsilon_{p_3}} + \sqrt{\epsilon_{p_4}})$. This completes the proof. □

Now, we move on to bounding the term T_2. Firstly, note that the summation in T_2 can be split into seven different summations based on how many indices within the summation over the triple (m_1', a', l') are equal to (m_1, a, l). However, only three of these seven provide binding constraints on the rate triple (R_1, R_2, R_3). Building on this we perform the split $T_2 = T_{21} + T_{22} + T_{23} + T_3$, where

$$T_{21} \triangleq \sum_{m_1' \neq m_1} \mathrm{Tr}\Big(\gamma^{a,l}_{m_1'} \pi^a_l \rho^{Y_1}_{c,\underline{m}} \pi^a_l\Big), \quad T_{22} \triangleq \sum_{a' \neq a, l' \neq l} \mathrm{Tr}\Big(\gamma^{a',l'}_{m_1} \pi^a_l \rho^{Y_1}_{c,\underline{m}} \pi^a_l\Big),$$
$$T_{23} \triangleq \sum_{\substack{m_1' \neq m_1, \\ a' \neq a, l' \neq l}} \mathrm{Tr}\Big(\gamma^{a',l'}_{m_1'} \pi^a_l \rho^{Y_1}_{c,\underline{m}} \pi^a_l\Big),$$

represents the rate constraining (binding) terms while $T_3 \triangleq T_2 - \sum_{i=1}^{3} T_{2i}$ represents the inactive terms (with respect to constraining the rate). We provide the following set of propositions bounding each of these terms $T_{2i} : i \in [3]$.

Proposition VII.12. *There exists $\epsilon_{T_{21}}(\delta), \delta_{T_{21}}(\delta)$, such that for all sufficiently small δ and sufficiently large n, we have $\mathbb{E}[T_{21}] \leq \epsilon_{T_{21}}(\delta)$ if $R_1 + \frac{2k}{n} \log q \leq 2 \log q - H(V_1, V_2) + I(Y_1; X_1|U)_{\sigma_1} + \delta_{T_{21}}$, where $\epsilon_{T_{21}}, \delta_{T_{21}} \searrow 0$ as $\delta \searrow 0$.*

Proof. The proof is provided in Appendix F.1 □

Now, we provide the proposition for T_{22} as follows.

Proposition VII.13. *There exists $\epsilon_{T_{22}}(\delta), \delta_{T_{22}}(\delta)$, such that for all sufficiently small δ and sufficiently large n, we have $\mathbb{E}[T_{22}] \leq \epsilon_{T_{22}}(\delta)$ if $\frac{3k+l}{n} \log q \leq 3 \log q - H(V_1, V_2) - H(U) + I(Y_1; U|X_1)_{\sigma_1} + \delta_{T_{22}}$, where $\epsilon_{T_{22}}, \delta_{T_{22}} \searrow 0$ as $\delta \searrow 0$.*

Proof. The proof is provided in Appendix F.2 □

Proposition VII.14. *There exists $\epsilon_{T_{23}}(\delta), \delta_{T_{23}}(\delta)$, such that for all sufficiently small δ and sufficiently large n, we have $\mathbb{E}[T_{23}] \leq \epsilon_{T_{23}}(\delta)$ if $R_1 + \frac{3k+l}{n} \log q \leq 3 \log q - H(V_1, V_2) - H(U) + I(Y_1; X_1, U)_{\sigma_1} + \delta_{T_{23}}$, where $\epsilon_{T_{23}}, \delta_{T_{23}} \searrow 0$ as $\delta \searrow 0$.*

Proof. The proof is provided in Appendix F.3 □

For the terms in the expression T_3, we split T_3 as $T_3 \triangleq T_{31} + T_{32} + T_{33} + T_{34}$, where

$$T_{31} \triangleq \sum_{a' \neq a} \mathrm{Tr}\Big(\gamma_{m_1}^{a',l} \pi_l^a \rho_{c,\underline{m}}^{Y_1} \pi_l^a\Big), \quad T_{32} \triangleq \sum_{l' \neq l} \mathrm{Tr}\Big(\gamma_{m_1}^{a,l'} \pi_l^a \rho_{c,\underline{m}}^{Y_1} \pi_l^a\Big),$$
$$T_{33} \triangleq \sum_{\substack{m_1' \neq m_1, \\ a' \neq a}} \mathrm{Tr}\Big(\gamma_{m_1'}^{a',l} \pi_l^a \rho_{c,\underline{m}}^{Y_1} \pi_l^a\Big), \quad T_{34} \triangleq \sum_{\substack{m_1' \neq m_1, \\ l' \neq l}} \mathrm{Tr}\Big(\gamma_{m_1'}^{a,l'} \pi_l^a \rho_{c,\underline{m}}^{Y_1} \pi_l^a\Big). \tag{7.6}$$

As mentioned earlier, the analysis of these term follows from the analysis performed for the terms $T_{2i}, i \in [3]$, and further these terms do not contribute to any new additional rate constraints. However, for the sake of completeness, we briefly indicate how each of these term scan be bounded using the ones corresponding to T_2. To begin with, consider the terms T_{31} and T_{32}. One can perform identical analysis as for the term T_{22} (see Appendix F.2) and obtain the following bounds.

$$\mathbb{E}[T_{31}] \leq \exp_2\left(\frac{3k}{n}\log q - (3\log q - H(V_1,V_2) - H(U) + I(Y_1;U|X_1)_{\sigma_1} + \delta_{T_{22}})\right), \tag{7.7}$$
$$\mathbb{E}[T_{32}] \leq \exp_2\left(\frac{2k+l}{n}\log q - (3\log q - H(V_1,V_2) - H(U) + I(Y_1;U|X_1)_{\sigma_1} + \delta_{T_{22}})\right), \tag{7.8}$$

where $\exp_2(x) \triangleq 2^x$. Note that the exponents in the right hand side terms (7.7) and (7.8) are always negative given the bound in Proposition VII.13 is true. Hence T_{31} and T_{32} can be made arbitrarily small for sufficiently large n without any additional constraints.

Similarly, consider the terms T_{33} and T_{34}. Using an identical analysis as for the term T_{23} (see Appendix F.3) we obtain

$$\mathbb{E}[T_{33}] \leq \exp_2\left(R_1 + \frac{3k}{n}\log q - (3\log q - H(V_1,V_2) - H(U) + I(Y_1;X_1,U)_{\sigma_1} + \delta_{T_{23}})\right), \tag{7.9}$$
$$\mathbb{E}[T_{32}] \leq \exp_2\left(R_1 + \frac{2k+l}{n}\log q - (3\log q - H(V_1,V_2) - H(U) + I(Y_1;X_1,U)_{\sigma_1} + \delta_{T_{23}})\right), \tag{7.10}$$

Again observe that the exponents in the right hand side of (7.9) and (7.10) are always negative given the bound in Proposition VII.14 is true. Hence T_{33} and T_{34} can be made arbitrarily small for sufficiently large n without any additional constraints. Having completed the proof for the terms in T, we now provide the result stating: NCC codes achieve capacity of a CQ-PTP channel (as discussed in the proof of Proposition VII.8). □

7.5 Rate-region using NCC and message splitting for 3to1−CQIC

Theorem VII.15. *Given a 3to1-CQIC* $(\rho_{\underline{x}} \in \mathcal{D}(\mathcal{H}_{\underline{Y}}) : \underline{x} \in \underline{\mathcal{X}})$ *and a PMF* $p_{U_2U_3V_2V_3X_2X_3} = p_{U_2V_2X_2}p_{U_3V_3X_3}$ *on* $\mathcal{U}_1 \times \mathcal{V}_1 \times \mathcal{X}_1 \times \mathcal{U}_2 \times \mathcal{V}_2 \times \mathcal{X}_2$ *where* $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{F}_q$*, a rate triple is achievable if it satisfies the following:* $R_j \leq I(U_jX_j; Y_j)_{\sigma_j}$,

$$R_1 \leq \min_{j=2,3}\{0, H(U_j) - H(W|Y_1)_{\sigma_1}\} + I(X_1; WY_1)_{\sigma_1}$$

$$R_1 + R_j \leq I(X_j; Y_j|U_j)_{\sigma_j} + I(X_1; W, Y_1)_{\sigma_1} + H(U_j) - H(W|Y_1)_{\sigma_1}$$

for $j = 2, 3$*, where*

$$\sigma_1^{\underline{Y}} \triangleq \sum_{x_1 \in \mathcal{X}_1, w \in \mathcal{F}_q} p_{X_1}(x_1)p_W(w)\rho^{\underline{Y}}_{x_1,w} \otimes |x_1\rangle\langle x_1| \otimes |w\rangle\langle w|,$$

$$\rho^{\underline{Y}}_{x_1,w} \triangleq \sum_{\substack{u_2,v_2,x_2\\ u_3,v_3,x_3}} p_{V_2,V_3U_2U_3X_2X_3|W}(v_2, v_3, u_2, u_3, x_2, x_3|w)\rho^{\underline{Y}}_{\underline{x}},$$

$$\sigma_2 \triangleq \sum_{v_1,v_2,v_3} p_{U_2U_3V_2V_3\underline{X}}(u_2, u_3, v_2, v_3, \underline{x})\rho^{\underline{Y}}_{\underline{x}} \otimes_{j=2}^{3} |u_j, x_j\rangle\langle u_j, x_j|,$$

for $W \triangleq U_2 \oplus U_3$*, and* $\{|u_j\rangle\}$ *and* $\{|x_j\rangle\}$ *as some orthonormal basis on* $\mathcal{H}_Y$ *for* $j = 2, 3$.

Proof. In view of the detailed proof provided for Thms. VII.5, VI.3, we only provide an outline. A complete proof of this theorem is beyond the scope of this thesis and can be constructed using techniques developed in *Sen* (2021); *Sohail et al.* (2022).

In the coding scheme of Thm. VII.5, Rx 1 decodes a bivariate function of Tx 2 and Tx 3's inputs. In general, decoding just a bivariate function of Tx 2 and Tx 3's inputs is insufficient. It is necessary for the coding scheme to permit Rx 1 decode univariate functions of the Tx 2 and Tx 3's inputs as well. Therefore an enhanced coding scheme, will split Tx 2 and Tx 3's transmissions into two parts respectively. For $j = 2, 3$, let U_j, V_j denote the splitting of Tx j's input. Here, $U_2, U_3 \in \mathcal{F}_q$ take values in a common finite field. U_2 and U_3 are communicated via a common nested coset code. V_2, V_3, X_1 are built via conventional unstructured codes. Since this is a 3to1−IC, Tx 1 does not split its input X_1.

Observe that, for $j \in 2, 3$, Rx j has to decode a U_j, V_j, one component of which is encoded via a nested coset code, and the other component which is encoded via a conventional unstructured code.

The analysis of its decoding is similar to the analysis of Tx 1's decoding in proof of Theorem VII.5. Indeed, in proof of Theorem VII.5, Tx 1 decoded from its unstructured code and the bivariate component of the interference that was encoded via a nested coset code. This provides the outline for the analysis of Rx 2 and 3. Rx 1 has to decode 4 components - 1 structured $(U_2 \oplus U_3)$ and 3 unstructured V_2, V_3, X_1. We adopt successive-simultaneous decoding wherein two code words are decoded at each stage of a 2 stage process. □

By choosing $W = \phi$, we can recover the $\mathscr{USB}-$rate region from the above inner bound.

7.6 Conclusion

In this chapter, we considered the problem of characterizing an inner bound to the capacity region of a 3−user classical-quantum interference channel (3−CQIC). The best known coding scheme for communicating over CQICs is based on unstructured random codes and employs the techniques of message splitting and superposition coding. For classical 3−user interference channels (ICs), it has been proven that coding techniques based on coset codes - codes possessing algebraic closure properties - strictly outperform all coding techniques based on unstructured codes. In this work, we developed analogous techniques based on coset codes for 3to1−CQICs - a subclass of 3−user CQICs. We analyzed its performance and derived a new inner bound to the capacity region of 3to1−CQICs that subsumes the current known largest and strictly enlarges the same for identified examples.

APPENDICES

APPENDIX A

Proofs of Chapter II

A.1 Proof of Theorem II.9

Note that $\theta = 1 - \sum_{x\in\mathcal{X}} \lambda_x$. Define $\tilde{\rho}'_x \triangleq \Pi_\rho \Pi_x \hat{\rho}_x \Pi_x \Pi_\rho$, and $\sigma' \triangleq \frac{1}{1-\theta} \sum_{x\in\mathcal{X}} \lambda_x \tilde{\rho}'_x$. Further let $\hat{\Pi}$ be the projector onto the subspace spanned by the eigenspace of σ' corresponding to the eigenvalues greater than ϵ/D. Let $\tilde{\rho}_x \triangleq \hat{\Pi}\tilde{\rho}'_x\hat{\Pi}$, and $\sigma \triangleq \hat{\Pi}\sigma'\hat{\Pi}$.

Construction of Random POVMS: Define a collection of of random codes $\mathcal{C} \triangleq \{\mathcal{C}^{(\mu)}\}$ for $\mu \in [1, N]$, where $\mathcal{C}^{(\mu)} \triangleq \{X(l,\mu)\}_{l\in[1,K]}$, and $X(l,\mu)$ are chosen randomly, independently according to the distribution $\{\lambda_x/(1-\theta)\}_{x\in\mathcal{X}}$. Using this, define

$$\gamma_x^{(\mu)} \triangleq \frac{(1-\theta)}{(1+\epsilon)} \frac{1}{K} |\{l : X(l,\mu) = x\}| = \frac{(1-\theta)}{(1+\epsilon)} \frac{1}{K} \sum_{l=1}^{K} \mathbb{1}_{\{X(l,\mu)=x\}},$$

and $A_x^{(\mu)} \triangleq \gamma_x^{(\mu)} \sqrt{\rho}^{-1} \tilde{\rho}_x \sqrt{\rho}^{-1}$, where $\sqrt{\rho}^{-1}$ refers to the generalized inverse as defined in (*Holevo*, 2012, Section 5.6). Now for each $\mu \in [1, N]$, construct a collection of non-negative operators $\tilde{M}^{(\mu)} \triangleq \{A_x^{(\mu)}\}_{x\in\mathcal{X}}$.

Proposition A.1. *$\tilde{M}^{(\mu)}$ forms a sub-POVM for all $\mu \in [1, N]$ with probability exceeding $1 - 2ND \exp\left[-\frac{K\epsilon^2 d\epsilon D^{-1}}{4\ln 2}\right]$.*

Proof. We use the operator Chernoff bound *Wilde* (2013a). Note that

$$\tilde{\rho}_x \leq d^{-1}\hat{\Pi}\Pi_\rho\Pi_x\Pi_\rho\hat{\Pi} \leq d^{-1}\hat{\Pi},$$

where we used the hypothesis (2.7d) assumed in the theorem statement. Moreover,

$$\mathbb{E}[\tilde{\rho}_{X(l,\mu)}] = \hat{\Pi}\sigma'\hat{\Pi} \geq \frac{\epsilon}{D}\hat{\Pi}.$$

Applying the operator Chenoff bound on $\{d\tilde{\rho}_{X(l,\mu)}\}_{l\in[1,K]}$, we obtain

$$P\left\{(1-\epsilon)\sigma \leq \frac{1}{K}\sum_{l=1}^{K}\tilde{\rho}_{X(l,\mu)} \leq (1+\epsilon)\sigma\right\} \geq 1 - 2D\exp\left[-\frac{K\epsilon^3 dD^{-1}}{4\ln 2}\right],$$

for all $\mu \in [1, N]$, where we used the fact that $\mathrm{Tr}\left(\hat{\Pi}\right) \leq \mathrm{Tr}(\Pi_\rho) \leq D$ (using the hypothesis (2.7c) of the theorem statement). Now we have

$$\sigma = \hat{\Pi}\sigma'\hat{\Pi} \leq \sigma' \leq \frac{1}{1-\theta}\Pi_\rho\rho\Pi_\rho \leq \frac{1}{1-\theta}\rho,$$

using the hypothesis (2.7e) and (2.7f) of the theorem statement. This results in $(1-\theta)\sqrt{\rho}^{-1}\sigma\sqrt{\rho}^{-1} \leq I$. This implies that with probability exceeding $1 - 2D\exp\left[-\frac{K\epsilon^3 dD^{-1}}{4\ln 2}\right]$, we have

$$\sum_{x\in\mathcal{X}} A_x^{(\mu)} = \sum_{x\in\mathcal{X}} \gamma_x^{(\mu)}\sqrt{\rho}^{-1}\tilde{\rho}_x\sqrt{\rho}^{-1} = \frac{1}{K}\frac{(1-\theta)}{(1+\epsilon)}\sqrt{\rho}^{-1}\left(\sum_{l=1}^{K}\tilde{\rho}_{X(l,\mu)}\right)\sqrt{\rho}^{-1} \leq I.$$

Hence using the union bound, we see that with probability exceeding $1 - 2ND\exp\left[-\frac{K\epsilon^2 d\epsilon D^{-1}}{4\ln 2}\right]$, we have $\{A_x^{(\mu)}\}_{x\in\mathcal{X}}$ forming a sub-POVM for all $\mu \in [1, N]$. □

Let $\tilde{M} \triangleq \{\tilde{\Lambda}_x\}_{x\in\mathcal{X}}$, $\tilde{M}^{(\mu)} \triangleq \{A_x^{(\mu)}\}_{x\in\mathcal{X}}$, where $\tilde{\Lambda}_x \triangleq \frac{1}{N}\sum_{\mu=1}^{N} A_x^{(\mu)}$. Let $\mathbb{1}_{\{\mathrm{sP}\}}$ denote the indicator random variable corresponding to the event that $\tilde{M}^{(\mu)}$ forms a sub-POVM for all $\mu \in [1, N]$. The completion of the sub-POVM is given by $I - \sum_{x\in\mathcal{X}}\tilde{\Lambda}_x$. We use the trivial POVM $\{I\}$ in the case of the complementary event. Using this construction, we have

$$\begin{aligned}
&\Xi_\rho(M, \tilde{M}) \\
&\leq \mathbb{1}_{\{\mathrm{sP}\}}\left[\sum_{x\in\mathcal{X}}\|\sqrt{\rho}(\Lambda_x - \tilde{\Lambda}_x)\sqrt{\rho}\|_1 + \mathrm{Tr}\left((I - \sum_{x\in\mathcal{X}}\tilde{\Lambda}_x)\rho\right)\right] + 2(1 - \mathbb{1}_{\{\mathrm{sP}\}}) + \theta \\
&\leq \sum_{x\in\mathcal{X}}\left\|\lambda_x\hat{\rho}_x - \frac{1}{N}\sum_{\mu=1}^{N}\gamma_x^{(\mu)}\tilde{\rho}_x\right\|_1 + \left\|\rho - \frac{1}{N}\sum_{\mu,x}\gamma_x^{(\mu)}\tilde{\rho}_x\right\|_1 + 2(1 - \mathbb{1}_{\{\mathrm{sP}\}}) + \theta
\end{aligned}$$

$$
\begin{aligned}
&\overset{(a)}{\leq} \sum_{x\in\mathcal{X}} \left\| \lambda_x \hat{\rho}_x - \frac{1}{N}\sum_{\mu=1}^{N} \gamma_x^{(\mu)} \tilde{\rho}_x \right\|_1 + \sum_{x\in\mathcal{X}} \left\| \lambda_x \hat{\rho}_x - \frac{1}{N}\sum_{\mu=1}^{N} \gamma_x^{(\mu)} \tilde{\rho}_x \right\|_1 + 2(1-\mathbb{1}_{\{\mathrm{sP}\}}) + 2\theta \\
&\overset{(b)}{\leq} 2\sum_{x\in\mathcal{X}} \left\| \lambda_x \hat{\rho}_x - \frac{1}{N}\sum_{\mu=1}^{N} \gamma_x^{(\mu)} \hat{\rho}_x \right\|_1 + 2\sum_{x\in\mathcal{X}} \frac{1}{N}\sum_{\mu=1}^{N} \gamma_x^{(\mu)} \|\hat{\rho}_x - \tilde{\rho}_x\|_1 + 2(1-\mathbb{1}_{\{\mathrm{sP}\}}) + 2\theta \\
&\overset{(c)}{=} 2[S_1 + S_2] + 2(1-\mathbb{1}_{\{\mathrm{sP}\}}) + 2\theta,
\end{aligned}
$$

where (a) follows by triangle inequality, (b) follows by adding and subtracting $\frac{1}{N}\sum_{\mu}\gamma_x^{(\mu)}\hat{\rho}_x$, and (c) follows by defining

$$
S_1 \triangleq \sum_{x\in\mathcal{X}} \left\| \lambda_x \hat{\rho}_x - \frac{1}{N}\sum_{\mu=1}^{N} \gamma_x^{(\mu)} \hat{\rho}_x \right\|_1, \quad S_2 \triangleq \sum_{x\in\mathcal{X}} \frac{1}{N}\sum_{\mu=1}^{N} \gamma_x^{(\mu)} \|\hat{\rho}_x - \tilde{\rho}_x\|_1.
$$

We work on the first term S_1 as follows. Note that

$$
S_1 \leq S_1' + \left(\frac{\epsilon}{1+\epsilon}\right) \sum_{x\in\mathcal{X}} \lambda_x \leq S_1' + \left(\frac{\epsilon}{1+\epsilon}\right),
$$

where

$$
S_1' \triangleq \frac{1}{1+\epsilon} \sum_{x\in\mathcal{X}} \left| \lambda_x - \frac{(1-\theta)}{NK} \sum_{\mu,l} \mathbb{1}_{\{X(l,\mu)=x\}} \right|.
$$

Note that

$$
\mathbb{E}[S_1'] = \frac{1}{(1+\epsilon)} \sum_{x\in\mathcal{X}} \mathbb{E}\left[\left| \hat{P}(x) - \mathbb{E}[\hat{P}(x)] \right| \right] \leq \frac{1}{(1+\epsilon)} \sum_{x\in\mathcal{X}} \sqrt{\mathrm{Var}(\hat{P}(x))} \leq \frac{1}{(1+\epsilon)} \sum_{x\in\mathcal{X}} \sqrt{\frac{\lambda_x}{NK}},
$$

where we have defined $\hat{P}(x) \triangleq \frac{(1-\theta)}{NK}\sum_{l,\mu}\mathbb{1}_{\{X(l,\mu)=x\}}$. Hence

$$
\mathbb{E}[S_1] \leq \frac{1}{(1+\epsilon)\sqrt{NK}} \sum_{x\in\mathcal{X}} \sqrt{\lambda_x} + \left(\frac{\epsilon}{1+\epsilon}\right).
$$

Moving on to S_2, consider the following.

$$
\begin{aligned}
2\mathbb{E}[S_2] &\leq \frac{2}{(1+\epsilon)} \left[\sum_{x\in\mathcal{X}} \lambda_x \|\hat{\rho}_x - \tilde{\rho}_x'\|_1 + \sum_{x\in\mathcal{X}} \lambda_x \|\tilde{\rho}_x' - \tilde{\rho}_x\|_1 \right] \\
&\leq \frac{1}{(1+\epsilon)} \left[4\sqrt{\epsilon} + 4\sqrt{\epsilon + 2\sqrt{\epsilon}} + 4\sqrt{2}(1-\theta)\sqrt{\epsilon+\sqrt{\epsilon}} \right]
\end{aligned}
$$

$$= f(\epsilon, \theta),$$

where we have used the ensemble gentle measurement lemma *Wilde* (2013a). Combining all the arguments, we see that

$$\mathbb{E}(\Xi_\rho(M, \tilde{M})) \leq \frac{2}{(1+\epsilon)\sqrt{NK}} \sum_{x \in \mathcal{X}} \sqrt{\lambda_x} + \frac{2\epsilon}{\epsilon + 1} + f(\epsilon, \theta) + 4DN \exp\left[-\frac{K\epsilon^3 dD^{-1}}{4 \ln 2}\right] + 2\theta.$$

There must exists a collection of sub-POVMs whose average performance is at least as good.

A.2 Proof of Lemma II.11

Consider the left hand side of (2.18). We define an operator Λ_{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 $\mathcal{Y}^+ \triangleq \mathcal{Y} \bigcup \{y_0\}$. Since trace norm is invariant to transposition with respect to ρ_{AB}, we can write for any $y \in \mathcal{Y}^+$,

$$\begin{aligned} \left\| \sqrt{\rho_{AB}} \left(\Gamma^A \otimes \Lambda_y^B\right) \sqrt{\rho_{AB}} \right\|_1 &= \left\| \left[\sqrt{\rho_{AB}} \left(\Gamma^A \otimes \Lambda_y^B\right) \sqrt{\rho_{AB}} \right]^T \right\|_1 \\ &= \left\| \sqrt{\rho_{AB}} \left(\left(\Gamma^A\right)^T \otimes \left(\Lambda_y^B\right)^T \right) \sqrt{\rho_{AB}} \right\|_1. \end{aligned} \tag{A.1}$$

One can easily prove for any Γ_A (not necessarily positive) that

$$\left(\sqrt{\rho_{AB}} \left(\left(\Gamma^A\right)^T \otimes \left(\Lambda_y^B\right)^T \right) \sqrt{\rho_{AB}} \right)^R = \mathrm{Tr}_{AB} \left\{ \left(\mathrm{id} \otimes \Gamma^A \otimes \Lambda_y^B \right) \Psi_{RAB} \right\}, \tag{A.2}$$

where Ψ_{RAB} is the canonical purification of ρ_{AB} defined as $\Psi_{RAB} \triangleq \sum_{x,x'} \sqrt{\lambda_x \lambda_{x'}} \left|x\right\rangle \left\langle x'\right|_{AB} \otimes \left|x\right\rangle \left\langle x'\right|_R$ for the spectral decomposition of ρ_{AB} given as $\rho_{AB} = \sum_x \lambda_x \left|x\right\rangle\!\left\langle x\right|_{AB}$ and $(\cdot)^R$ represents a state in the reference Hilbert space R. Now, using (A.2) we perform the following simplification

$$\begin{aligned} &\sum_{y \in \mathcal{Y}} \left\| \sqrt{\rho_{AB}} \left(\Gamma^A \otimes \Lambda_y^B\right) \sqrt{\rho_{AB}} \right\|_1 \overset{(a)}{\leq} \sum_{y \in \mathcal{Y}^+} \left\| \sqrt{\rho_{AB}} \left(\Gamma^A \otimes \Lambda_y^B\right) \sqrt{\rho_{AB}} \right\|_1 \\ &= \sum_{y \in \mathcal{Y}^+} \left\| \mathrm{Tr}_{AB} \left\{ \left(\mathrm{id}_R \otimes \Gamma^A \otimes \Lambda_y^B \right) \Psi_{RAB} \right\} \right\|_1 \\ &\overset{(b)}{=} \left\| \sum_{y \in \mathcal{Y}^+} \mathrm{Tr}_{AB} \left\{ \left(\mathrm{id}_{RB} \otimes \Gamma^A \right) \left(\mathrm{id}_{RA} \otimes \Lambda_y^B \right) \Psi_{RAB} \otimes \left|y\right\rangle\!\left\langle y\right| \right\} \right\|_1 \end{aligned}$$

$$
\begin{aligned}
&= \left\| \mathrm{Tr}_A \left\{ \left(\mathrm{id}_{RY} \otimes \Gamma^A\right) \left(\sum_{y \in \mathcal{Y}^+} |y\rangle\langle y| \otimes \mathrm{Tr}_B \left\{ \left(\mathrm{id}_{RA} \otimes \Lambda_y^B\right) \Psi_{RAB} \right\} \right) \right\} \right\|_1 \\
&\overset{(c)}{=} \left\| \mathrm{Tr}_A \left\{ \left(\mathrm{id}_{RY} \otimes \Gamma^A\right) \sigma_{RAY} \right\} \right\|_1 \\
&\overset{(d)}{=} \left\| \mathrm{Tr}_{AZ} \left\{ \left(\mathrm{id}_{RY} \otimes \Gamma^A \otimes \mathrm{id}_Z\right) \Phi_{RAYZ}^{\sigma_{RAY}} \right\} \right\|_1 , \qquad \text{(A.3)}
\end{aligned}
$$

where (a) follows from the fact that $\left\| \sqrt{\rho_{AB}} \left(\Gamma^A \otimes \Lambda_{y_0}^B\right) \sqrt{\rho_{AB}} \right\|_1$ is always non-negative, (b) uses the triangle inequality for block diagonal operators, (c) uses σ_{RAY} defined as

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

and finally, (d) uses $\Phi_{RAYZ}^{\sigma_{RAY}}$ defined as the canonical purification of σ_{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 (A.1) and (A.2), we have

$$
\begin{aligned}
&\left\| \mathrm{Tr}_{AZ} \left\{ \left(\mathrm{id}_{RY} \otimes \Gamma^A \otimes \mathrm{id}_Z\right) \Phi_{RAYZ}^{\sigma_{RAY}} \right\} \right\|_1 \\
&\quad = \left\| \sqrt{\mathrm{Tr}_{RYZ} \left\{ \Phi_{RAYZ}^{\sigma_{RAY}} \right\}} \Gamma^A \sqrt{\mathrm{Tr}_{RYZ} \left\{ \Phi_{RAYZ}^{\sigma_{RAY}} \right\}} \right\|_1 = \| \sqrt{\rho_A}\, \Gamma^A \sqrt{\rho_A} \|_1 .
\end{aligned}
$$

This completes the proof.

A.3 Proof of Lemma II.12

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

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

Next, by adding and subtracting appropriate terms, we get

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

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

where the second inequality follows by applying Lemma II.11 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.

A.4 Proof of Lemma II.24

Proof. Using the chain rule of quantum mutual information we see that

$$I(A;B|C,J)_\sigma=S(ACJ)_\sigma+S(BCJ)_\sigma-S(ABCJ)_\sigma-S(CJ)_\sigma.$$

The eigenvectors of the state σ_{ABCJ} are of the form $(0,\ldots,0,|j\rangle\otimes\left|v_i^j\right\rangle,0,\ldots,0)$, with eigenvalue $P_J(j)\lambda_i^j$, where $\left|v_i^j\right\rangle$ is an eigenvector of state ρ_{ABC}^j with eigenvalue λ_i^j. Hence

$$
\begin{aligned}
S(ABCJ)&=-\sum_{j,i}P_J(j)\lambda_i^j\log\Big(P_J(j)\lambda_i^j\Big)\\
&\overset{(a)}{=}H(P_J)+\sum_{j=1}^n P_J(j)\sum_i[-\lambda_i^j\log\lambda_i^j]
\end{aligned}
$$

$$= S(J)_\sigma + \sum_{j=1}^{n} P_J(j) S(ABC)_{\rho^j},$$

where in (a) we used the grouping axiom of entropy. Applying similar arguments for $S(ACJ)$, $S(BCJ)$, and $S(CJ)$ we get the desired result. □

A.5 Proof of Lemma II.30

Proof. Consider the trace norm expression given in (2.33). This expression can be bounded from above using the triangle inequality as

$$\begin{aligned}
&\left\| \sum_{w^n} \lambda_{w^n} \mathcal{T}_{w^n} - \frac{1}{2^{n(R+C)}} \frac{(1-\varepsilon)}{(1+\eta)} \sum_{l,\mu} \mathcal{T}_{W^{n,(\mu)}(l)} \right\|_1 \\
&\qquad \leq \left\| \sum_{w^n} \lambda_{w^n} \mathcal{T}_{w^n} - \frac{(1-\varepsilon)}{(1+\eta)} \sum_{\substack{w^n \in \\ \mathcal{T}_\delta^{(n)}(W)}} \tilde{P}_{W^n}(w^n) \mathcal{T}_{w^n} \right\|_1 \\
&\qquad\quad + \frac{(1-\varepsilon)}{(1+\eta)} \left\| \sum_{\substack{w^n \in \\ \mathcal{T}_\delta^{(n)}(W)}} \tilde{P}_{W^n}(w^n) \mathcal{T}_{w^n} - \frac{1}{2^{n(R+C)}} \sum_{l,\mu} \mathcal{T}_{W^{n,(\mu)}(l)} \right\|_1 .
\end{aligned} \tag{A.4}$$

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

$$\begin{aligned}
&\Big\| \sum_{w^n} \lambda_{w^n} \mathcal{T}_{w^n} - \frac{(1-\varepsilon)}{(1+\eta)} \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \tilde{P}_{W^n}(w^n) \mathcal{T}_{w^n} \Big\|_1 \\
&\qquad \leq \Big\| \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \left(1 - \frac{1}{(1+\eta)} \right) \mathcal{T}_{w^n} \Big\|_1 + \Big\| \sum_{w^n \notin \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \mathcal{T}_{w^n} \Big\|_1 \\
&\qquad \leq \left(\frac{\eta}{1+\eta} \right) \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \underbrace{\|\mathcal{T}_{w^n}\|_1}_{=1} + \sum_{w^n \notin \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \underbrace{\|\mathcal{T}_{w^n}\|}_{=1} \\
&\qquad \leq \left(\frac{\eta}{1+\eta} \right) + \varepsilon \leq \eta + \varepsilon \leq \frac{\epsilon}{2},
\end{aligned} \tag{A.5}$$

for all η sufficiently small and n sufficiently large. Now consider the second term in (A.4). Using the covering lemma from *Wilde* (2013a), this can be bounded as follows. For $w^n \in \mathcal{T}_\delta^{(n)}(W)$,

let Π and Π_{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 $w^n \in \mathcal{T}_\delta^{(n)}(W)$:

$$\begin{aligned} \operatorname{Tr}\{\Pi \mathcal{T}_{w^n}\} &\geq 1 - \epsilon_1, \\ \operatorname{Tr}\{\Pi_{w^n} \mathcal{T}_{w^n}\} &\geq 1 - \epsilon_1, \\ \operatorname{Tr}\{\Pi\} &\leq D, \\ \Pi_{w^n} \mathcal{T}_{w^n} \Pi_{w^n} &\leq \frac{1}{d} \Pi_{w^n}, \end{aligned} \tag{A.6}$$

where $D = 2^{n(S(\mathcal{T})+\delta_1)}$ and $d = 2^{n\left[\left(\sum_w \lambda_w S(\mathcal{T}_w)\right)-\delta_2\right]}$, and $\delta_1(\delta) \searrow 0, \delta_2(\delta) \searrow 0$ as $\delta \searrow 0$. 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 Π and Π_{w^n} such that they satisfy the set of inequalities in (A.6), then for all sufficiently large n, if $n(R+C) > \log_2 \frac{D}{d}$, the obfuscation error, defined as

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

can be made smaller than $\epsilon_1 + 4\sqrt{\epsilon_1} + 24\sqrt[4]{\epsilon_1}$ with high probability. This gives us the the following rate constraints $R + C > S(\sum_w \lambda_w \mathcal{T}_w) - \sum_w \lambda_w S(\mathcal{T}_w) + \delta_1 + \delta_2 = \chi\left(\{\lambda_w\}, \{\hat{\rho}_w \otimes \sigma_w\}\right) + \delta_1 + \delta_2$. Using this constraint and the bound from (A.5), the result follows. □

A.6 Proof of Proposition II.14

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

$$\begin{aligned} &(\mathrm{id} \otimes [\tilde{M}_{AB}])(\Psi_{RAB}^\rho) \\ &= \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{i,j} \Phi_{F^{(\mu_1,\mu_2)}(i,j)} \otimes \operatorname{Tr}_{AB} \left\{ (\mathrm{id} \otimes \Gamma_i^{A,(\mu_1)} \otimes \Gamma_j^{B,(\mu_2)}) \Psi_{RAB}^\rho \right\} \\ &\quad + (1 - \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}}) \Phi_{(0_U,0_V)} \otimes \operatorname{Tr}_{AB} \left\{ (\mathrm{id} \otimes \mathrm{id} \otimes \mathrm{id}) \Psi_{RAB}^\rho \right\} \\ &= \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \end{aligned}$$

$$
\begin{aligned}
\times \Bigg[& \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{i,j\geq 1} \sum_{(u,v)\in \mathcal{B}_1^{(\mu_1)}(i)\times \mathcal{B}_2^{(\mu_2)}(j)} \Phi_{e^{(\mu_1,\mu_2)}(u,v)} \otimes \mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi_{RAB}^{\rho}\right\} \\
& + \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{j\geq 1} \sum_{v\in \mathcal{B}_2^{(\mu_2)}(j)} \Phi_{(0_U,0_V)} \otimes \mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes A_{0_U}^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi_{RAB}^{\rho}\right\} \\
& + \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{i\geq 1} \sum_{u\in \mathcal{B}_1^{(\mu_1)}(i)} \Phi_{(0_U,0_V)} \otimes \mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_{0_V}^{(\mu_2)})\Psi_{RAB}^{\rho}\right\} \\
& + \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \Phi_{(0_U,0_V)} \otimes \mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes A_{0_U}^{(\mu_1)}\otimes B_{0_V}^{(\mu_2)})\Psi_{RAB}^{\rho}\right\}\Bigg] \\
& + (1-\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}})\Phi_{(0_U,0_V)} \otimes \mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes \mathrm{id}\otimes \mathrm{id})\Psi_{RAB}^{\rho}\right\}. \qquad \text{(A.7)}
\end{aligned}
$$

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

$$
\begin{aligned}
& \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} (\mathrm{id}\otimes [M_1^{(\mu_1)}]\otimes [M_2^{(\mu_2)}])(\Psi_{RAB}^{\rho}) \\
& = \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\Bigg[\frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{u\in\mathcal{U}} \sum_{v\in\mathcal{V}} \Phi_{(u,v)} \otimes \mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi_{RAB}^{\rho}\right\} \\
& \quad + \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{v\in\mathcal{V}} \Phi_{(0_U,v)} \otimes \mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes A_{0_U}^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi_{RAB}^{\rho}\right\} \\
& \quad + \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{u\in\mathcal{U}} \Phi_{(u,0_V)} \otimes \mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_{0_V}^{(\mu_2)})\Psi_{RAB}^{\rho}\right\} \\
& \quad + \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \Phi_{(0_U,0_V)} \otimes \mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes A_{0_U}^{(\mu_1)}\otimes B_{0_V}^{(\mu_2)})\Psi_{RAB}^{\rho}\right\}\Bigg] \\
& \quad + (1-\mathbb{1}_{\{\text{sP}-1\}})\mathbb{1}_{\{\text{sP}-2\}} \frac{1}{N_2} \sum_{\mu_2} \sum_{v\in\mathcal{V}\cup\{0_V\}} \Phi_{(0_U,v)} \otimes \mathrm{Tr}_{AB}\{\mathrm{id}\otimes \mathrm{id}\otimes B_v^{(\mu_2)})\Psi_{RAB}\} \\
& \quad + (1-\mathbb{1}_{\{\text{sP}-2\}})\mathbb{1}_{\{\text{sP}-1\}} \frac{1}{N_1} \sum_{\mu_1} \sum_{u\in\mathcal{U}\cup\{0_U\}} \Phi_{(u,0_V)} \otimes \mathrm{Tr}_{AB}\{\mathrm{id}\otimes A_u^{(\mu_1)}\otimes \mathrm{id})\Psi_{RAB}\} \\
& \quad + (1-\mathbb{1}_{\{\text{sP}-2\}})(1-\mathbb{1}_{\{\text{sP}-1\}})\Phi_{(0_U,0_V)}\,\mathrm{Tr}_{AB}\{(\mathrm{id}\otimes \mathrm{id}\otimes \mathrm{id})\Psi_{RAB}\}. \qquad \text{(A.8)}
\end{aligned}
$$

By replacing the terms in S_2 using the corresponding expansions from (A.7) and (A.8), we observe that the fourth terms on the right hand side of (A.7) get canceled with the corresponding terms on the right hand side of (A.8). Next we take the second term in (A.7) and apply the triangle

inequality and bound from above its l_1 norm by

$$\frac{1}{N_1}\sum_{\mu_1}\mathrm{Tr}\left(\left(I-\sum_{u\in\mathcal{U}}A_u^{(\mu_1)}\right)\rho_A\right).$$

Similarly, we can bound the rest of the terms in (A.7), (A.8), except the first terms. The l_1 norm of the difference of the first terms in (A.7), (A.8) can be written as

$$\begin{aligned}
&\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left\|(\Phi_{(u,v)}-\Phi_{e^{(\mu_1,\mu_2)}(u,v)})\otimes\mathrm{Tr}_{AB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi_{RAB}^{\rho}\right\}\right\|_1\\
&=\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left\|\Phi_{(u,v)}-\Phi_{e^{(\mu_1,\mu_2)}(u,v)}\right\|_1\times\mathrm{Tr}_{RAB}\left\{(\mathrm{id}\otimes A_u^{(\mu_1)}\otimes B_v^{(\mu_2)})\Psi_{RAB}^{\rho}\right\}\\
&=\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\times\sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left\|\Phi_{u,v}-\Phi_{e^{(\mu_1,\mu_2)}(u,v)}\right\|_1\gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v},
\end{aligned}$$

where the first equality is obtained by using the definition of trace norm and the last equality follows from the definition of $A_u^{(\mu_1)}$ and $B_v^{(\mu_2)}$, with $\Omega_{u,v}$ as given in the statement of the proposition. This completes the proof.

A.7 Proof of Proposition II.15

Using the proof of Theorem II.9, one can show that

$$\frac{2}{N_1}\sum_{\mu_1}\mathrm{Tr}\left((I-\sum_{u\in\mathcal{U}}A_u^{(\mu_1)})\rho_A\right)+\frac{2}{N_2}\sum_{\mu_2}\mathrm{Tr}\left((I-\sum_{v\in\mathcal{V}}B_v^{(\mu_2)})\rho_B\right)+2(2-\mathbb{1}_{\{\text{sP-1}\}}-\mathbb{1}_{\{\text{sP-2}\}})\leq\alpha_A+\alpha_B.$$

Recall from Proposition II.14 that S_3 can be simplified as

$$S_3=\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left\|\Phi_{u,v}-\Phi_{e^{(\mu_1,\mu_2)}(u,v)}\right\|_1\gamma_{u^n}^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v},$$

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

$$\|\Phi_{u,v}-\Phi_{e^{(\mu_1,\mu_2)}(u,v)}\|_1\leq 2[\mathbb{1}_{\{(u,v)\notin\mathcal{W}\}}+\mathbb{1}^{(\mu_1,\mu_2)}(u,v)],$$

where $\mathbb{1}^{(\mu_1,\mu_2)}(u,v) \triangleq$

$$\mathbb{1}\Big\{\exists(\tilde{u},\tilde{v},i,j):(u,v)\in\mathcal{B}_1^{(\mu_1)}(i)\times\mathcal{B}_2^{(\mu_2)}(j),(\tilde{u},\tilde{v})\in\mathcal{C}^{(\mu_1,\mu_2)}\bigcap\mathcal{W},$$
$$(\tilde{u},\tilde{v})\in\mathcal{B}_1^{(\mu_1)}(i)\times\mathcal{B}_2^{(\mu_2)}(j),(\tilde{u},\tilde{v})\neq(u,v)\Big\}.$$

Using such indicator functions, S_3 can be bounded from above as $S_3 \leq S_4 + S_5$, where

$$S_4 \triangleq \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{2}{N_1N_2}\sum_{(u,v)}\Omega_{u,v}\sum_{\mu_1,\mu_2}\mathbb{1}_{\{(u,v)\notin\mathcal{W}\}}\gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)},$$

$$S_5 \triangleq \frac{(1-\theta_1)(1-\theta_2)}{(1+\epsilon_1)(1+\epsilon_2)K_1K_2}\sum_{l,k}\sum_{(u,v)}\Omega_{u,v}\frac{2}{N_1N_2}\sum_{\mu_1,\mu_2}\mathbb{1}^{(\mu_1,\mu_2)}(u,v)\mathbb{1}\{U^{(\mu_1)}(l)=u,V^{(\mu_2)}(k)=v\},$$

where we have bounded the indicator random variables in S_5. We provide bounds on the expectation of S_4 and S_5. For that we take the expectation of the indicator functions with respect to random variables which are independent of each other and distributed according to $\{\lambda_u^A\}_{u\in\mathcal{U}}$, and $\{\lambda_v^B\}_{v\in\mathcal{V}}$. First consider the following argument:

$$\begin{aligned}
S_4 &\leq \left|S_4 - 2\sum_{(u,v)\notin\mathcal{W}}\lambda_{u,v}^{AB}\right| + 2\sum_{(u,v)\notin\mathcal{W}}\lambda_{u,v}^{AB}\\
&\overset{(a)}{\leq} \sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\left|\lambda_{u,v}^{AB} - \mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\gamma_u^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v}\right|\\
&\qquad + \mathbb{1}_{\{\text{sP}-1\}}\mathbb{1}_{\{\text{sP}-2\}}(1-\frac{1}{N_1N_2}\sum_{u,v}\sum_{\mu_1,\mu_2}\gamma_v^{(\mu_1)}\zeta_v^{(\mu_2)}\Omega_{u,v})\\
&\qquad + (1-\mathbb{1}_{\{\text{sP}-1\}}\mathbb{1}_{\{\text{sP}-2\}}) + 2\sum_{(u,v)\notin\mathcal{W}}\lambda_{u,v}^{AB}\\
&\overset{(b)}{\leq} S_1 + 2\sum_{(u,v)\notin\mathcal{W}}\lambda_{u,v}^{AB},
\end{aligned}$$

where (a) follows from the two different definitions of variational distance between probability distributions, (b) follows from Lemma II.13. Taking expectation we obtain

$$\mathbb{E}[S_4] \leq (\alpha_A+\alpha_B) + 2\sum_{(u,v)\notin\mathcal{W}}\lambda_{u,v}^{AB}, \tag{A.9}$$

where we used the bound developed earlier on S_1 in the mutual covering lemma. For S_5 we have

$$\begin{aligned}
&\mathbb{E}\Big[\mathbb{1}^{(\mu_1,\mu_2)}(u,v)\mathbb{1}\left\{U^{(\mu_1)}(l)=u, V^{(\mu_2)}(k)=v\right\}\Big] \\
&\overset{(a)}{\leq} \sum_{\substack{(\tilde{u},\tilde{v})\in\mathcal{W}\\(\tilde{u},\tilde{v})\neq(u,v)}} \sum_{i,j}\sum_{(\tilde{l},\tilde{k})} \mathbb{E}\left[\mathbb{1}\left\{(u,v)\in\mathcal{B}_1^{(\mu_1)}(i)\times\mathcal{B}_2^{(\mu_2)}(j)\right\}\right. \\
&\quad \left.\times\mathbb{1}\left\{(\tilde{u},\tilde{v})\in\mathcal{B}_1^{(\mu_1)}(i)\times\mathcal{B}_2^{(\mu_2)}(j)\right\}\mathbb{1}\left\{U^{(\mu_1)}(l)=u, V^{(\mu_2)}(k)=v\right\}\mathbb{1}\left\{U^{(\mu_1)}(\tilde{l})=\tilde{u}, V^{(\mu_2)}(\tilde{k})=\tilde{v}\right\}\right] \\
&\overset{(b)}{\leq} \frac{\lambda_u^A\lambda_v^B}{(1-\theta_1)(1-\theta_2)}\Bigg[\frac{\lambda_{\mathrm{m}}^A\lambda_{\mathrm{m}}^B|\mathcal{W}|K_1K_2}{(1-\theta_1)(1-\theta_2)T_1T_2}+\frac{K_1W_A\lambda_{\mathrm{m}}^A}{(1-\theta_1)T_1} \\
&\qquad\qquad \times\left(1+\frac{\lambda_{\mathrm{m}}^BK_2}{(1-\theta_2)}\right)+\frac{K_2W_B\lambda_{\mathrm{m}}^B}{(1-\theta_2)T_2}\left(1+\frac{\lambda_{\mathrm{m}}^AK_1}{(1-\theta_1)}\right)\Bigg], \qquad \text{(A.10)}
\end{aligned}$$

where (a) follows from the union bound, and (b) follows by noting that there are 5 cases to consider, and by evaluating the expectation of the indicator functions while recalling $W_A = \max_{v\in\mathcal{V}}|\{u : (u,v)\in\mathcal{W}\}|$, and $W_B = \max_{u\in\mathcal{U}}|\{v : (u,v)\in\mathcal{W}\}|$, $\lambda_{\mathrm{m}}^A = \max_u \lambda_u^A$, $\lambda_{\mathrm{m}}^B = \max_v \lambda_v^B$. This implies that

$$\begin{aligned}
\mathbb{E}[S_{\bar{5}}] \leq \frac{2}{(1+\epsilon_1)(1+\epsilon_2)}\Bigg[&\frac{\lambda_{\mathrm{m}}^A\lambda_{\mathrm{m}}^B|\mathcal{W}|K_1K_2}{(1-\theta_1)(1-\theta_2)T_1T_2} \\
&+\frac{K_1W_A\lambda_{\mathrm{m}}^A}{(1-\theta_1)T_1}\left(1+\frac{\lambda_{\mathrm{m}}^BK_2}{(1-\theta_2)}\right)+\frac{K_2W_B\lambda_{\mathrm{m}}^B}{(1-\theta_2)T_2}\times\left(1+\frac{\lambda_{\mathrm{m}}^AK_1}{(1-\theta_1)}\right)\Bigg]\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\Omega_{u,v}\lambda_u^A\lambda_v^B.
\end{aligned}$$

We have the following lemma.

Lemma A.2. *We have*

$$\sum_{u\in\mathcal{U}}\sum_{v\in\mathcal{V}}\Omega_{u,v}\lambda_u^A\lambda_v^B \leq \frac{f_1f_2}{F_1F_2}.$$

Proof. Firstly, note that

$$\sum_{\substack{u\in\mathcal{U}\\v\in\mathcal{V}}}\Omega_{u,v}\lambda_u^A\lambda_v^B = \mathrm{Tr}\left\{\left[\sqrt{\rho_A}^{-1}\Big(\sum_{u\in\mathcal{U}}\lambda_u^A\tilde{\rho}_u^A\Big)\sqrt{\rho_A}^{-1}\otimes\sqrt{\rho_B}^{-1}\Big(\sum_{v\in\mathcal{V}}\lambda_v^B\tilde{\rho}_v^B\Big)\sqrt{\rho_B}^{-1}\right]\rho_{AB}\right\}. \qquad \text{(A.11)}$$

Consider,

$$\sum_{u\in\mathcal{U}}\lambda_u^A\tilde{\rho}_u^A = \hat{\Pi}^A\Pi_{\rho_A}\left(\sum_{u\in\mathcal{U}}\lambda_u^A\Pi_u^A\hat{\rho}_u^A\Pi_u^A\right)\Pi_{\rho_A}\hat{\Pi}^A$$

$$\overset{(a)}{\leq} \hat{\Pi}^A \Pi_{\rho_A} \left(\sum_u \lambda_u^A \hat{\rho}_u^A \right) \Pi_{\rho_A} \hat{\Pi}^A$$
$$\overset{(b)}{\leq} \hat{\Pi}^A \Pi_{\rho_A} \rho_A \Pi_{\rho_A} \hat{\Pi}^A \overset{(c)}{\leq} \frac{1}{F_1} \hat{\Pi}^A \Pi_{\rho_A} \hat{\Pi}^A \overset{(d)}{\leq} \frac{1}{F_1} \Pi_{\rho_A}.$$

where (a) follows from the hypothesis $\Pi_u^A \hat{\rho}_u^A \Pi_u^A \leq \hat{\rho}_u^A$, (b) from the fact that M_A is a sub-POVM, and (c) from the hypothesis $\Pi_{\rho_A} \rho_A \Pi_{\rho_A} \leq \frac{1}{F_1} \Pi_{\rho_A}$, and (d) from the commutativity of $\hat{\Pi}^A$ and Π_{ρ_A}, where the commutativity follows from the fact that $\hat{\Pi}^A$ is a cut-off projector on the subspace determined by Π_{ρ_A}. This implies that

$$\sqrt{\rho_A}^{-1} \left(\sum_{u \in \mathcal{U}} \lambda_{u^n}^A \tilde{\rho}_u^A \right) \sqrt{\rho_A}^{-1} \leq \frac{1}{F_1} \sqrt{\rho_A}^{-1} \Pi_{\rho_A} \sqrt{\rho_A}^{-1} \leq \frac{f_1}{F_1} \Pi_{\rho_A}, \tag{A.12}$$

where the last inequality follows by using the hypothesis $\sqrt{\rho_A}^{-1} \Pi_{\rho_A} \sqrt{\rho_A}^{-1} \leq f_1 \Pi_{\rho_A}$. Using the same arguments for the operators acting on $\mathcal{H}_B$, we have

$$\sqrt{\rho_B}^{-1} \left(\sum_{v \in \mathcal{V}} \lambda_{v^n}^B \tilde{\rho}_v^B \right) \sqrt{\rho_B}^{-1} \leq \frac{1}{F_2} \sqrt{\rho_B}^{-1} \Pi_{\rho_B} \sqrt{\rho_B}^{-1} \leq \frac{f_2}{F_2} \Pi_{\rho_B}. \tag{A.13}$$

Using (A.12) and (A.13) in (A.11), gives

$$\sum_{u,v} \Omega_{u,v} \lambda_u^A \lambda_v^B \leq \frac{f_1 f_2}{F_1 F_2} \operatorname{Tr}\{(\Pi_{\rho_A} \otimes \Pi_{\rho_B}) \rho_{AB}\} \leq \frac{f_1 f_2}{F_1 F_2} \operatorname{Tr}\{\rho_{AB}\} = \frac{f_1 f_2}{F_1 F_2},$$

which is the desired result.

□

A.8 Proof of Proposition II.37

Fix an arbitrary $\epsilon > 0$, and $\eta, \delta \in (0,1)$ sufficiently small. Recalling $S_2(\tilde{\mu}_1, \tilde{\mu}_2)$, we have

$$S_2(\tilde{\mu}_1, \tilde{\mu}_2) \leq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{z^n} \sum_{u^n, v^n} \left| P_{Z|U,V}^n(z^n|u^n, v^n) - P_{Z|U,V}^n \left(z^n | e^{(\bar{\mu}_1, \bar{\mu}_2)}(u^n, v^n) \right) \right|$$
$$\times \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(A_{u^n}^{(\bar{\mu}_1)} \otimes B_{v^n}^{(\bar{\mu}_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1$$
$$\leq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{u^n, v^n} \gamma_{u^n}^{(\bar{\mu}_1)} \zeta_{v^n}^{(\bar{\mu}_2)} \Omega_{u^n, v^n} \sum_{z^n} \left| P_{Z|U,V}^n(z^n|u^n, v^n) - P_{Z|U,V}^n \left(z^n | e^{(\bar{\mu}_1, \bar{\mu}_2)}(u^n, v^n) \right) \right|$$

$$\leq \frac{2}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{u^n,v^n} \left(\mathbb{1}_{\{(u^n,v^n)\notin \mathcal{T}_\delta^{(n)}(U,V)\}} + \mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n) \right) \gamma_{u^n}^{(\bar{\mu}_1)} \zeta_{v^n}^{(\bar{\mu}_2)} \Omega_{u^n,v^n}, \tag{A.14}$$

where Ω_{u^n,v^n} and $\mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)$ are defined as

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

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

Now we can use the bound $S_2 \leq S_{21} + S_{22}$, where

$$S_{21}(\tilde{\mu}_1,\tilde{\mu}_2) \triangleq \frac{2}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{u^n,v^n} \mathbb{1}_{\{(u^n,v^n)\notin \mathcal{T}_\delta^{(n)}(U,V)\}} \gamma_{u^n}^{(\bar{\mu}_1)} \zeta_{v^n}^{(\bar{\mu}_2)} \Omega_{u^n,v^n},$$

$$S_{22}(\tilde{\mu}_1,\tilde{\mu}_2) \triangleq \frac{2}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{u^n,v^n} \mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n) \gamma_{u^n}^{(\bar{\mu}_1)} \zeta_{v^n}^{(\bar{\mu}_2)} \Omega_{u^n,v^n}.$$

We begin by bounding the term corresponding to S_{21}. Consider the following argument.

$$\begin{aligned}
&\frac{1}{\tilde{N}_1 \tilde{N}_2} \sum_{\tilde{\mu}_1,\tilde{\mu}_2} S_{21} \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \\
&\leq \left| \frac{1}{\tilde{N}_1 \tilde{N}_2} \sum_{\tilde{\mu}_1,\tilde{\mu}_2} S_{21} \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} - \sum_{\substack{(u^n,v^n)\notin \mathcal{T}_\delta^{(n)}(UV) \\ u^n \in \mathcal{T}_\delta^{(n)}(U), v^n \in \mathcal{T}_\delta^{(n)}(V)}} 2\lambda_{u^n,v^n}^{AB} \right| + \sum_{(u^n,v^n)\notin \mathcal{T}_\delta^{(n)}(U,V)} 2\lambda_{u^n,v^n}^{AB} \\
&\overset{(a)}{\leq} 2 \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \sum_{v^n \in \mathcal{T}_\delta^{(n)}(V)} \left| \lambda_{u^n,v^n}^{AB} - \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \frac{1}{\bar{N}_1 \bar{N}_2} \sum_{\bar{\mu}_1,\bar{\mu}_2} \gamma_{u^n}^{(\bar{\mu}_1)} \zeta_{v^n}^{(\bar{\mu}_2)} \Omega_{u^n,v^n} \right| + \sum_{(u^n,v^n)\notin \mathcal{T}_\delta^{(n)}(UV)} 2\lambda_{u,v}^{AB} \\
&\overset{(b)}{\leq} 2\, \tilde{S}_1 + 2 \sum_{(u^n,v^n)\notin \mathcal{T}_\delta^{(n)}(UV)} \lambda_{u^n,v^n}^{AB},
\end{aligned}$$

where

$$\tilde{S}_1 \triangleq \left\| (\mathrm{id} \otimes M_A^{\otimes n} \otimes M_B^{\otimes n})(\Psi_{RAB}^{\rho})^{\otimes n} - \frac{1}{\bar{N}_1 \bar{N}_2} \sum_{\bar{\mu}_1,\bar{\mu}_2} (\mathrm{id} \otimes [M_1^{(\bar{\mu}_1)}] \otimes [M_2^{(\bar{\mu}_2)}])(\Psi_{RAB}^{\rho})^{\otimes n} \right\|_1,$$

(a) follows by applying the triangle inequality, and (b) follows from Lemma II.13. Note that in $\tilde{S}_1$, the average over the entire common information sequence $(\bar{\mu}_1,\bar{\mu}_2)$ is inside the norm. Using the

Lemmas II.10 and II.12, and the proof of Theorem II.9, for any $\epsilon \in (0,1)$, and any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, if

$$\begin{aligned} \tilde{R}_1 > I(U;RB)_{\sigma_1}, \quad \tilde{R}_2 > I(V;RA)_{\sigma_2}, \\ \tilde{R}_1 + \frac{1}{n}\log\big(\bar{N}_1\big) > S(U)_{\sigma_3}, \quad \tilde{R}_2 + \frac{1}{n}\log\big(\bar{N}_2\big) > S(V)_{\sigma_3}, \end{aligned} \tag{A.15}$$

then $\mathbb{E}[\tilde{S}_1] \leq \epsilon$. Consequently, we have

$$\frac{1}{\tilde{N}_1\tilde{N}_2}\sum_{\tilde{\mu}_1,\tilde{\mu}_2} \mathbb{E}\left[S_{21}(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\right] \leq 4\epsilon.$$

Now considering the term S_{22}, using a simplification similar to (A.10) we obtain

$$\begin{aligned} &\mathbb{E}\left[\mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n,v^n)\mathbb{1}_{\left\{U^{n,(\mu_1)}(l)=u^n\right\}}\mathbb{1}_{\left\{V^{n,(\mu_2)}(k)=v^n\right\}}\right] \\ &\quad\leq \frac{5\ \lambda^A_{u^n}\lambda^B_{v^n}}{(1-\varepsilon)^2(1-\varepsilon')^2}2^{-n(I(U;V)-3\delta_1)}2^{n(\tilde{R}_1-R_1)}2^{n(\tilde{R}_2-R_2)}. \end{aligned}$$

Substituting this in the expression for S_{22} gives

$$\begin{aligned} \mathbb{E}[S_{22}] &\leq 10\frac{2^{-n(I(U;V)-3\delta_1)}2^{n(\tilde{R}_1-R_1)}2^{n(\tilde{R}_2-R_2)}}{(1+\eta)^2(1-\varepsilon)^2(1-\varepsilon')^2}\sum_{u^n,v^n}\Omega_{u^n,v^n}\lambda^A_{u^n}\lambda^B_{v^n} \\ &\leq 10\frac{2^{-n(I(U;V)-3\delta_1-\delta_{AB})}2^{n(\tilde{R}_1-R_1)}2^{n(\tilde{R}_2-R_2)}}{(1+\eta)^2(1-\varepsilon)^2(1-\varepsilon')^2}, \end{aligned}$$

where the second inequality above uses arguments similar to Lemma A.2. Therefore, if

$$\tilde{R}_1 + \tilde{R}_2 - R_1 - R_2 \leq I(U;V)_{\sigma_3} - 3\delta_1 - \delta_{AB} - \delta, \tag{A.16}$$

then we have $\mathbb{E}[S_{22}] \leq 10\frac{2^{-n\delta}}{(1+\eta)^2(1-\varepsilon)(1-\varepsilon')} < \epsilon$, for all sufficiently large n. Hence

$$\frac{1}{\tilde{N}_1\tilde{N}_2}\sum_{\tilde{\mu}_1,\tilde{\mu}_2}\mathbb{E}(S_2(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-1}\}}(\tilde{\mu}_1,\tilde{\mu}_2)\mathbb{1}_{\{\text{sP-2}\}}(\tilde{\mu}_1,\tilde{\mu}_2)) < 5\epsilon,$$

for all sufficiently large n, if (A.15) and (A.16) are satisfied.

A.9 Proof of Proposition II.39

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

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

Analysis of $\widetilde{S}_2$: We have

$$\begin{aligned}
&\widetilde{S}_2 \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \\
&\leq \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{z^n} P_{Z|U,V}^n(z^n|u_0^n, v_0^n) \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(\sum_{i>0} \Gamma_i^{A,(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\
&\overset{(a)}{=} \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(\sum_{u^n} A_{u^n}^{(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\
&\leq \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{u^n} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(A_{u^n}^{(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\
&\overset{(b)}{\leq} \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \left\| \sqrt{\rho_B^{\otimes n}} \Gamma_0^{B,(\mu_2)} \sqrt{\rho_B^{\otimes n}} \right\|_1 \\
&= \frac{1}{N_2} \sum_{\mu_2} \left\| \sum_{v^n} \lambda_{v^n}^B \hat{\rho}_{v^n}^B - \sum_{v^n} \sqrt{\rho_B^{\otimes n}} B_{v^n}^{(\mu_2)} \sqrt{\rho_B^{\otimes n}} \right\|_1 \\
&\overset{(c)}{\leq} \frac{1}{N_2} \sum_{\mu_2} \left\| \sum_{v^n} \lambda_{v^n}^B \hat{\rho}_{v^n}^B - \frac{(1-\varepsilon')}{(1+\eta)} \frac{1}{2^{n\tilde{R}_2}} \sum_{k=1}^{2^{n\tilde{R}_2}} \hat{\rho}_{V^{n,(\mu_2)}(k)}^B \right\|_1 + \underbrace{\frac{1}{N_2} \sum_{\mu_2} \sum_{v^n} \zeta_{v^n}^{(\mu_2)} \left\| \hat{\rho}_{v^n}^B - \Lambda_{v^n}^B \right\|_1}_{\widetilde{S}_{22}},
\end{aligned} \tag{A.17}$$

where (a) uses the fact that $\sum_{i>0} \Gamma_i^{A,(\mu_1)} = \sum_{u^n} A_{u^n}^{(\mu_1)}$, (b) uses the fact that under the event $\{\mathbb{1}_{\{\text{sP-1}\}} = 1\}$, we have $\sum_{u^n} A_{u^n}^{(\mu_1)} \leq I$, and Lemma II.11. Finally (c) follows from adding and subtracting an appropriate term. Regarding the first term in (A.17) using Lemma II.30 we claim that for for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, the term can be made smaller than ϵ, if $\tilde{R}_2 > I(V;RA)_{\sigma_2}$, where σ_2 is as defined in the statement of the theorem. Note that the requirement we obtain on $\tilde{R}_2$ here was already imposed earlier in

Proposition II.35. And as for the second term, we use the gentle measurement lemma and bound its expected value as

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

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 $n\to\infty$ and $\varepsilon'_B=\varepsilon'_p+2\sqrt{\varepsilon'_p}$ and $\varepsilon''_B=2\varepsilon'_p+2\sqrt{\varepsilon'_p}$ for $\varepsilon'_p\triangleq 1-\min\left\{\mathrm{Tr}\{\Pi_{\rho_B}\hat{\rho}_{v^n}^B\},\mathrm{Tr}\{\Pi_{v^n}\hat{\rho}_{v^n}^B\},1-\varepsilon'\right\}$. Hence $\mathbb{E}\left[\tilde{S}_2\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\right]\leq 2\epsilon$.

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

Analysis of $\widetilde{S}_4$: We have

$$\begin{aligned}
&\widetilde{S}_4\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\\
&\leq\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{z^n}P_{Z|U,V}^n(z^n|u_0^n,v_0^n)\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\otimes\Gamma_0^{B,(\mu_2)}\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1\\
&\leq\frac{\mathbb{1}_{\{\text{sP-1}\}}}{N_1N_2}\sum_{\mu_1,\mu_2}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\otimes I\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1\\
&\qquad+\frac{\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{v^n}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\otimes B_{v^n}^{(\mu_2)}\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1,
\end{aligned}\tag{A.18}$$

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 (A.18) as

$$\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\otimes I\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1=\frac{1}{N_1}\sum_{\mu_1}\left\|\sqrt{\rho_A^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\right)\sqrt{\rho_A^{\otimes n}}\right\|_1.$$

Similarly, the second term in (A.18) simplifies using Lemma II.11 as

$$\frac{\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{v^n}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\otimes B_{v^n}^{(\mu_2)}\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1\leq\frac{1}{N_1}\sum_{\mu_1}\left\|\sqrt{\rho_A^{\otimes n}}\left(\Gamma_0^{A,(\mu_1)}\right)\sqrt{\rho_A^{\otimes n}}\right\|_1.$$

Using these simplifications, we have

$$\widetilde{S}_4 \mathbb{1}_{\{\text{sP-1}\}} \mathbb{1}_{\{\text{sP-2}\}} \leq \frac{2}{N_1} \sum_{\mu_1} \left\| \sqrt{\rho_A^{\otimes n}} \left(\Gamma_0^{A,(\mu_1)} \right) \sqrt{\rho_A^{\otimes n}} \right\|_1 .$$

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

A.10 Proof of Proposition II.40

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 and J_2 are defined as

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

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

Now with the intention of employing Lemma II.30, we express J_1 as

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

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

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

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

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

Now, using Lemma II.30 we get the following bound. For any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}(J_1) < \epsilon$ if

$$\tilde{R}_1 + C_1 > S\left(\sum_{u \in \mathcal{U}} \lambda^A_u \mathcal{T}_u \right) + \sum_{u \in \mathcal{U}} \lambda^A_u S(\mathcal{T}_u) = I(U; RZV)_{\sigma_3}, \tag{A.19}$$

where $\sigma_3 = \sum_{u \in \mathcal{U}} \lambda^A_u \mathcal{T}_u \otimes |u\rangle\langle u|$.

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

$$\begin{aligned} J_2 &\leq \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sum_{u^n, v^n} \sum_{z^n} P^n_{Z|U,V}(z^n|u^n, v^n) \left\| \sqrt{\rho^{\otimes n}_{AB}} \left(\frac{\gamma^{(\mu_1)}_{u^n}}{\lambda^A_{u^n}} \Lambda^A_{u^n} \otimes \Lambda^B_{v^n} - A^{(\mu_1)}_{u^n} \otimes \Lambda^B_{v^n} \right) \sqrt{\rho^{\otimes n}_{AB}} \right\|_1 \\ &= \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sum_{u^n, v^n} \gamma^{(\mu_1)}_{u^n} \left\| \sqrt{\rho^{\otimes n}_{AB}} \left(\left(\frac{1}{\lambda^A_{u^n}} \Lambda^A_{u^n} - \sqrt{\rho^{\otimes n}_{A}}^{-1} \Lambda^A_{u^n} \sqrt{\rho^{\otimes n}_{A}}^{-1} \right) \otimes \Lambda^B_{v^n} \right) \sqrt{\rho^{\otimes n}_{AB}} \right\|_1 , \end{aligned}$$

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

$$\mathbb{E}[J_2] \leq \frac{1}{(1+\eta)} \sum_{u^n \in \mathcal{T}^{(n)}_{\delta}(U)} \lambda^A_{u^n} \sum_{v^n} \left\| \sqrt{\rho^{\otimes n}_{AB}} \left(\left(\frac{1}{\lambda^A_{u^n}} \Lambda^A_{u^n} - \sqrt{\rho^{\otimes n}_{A}}^{-1} \Lambda^A_{u^n} \sqrt{\rho^{\otimes n}_{A}}^{-1} \right) \otimes \Lambda^B_{v^n} \right) \sqrt{\rho^{\otimes n}_{AB}} \right\|_1 ,$$

where we have used the fact that $\mathbb{E}[\gamma^{(\mu_1)}_{u^n}] = \frac{\lambda^A_{u^n}}{(1+\eta)}$. To simplify the above equation, we employ Lemma II.11 from Section 2.2.4.2 that completely discards the effect of Bob's measurement. Since

$\sum_{v^n} \Lambda^B_{v^n} = I$, from Lemma II.11 we have for every $u^n \in \mathcal{T}^{(n)}_\delta(A)$,

$$\begin{aligned}
&\sum_{v^n} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(\left(\frac{1}{\lambda^A_{u^n}} \Lambda^A_{u^n} - \sqrt{\rho_A^{\otimes n}}^{-1} \Lambda^A_{u^n} \sqrt{\rho_A^{\otimes n}}^{-1} \right) \otimes \Lambda^B_{v^n} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\
&= \left\| \sqrt{\rho_A^{\otimes n}} \left(\frac{1}{\lambda^A_{u^n}} \Lambda^A_{u^n} - \sqrt{\rho_A^{\otimes n}}^{-1} \Lambda^A_{u^n} \sqrt{\rho_A^{\otimes n}}^{-1} \right) \sqrt{\rho_A^{\otimes n}} \right\|_1 .
\end{aligned}$$

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

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

where the last inequality is obtained by the repeated usage of the average gentle measurement lemma by setting $\epsilon_{J_2} = \frac{(1-\varepsilon_p)}{(1+\eta)}(2\sqrt{\varepsilon'_A} + 2\sqrt{\varepsilon''_A})$ with $\epsilon_{J_2} \searrow 0$ as $n \to \infty$ and $\varepsilon'_A = \varepsilon_p + 2\sqrt{\varepsilon_p}$ and $\varepsilon''_A = 2\varepsilon_p + 2\sqrt{\varepsilon_p}$ for $\varepsilon_p \triangleq 1 - \min\left\{ \mathrm{Tr}\left\{ \Pi_{\rho_A} \hat{\rho}^A_{u^n} \right\}, \mathrm{Tr}\left\{ \Pi_{u^n} \hat{\rho}^A_{u^n} \right\}, 1-\varepsilon \right\}$. 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 $\tilde{R}_1 + C_1 > I(U;RZV)_{\sigma_3}$.

A.11 Proof of Proposition II.41

We start by adding and subtracting the following terms in Q_2

$$(i) \sum_{u^n, v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda^A_{u^n} \otimes \Lambda^B_{v^n} \right) \sqrt{\rho_{AB}^{\otimes n}} P^n_{Z|U,V}(z^n|u^n, v^n)$$

$$(ii) \sum_{u^n, v^n} \frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda^A_{u^n} \otimes \frac{\zeta^{(\mu_2)}_{v^n}}{\lambda^B_{v^n}} \Lambda^B_{v^n} \right) \sqrt{\rho_{AB}^{\otimes n}} P^n_{Z|U,V}(z^n|u^n, v^n)$$

$$(iii) \sum_{u^n, v^n} \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sqrt{\rho_{AB}^{\otimes n}} \left(A^{(\mu_1)}_{u^n} \otimes \frac{\zeta^{(\mu_2)}_{v^n}}{\lambda^B_{v^n}} \Lambda^B_{v^n} \right) \sqrt{\rho_{AB}^{\otimes n}} P^n_{Z|U,V}(z^n|u^n, v^n).$$

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

$$Q_{21} \triangleq \sum_{z^n} \left\| \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\left(\frac{1}{N_1} \sum_{\mu_1=1}^{N_1} A_{u^n}^{(\mu_1)} \right) \otimes \Lambda_{v^n}^B - \Lambda_{u^n}^A \otimes \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n,v^n) \right\|_1,$$

$$Q_{22} \triangleq \sum_{z^n} \left\| \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda_{u^n}^A \otimes \Lambda_{v^n}^B - \Lambda_{u^n}^A \otimes \left(\frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \Lambda_{v^n}^B \right) \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n,v^n) \right\|_1,$$

$$Q_{23} \triangleq \sum_{z^n} \left\| \sum_{u^n,v^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\Lambda_{u^n}^A \otimes \left(\frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \Lambda_{v^n}^B \right) \right. \right.$$
$$\left. \left. - \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} A_{u^n}^{(\mu_1)} \otimes \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \Lambda_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n,v^n) \right\|_1,$$

$$Q_{24} \triangleq \sum_{z^n} \left\| \sum_{u^n,v^n} \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sqrt{\rho_{AB}^{\otimes n}} \left(A_{u^n}^{(\mu_1)} \otimes \frac{\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B} \Lambda_{v^n}^B - A_{u^n}^{(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|U,V}^n(z^n|u^n,v^n) \right\|_1.$$

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

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

where the equality uses the triangle inequality for block operators. From here on, we use Lemma II.30 to bound Q_{22}. For this, let us define $\mathcal{T}_{v^n}$ as

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

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

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

Using Lemma II.30, for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}(Q_{22}) \leq \epsilon$, if

$$\tilde{R}_2 + C_2 > S\left(\sum_{v\in\mathcal{V}} \lambda_v^B \mathcal{T}_v\right) - \sum_{v\in\mathcal{V}} \lambda_v^B S(\mathcal{T}_v) = I(RZ;V)_{\sigma_3}, \tag{A.20}$$

where σ_3 is defined in the statement of the theorem.

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

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

where the inequality 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 (2.50). Hence, we use the result obtained in bounding $\mathbb{E}[J]$ in the proof of Proposition II.40.

Finally, we consider Q_{24}.

$$
\begin{aligned}
Q_{24} \leq \sum_{u^n,v^n}\sum_{z^n} P^n_{Z|U,V}(z^n|u^n,v^n)\Bigg\| \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho_{AB}^{\otimes n}}\left(A^{(\mu_1)}_{u^n}\otimes\frac{\zeta^{(\mu_2)}_{v^n}}{\lambda^B_{v^n}}\Lambda^B_{v^n}\right)\sqrt{\rho_{AB}^{\otimes n}} \\
- \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho_{AB}^{\otimes n}}\left(A^{(\mu_1)}_{u^n}\otimes\zeta^{(\mu_2)}_{v^n}\left(\sqrt{\rho_B^{\otimes n}}^{-1}\Lambda^B_{v^n}\sqrt{\rho_B^{\otimes n}}^{-1}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}\Bigg\|_1 \\
\leq \frac{1}{N_2}\sum_{\mu_2}\sum_{u^n,v^n}\zeta^{(\mu_2)}_{v^n}\Bigg\|\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}A^{(\mu_1)}_{u^n}\otimes\frac{1}{\lambda^B_{v^n}}\Lambda^B_{v^n}\right)\sqrt{\rho_{AB}^{\otimes n}} \\
- \sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}A^{(\mu_1)}_{u^n}\otimes\left(\sqrt{\rho_B^{\otimes n}}^{-1}\Lambda^B_{v^n}\sqrt{\rho_B^{\otimes n}}^{-1}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}\Bigg\|_1,
\end{aligned}
$$

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

$$
\begin{aligned}
\mathbb{E}_{\mathcal{C}_2}\Big[Q_{24}\mathbb{1}_{\{\text{sP-1}\}}\mathbb{1}_{\{\text{sP-2}\}}\Big] &\leq \mathbb{1}_{\{\text{sP-1}\}}\sum_{\substack{u^n\\ v^n\in\mathcal{T}^{(n)}_\delta(V)}}\frac{\lambda^B_{v^n}}{(1+\eta)}\Bigg\|\sqrt{\rho_{AB}^{\otimes n}}\Big(\frac{1}{N_1}\sum_{\mu_1}A^{(\mu_1)}_{u^n}\otimes\frac{1}{\lambda^B_{v^n}}\Lambda^B_{v^n}\Big)\sqrt{\rho_{AB}^{\otimes n}} \\
&\qquad - \sqrt{\rho_{AB}^{\otimes n}}\Big(\frac{1}{N_1}\sum_{\mu_1}A^{(\mu_1)}_{u^n}\otimes\Big(\sqrt{\rho_B^{\otimes n}}^{-1}\Lambda^B_{v^n}\sqrt{\rho_B^{\otimes n}}^{-1}\Big)\Big)\sqrt{\rho_{AB}^{\otimes n}}\Bigg\|_1 \\
&\overset{(a)}{\leq} \sum_{v^n\in\mathcal{T}^{(n)}_\delta(V)}\frac{\lambda^B_{v^n}}{(1+\eta)}\Bigg\|\sqrt{\rho_B^{\otimes n}}\Big(\frac{1}{\lambda^B_{v^n}}\Lambda^B_{v^n}-\sqrt{\rho_B^{\otimes n}}^{-1}\Lambda^B_{v^n}\sqrt{\rho_B^{\otimes n}}^{-1}\Big)\sqrt{\rho_B^{\otimes n}}\Bigg\|_1 \\
&= \sum_{v^n\in\mathcal{T}^{(n)}_\delta(V)}\frac{\lambda^B_{v^n}}{(1+\eta)}\left\|\hat{\rho}^B_{v^n}-\Lambda^B_{v^n}\right\|_1 \\
&\overset{(b)}{\leq} \frac{(1-\varepsilon')}{(1+\eta)}(2\sqrt{\varepsilon'_B}+2\sqrt{\varepsilon''_B}) = \epsilon_{Q_{24}},
\end{aligned}
\tag{A.21}
$$

where (a) follows by using Lemma II.11 and the fact that under the event $\{\mathbb{1}_{\{\text{sP-1}\}}=1\}$ we have $\sum_{u^n}\frac{1}{N_1}\sum_{\mu_1}A^{(\mu_1)}_{u^n}\leq I$, and (b) 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 $n\to\infty$ and $\varepsilon'_B=\varepsilon_p+2\sqrt{\varepsilon_p}$ and $\varepsilon''_B=2\varepsilon_p+2\sqrt{\varepsilon_p}$, for $\varepsilon_p \triangleq 1-\min\left\{\mathrm{Tr}\{\Pi_{\rho_B}\hat{\rho}^B_{v^n}\},\mathrm{Tr}\{\Pi_{v^n}\hat{\rho}^B_{v^n}\},1-\varepsilon'\right\}$. This completes the proof for Q_{24} and hence for all the terms corresponding to Q_2.

APPENDIX B

Proofs of Chapter III

B.1 Proof of Lemma III.12

Proof. We begin by defining the ensemble $\{\lambda_x, \tilde{\sigma}_x\}_{x\in\mathcal{X}}$ where $\tilde{\sigma}_x = \Pi\Pi_x\sigma_x\Pi_x\Pi$ for all $x \in \mathcal{X}$. Further, let S be defined as

$$S \triangleq \Big\| \sum_{x\in\mathcal{X}} \lambda_x\sigma_x - \frac{1}{M}\sum_{x\in\mathcal{X}}\sum_{m=1}^{M} \frac{\lambda_x}{\mu_x}\sigma_x \mathbb{1}_{\{C_m=x\}} \Big\|_1.$$

By adding an subtracting appropriate terms within the trace norm of S and using the triangle inequality we obtain, $S \leq S_1 + S_2 + S_3$, where

$$S_1 \triangleq \| \sum_{x\in\mathcal{X}} \lambda_x\sigma_x - \sum_{x\in\mathcal{X}} \lambda_x\tilde{\sigma}_x\|_1, \quad S_2 \triangleq \|\frac{1}{M}\sum_{m=1}^{M} \frac{\lambda_{C_m}}{\mu_{C_m}}\tilde{\sigma}_{C_m} - \frac{1}{M}\sum_{m=1}^{M} \frac{\lambda_{C_m}}{\mu_{C_m}}\sigma_{C_m}\|_1, \quad \text{and}$$
$$S_3 \triangleq \| \sum_{x\in\mathcal{X}} \lambda_x\tilde{\sigma}_x - \frac{1}{M}\sum_{x\in\mathcal{X}}\sum_{m=1}^{M} \frac{\lambda_x}{\mu_x}\tilde{\sigma}_x \mathbb{1}_{\{C_m=x\}}\|_1.$$

We begin by bounding the term corresponding to S_1 and S_2 as follows:

$$\begin{aligned} S_1 &\leq \sum_{x\in\mathcal{X}} \lambda_x\|\sigma_x - \Pi\Pi_x\sigma_x\Pi_x\Pi\|_1 \leq \sum_{x\in\mathcal{X}} \lambda_x\|\sigma_x - \Pi\sigma_x\Pi\|_1 + \sum_{x\in\mathcal{X}} \lambda_x\|\Pi\sigma_x\Pi - \Pi\Pi_x\sigma_x\Pi_x\Pi\|_1 \\ &\leq 2\sqrt{\epsilon} + \sum_{x\in\mathcal{X}} \lambda_x\|\Pi\|_\infty\|\sigma_x - \Pi_x\sigma_x\Pi_x\|_1\|\Pi\|_\infty \leq 4\sqrt{\epsilon} = \delta(\epsilon), \end{aligned} \tag{B.1}$$

where the first two inequalities use the triangle inequality, the third uses the gentle measurement lemma (given the assumption (3.6a) from the statement of the Lemma) for the first term, and operator Holder's inequality (Exercise 12.2.1 in *Wilde* (2013a)) for the second term. The last inequality follows again from the gentle measurement given the assumption (3.6b). Similarly, for S_2 we have

$$\mathbb{E}_{\mathbb{C}}[S_2] \leq \mathbb{E}_{\mathbb{C}}\left[\frac{1}{M}\sum_{m=1}^{M}\sum_{x\in\mathcal{X}}\frac{\lambda_x}{\mu_x}\mathbb{1}_{\{C_m=x\}}\|\sigma_x-\tilde{\sigma}_x\|_1\right] = \frac{1}{M}\sum_{m=1}^{M}\sum_{x\in\mathcal{X}}\lambda_x\|\sigma_x-\tilde{\sigma}_x\|_1 \leq 4\sqrt{\epsilon} = \delta(\epsilon), \tag{B.2}$$

where we use the fact that $\mathbb{E}_{\mathbb{C}}[\mathbb{1}_{\{c_m=x\}}] = \mu_x$, and the last inequality uses similar arguments as in (B.1). Finally, we proceed to bound the term corresponding to S_3. Firstly, note that, $\mathbb{E}_{\mathbb{C}}[\frac{1}{M}\sum_m \frac{\lambda_{C_m}}{\mu_{C_m}}\tilde{\sigma}_{C_m}] = \sum_{x\in\mathcal{X}}\lambda_x\tilde{\sigma_x}$. This gives

$$\begin{aligned}
\mathbb{E}_{\mathbb{C}}[S_3] &\leq \mathrm{Tr}\left\{\sqrt{\mathbb{E}_{\mathbb{C}}\left[\left(\frac{1}{M}\sum_m\frac{\lambda_{C_m}}{\mu_{C_m}}\tilde{\sigma}_{C_m}-\mathbb{E}_{\mathbb{C}}\left[\frac{1}{M}\sum_m\frac{\lambda_{C_m}}{\mu_{C_m}}\tilde{\sigma}_{C_m}\right]\right)^2\right]}\right\} \\
&= \mathrm{Tr}\left\{\sqrt{\frac{1}{M^2}\sum_m\mathbb{E}_{\mathbb{C}}\left[\left(\frac{\lambda_{C_m}}{\mu_{C_m}}\tilde{\sigma}_{C_m}\right)^2\right]+\frac{1}{M^2}\sum_{\substack{m,m'\\m\neq m'}}\mathbb{E}_{\mathbb{C}}\left[\frac{\lambda_{C_m}\tilde{\sigma}_{C_m}}{\mu_{C_m}}\frac{\lambda_{C_{m'}}\tilde{\sigma}_{C_{m'}}}{\mu_{C_{m'}}}\right]-\left(\frac{1}{M}\sum_m\mathbb{E}_{\mathbb{C}}\left[\frac{\lambda_{C_m}\tilde{\sigma}_{C_m}}{\mu_{C_m}}\right]\right)^2}\right\} \\
&= \mathrm{Tr}\left\{\sqrt{\frac{1}{M}\mathbb{E}_{\mathbb{C}}\left[\left(\frac{\lambda_{C_1}\tilde{\sigma}_{C_1}}{\mu_{C_1}}\right)^2\right]-\frac{1}{M}\left(\mathbb{E}_{\mathbb{C}}\left[\frac{\lambda_{C_1}\tilde{\sigma}_{C_1}}{\mu_{C_1}}\right]\right)^2}\right\} \leq \mathrm{Tr}\left\{\sqrt{\frac{1}{M}\mathbb{E}_{\mathbb{C}}\left[\left(\frac{\lambda_{C_1}\tilde{\sigma}_{C_1}}{\mu_{C_1}}\right)^2\right]}\right\},
\end{aligned} \tag{B.3}$$

where the first inequality follows from concavity of operator square-root function (Löwner-Heinz theorem, see Theorem 2.6 in *Carlen* (2010)). The last equality uses the fact that codewords of the random code $\mathbb{C}$ are pairwise independent, and the last inequality follows from monotonicity of the operator square-root function (Theorem 2.6 in *Carlen* (2010)).

Moving on, we now bound the operator within the square root of (B.3) as

$$\mathbb{E}_{\mathbb{C}}\left[\left(\frac{\lambda_{C_1}\tilde{\sigma}_{C_1}}{\mu_{C_1}}\right)^2\right] = \sum_{x\in\mathcal{X}}\frac{\lambda_x^2}{\mu_x}\tilde{\sigma}_x^2 \leq \sum_{x\in\mathcal{X}}\kappa\lambda_x\tilde{\sigma}_x^2 = \kappa\sum_{x\in\mathcal{X}}\lambda_x\Pi\left(\Pi_x\sigma_x\Pi_x\right)\Pi\left(\Pi_x\sigma_x\Pi_x\right)\Pi,$$

where we use the assumption $\frac{\lambda_x}{\mu_x} \leq \kappa$ for all $x \in \mathcal{X}$.

Further since, $\Pi \leq I$, we have $(\Pi_x \sigma_x \Pi_x)\,\Pi\,(\Pi_x \sigma_x \Pi_x) \leq (\Pi_x \sigma_x \Pi_x)^2$, which gives

$$\mathbb{E}_{\mathbb{C}}\left[\left(\frac{\lambda_{C_1}\tilde{\sigma}_{C_1}}{\mu_{C_1}}\right)^2\right] \leq \kappa \sum_{x\in\mathcal{X}} \lambda_x \Pi\,(\Pi_x \sigma_x \Pi_x)^2\,\Pi.$$

Moreover, using the assumption 3.6d, i.e., $\Pi_x \sigma_x \Pi_x \leq \frac{1}{d}\Pi_x \leq \frac{1}{d} I$, we get

$$(\Pi_x \sigma_x \Pi_x)^2 = \sqrt{\Pi_x \sigma_x \Pi_x}\,(\Pi_x \sigma_x \Pi_x)\,\sqrt{\Pi_x \sigma_x \Pi_x} \leq \frac{1}{d}\Pi_x \sigma_x \Pi_x, \quad \text{for all } x \in \mathcal{X}.$$

Thus,
$$\mathbb{E}_{\mathbb{C}}\left[\left(\frac{\lambda_{C_1}\tilde{\sigma}_{C_1}}{\mu_{C_1}}\right)^2\right] \leq \frac{\kappa}{d}\Pi\Big(\sum_{x\in\mathcal{X}} \lambda_x \Pi_x \sigma_x \Pi_x\Big)\Pi \leq \frac{\kappa}{d}\Pi\sigma\Pi, \tag{B.4}$$

where the second inequality uses the assumption (3.6e) from the statement of the Lemma. Substituting the simplification obtained in (B.4) into (B.3) and using the monotonicity of square-root operator, we obtain

$$\mathbb{E}_{\mathbb{C}}[S_3] \leq \mathrm{Tr}\left\{\sqrt{\frac{\kappa}{Md}\Pi\sigma\Pi}\right\} \leq \sqrt{\frac{\kappa D}{Md}}, \tag{B.5}$$

where the second inequality uses the assumption (3.6c). Combining the bounds (B.1), (B.2), and (B.5) we get the desired result.

□

B.2 Proof of Lemma III.13

We begin by defining L as

$$L \triangleq \left\| \sum_{w^n} \lambda_{w^n} \theta_{w^n} - \frac{1}{(1+\eta)}\frac{p^n}{p^{k+l}N'} \sum_{\mu=1}^{N'} \sum_{a,m} \sum_{w^n} \lambda_{w^n} \theta_{w^n} \mathbb{1}_{\{W^{n,(\mu)}(a,m)=w^n\}} \right\|_1.$$

Further, let $\theta \triangleq \sum_{w\in\mathcal{W}} \lambda_w \theta_w$ and let Π_θ and $\Pi^{\theta}_{w^n}$ denote the δ-typical projector of θ and conditional typical projector of θ_{w^n}, respectively. Define $\tilde{\lambda}_{w^n} = \frac{\lambda_{w^n}}{1-\varepsilon}$ for $w^n \in \mathcal{T}^{(n)}_{\delta}(W)$, and 0 otherwise, where

$\varepsilon = \sum_{w^n \notin \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n}$. Using the triangle inequality we can bound L as $L \leq L_1 + L_2 + L_3$, where

$$
\begin{aligned}
L_1 &\triangleq \left\| \sum_{w^n} \lambda_{w^n} \theta_{w^n} - \sum_{w^n} \tilde{\lambda}_{w^n} \theta_{w^n} \right\|_1, \\
L_2 &\triangleq \left\| \sum_{w^n} \tilde{\lambda}_{w^n} \theta_{w^n} - \frac{p^n}{p^{k+l} N'} \sum_{\mu=1}^{N'} \sum_{a,m} \sum_{w^n} \tilde{\lambda}_{w^n} \theta_{w^n} \mathbb{1}_{\{W^{n,(\mu)}(a,m) = w^n\}} \right\|_1, \\
L_3 &\triangleq \left\| \frac{p^n}{p^{k+l} N'} \sum_{\mu} \sum_{a,m} \sum_{w^n} \left(\tilde{\lambda}_{w^n} - \frac{\lambda_{w^n}}{(1+\eta)} \right) \theta_{w^n} \mathbb{1}_{\{W^{n,(\mu)}(a,m) = w^n\}} \right\|_1.
\end{aligned}
$$

We begin by bounding the term corresponding to L_1 as

$$
L_1 \leq \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \frac{\varepsilon}{1-\varepsilon} \underbrace{\|\theta_{w^n}\|_1}_{=1} + \sum_{w^n \notin \mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \underbrace{\|\theta_{w^n}\|_1}_{=1} = 2\varepsilon. \tag{B.6}
$$

Now consider the term corresponding to L_2, for which we employ Lemma III.12. Toward this, we consider the following identification: λ_x with $\tilde{\lambda}_{w^n}$, σ_x with θ_{w^n}, $\mathcal{X}$ with $\mathcal{T}_\delta^{(n)}(W)$, $\bar{\mathcal{X}}$ with $\mathbb{F}_p^n$, σ with $\widetilde{\theta} \triangleq \sum_{w^n} \tilde{\lambda}_{w^n} \theta_{w^n}$, Π with Π_θ, Π_x with $\Pi_{w^n}^\theta$, and $\mu_x = \frac{1}{p^n}$ for all $x \in \bar{\mathcal{X}}$. Since the collection of random variables $\{W^{n,(\mu)}(a,m)\}$ are generated using Unionized Coset Codes, we have

$$
\mathbb{P}\left(\mathbb{1}_{\{W^{n,(\mu)}(a,m) = w^n\}} = 1 \right) = \frac{1}{p^n}, \quad \text{for all} \quad w^n \in \mathbb{F}_p^n.
$$

Note that $\frac{\tilde{\lambda}_{w^n}}{1/p^n} \leq 2^{-n(S(W)_{\sigma_\theta} - \log p - \delta_w)}$ for all $w^n \in \mathbb{F}_p^n$, where $\delta_w(\delta) \searrow 0$ as $\delta \searrow 0$, and σ_θ is defined in the statement of the lemma. With these, we check the hypotheses of Lemma III.12. Firstly, using the pinching arguments described in (*Wilde*, 2013a, Property 15.2.7), we have $\mathrm{Tr}\{\Pi_\theta \theta_{w^n}\} \geq 1 - \epsilon$ for all $\epsilon \in (0,1), \delta > 0$ and sufficiently large n, satisfying hypothesis (3.6a). Secondly, (3.6b) and (3.6e) are satisfied from the construction of $\Pi_{w^n}^\theta$. Next, we consider the hypothesis (3.6c). We have

$$
\left\| \Pi^\theta \sqrt{\tilde{\theta}} \right\|_1 = \mathrm{Tr}\left\{ \sqrt{\Pi^\theta \tilde{\theta} \Pi^\theta} \right\} \leq \frac{1}{\sqrt{(1-\varepsilon)}} \mathrm{Tr}\left\{ \sqrt{\Pi^\theta \theta^{\otimes n} \Pi^\theta} \right\} \leq 2^{\frac{n}{2}(S(R)_{\sigma_\theta} + \delta'_w)},
$$

where the first inequality above follows from the fact that $\sum_{w^n} \tilde{\lambda}_{w^n} \theta_{w^n} \leq \frac{1}{(1-\varepsilon)} \sum_{w^n} \lambda_{w^n} \theta_{w^n} = \frac{\theta^{\otimes n}}{(1-\varepsilon)}$ and using the operator monotonicty of the square-root function (Theorem 2.6 in *Carlen* (2010)). The second inequality follows from the property of the typical projector for some δ'_w such that $\delta'_w \searrow 0$ as $\delta \searrow 0$. This gives $D = 2^{n(S(R)_{\sigma_\theta} + \delta'_w)}$. Finally, the hypotheses (3.6d) is satisfied from

the property of conditional typical projectors for $d = 2^{n(S(R|W)_{\sigma_\theta} - \delta''_w)}$, where $\delta''_w \searrow 0$ as $\delta \searrow 0$ (see (*Wilde*, 2013a, Property 15.2.6)). Next we check the pairwise independence of $W^{n,(\mu)}(a,m)$ and $W^{n,(\mu)}(\tilde{a},\tilde{m})$. Since these are constructed using randomly and uniformly generated G and $h^{(\mu)}$, we have $\{W^{n,(\mu)}(a,m)\}_{a\in\mathbb{F}_p^k, m\in\mathbb{F}_p^l, \mu\in[1:N']}$ to be pairwise independent for each (see *Pradhan et al.* (2021) for details). Therefore, employing Lemma III.12 we get

$$\begin{aligned}\mathbb{E}[L_2] &\leq \sqrt{\frac{2^{n(S(R)_{\sigma_\theta}+\delta'_w)}2^{-n(S(W)_{\sigma_\theta}-\log p-\delta_w)}}{N'p^{k+l}2^{n(S(R|W)_{\sigma_\theta}-\delta''_w)}}} + 8\sqrt{\epsilon} \\ &\leq \exp_2\left[-\frac{n}{2}\left(\frac{k+l}{n}\log p + \frac{1}{n}\log N' - I(R;W)_{\sigma_\theta} - \log p + S(W)_{\sigma_\theta} - \delta_w - \delta'_w - \delta''_w\right)\right] + 8\sqrt{\epsilon},\end{aligned} \tag{B.7}$$

where $\exp_2(x) \triangleq 2^x$. As for L_3, taking expectation and using $\mathbb{E}[\mathbb{1}_{\{W^{n,(\mu)}(a,m)=w^n\}}] = \frac{1}{p^n}$ gives

$$\mathbb{E}[L_3] \leq \frac{\eta+\varepsilon}{(1+\eta)} + \frac{\varepsilon}{(1+\eta)} = \frac{\eta+2\varepsilon}{1+\eta}. \tag{B.8}$$

Combining the bounds from (B.6), (B.7) and (B.8) gives the desired result.

B.3 Proof of Proposition III.14

Applying the triangle inequality on $\widetilde{S}_1$ gives $\widetilde{S}_1 \leq \widetilde{S}_{11} + \widetilde{S}_{12}$, where

$$\widetilde{S}_{11} \triangleq \frac{1}{N}\sum_{\mu}\left\|\sum_{w^n}\lambda_{w^n}\hat{\rho}_{w^n} - \sum_{w^n}\alpha_{w^n}\gamma^{(\mu)}_{w^n}\hat{\rho}_{w^n}\right\|_1, \quad \widetilde{S}_{12} \triangleq \frac{1}{N}\sum_{\mu}\sum_{w^n}\alpha_{w^n}\gamma^{(\mu)}_{w^n}\left\|\hat{\rho}_{w^n} - \tilde{\rho}_{w^n}\right\|_1.$$

For the first term $\widetilde{S}_{11}$, we use Lemma II.30, and identify θ_{w^n} with $\hat{\rho}_{w^n}$ and $N' = 1$. Using this lemma, we obtain the following: For any $\epsilon > 0$, and any $\eta, \delta \in (0,1)$ sufficiently small and any n sufficiently large, $\mathbb{E}[\widetilde{S}_{11}] \leq \epsilon$, if the $\frac{k+l}{n}\log p > I(W;R)_\sigma - S(W)_\sigma + \log p$, where σ is defined in the statement of the theorem. As for the second term $\widetilde{S}_{12}$, we use the gentle measurement lemma and bound its expected value as

$$\mathbb{E}\left[\frac{1}{N}\sum_{\mu}\sum_{w^n}\alpha_{w^n}\gamma^{(\mu)}_{w^n}\left\|\hat{\rho}_{w^n} - \tilde{\rho}_{w^n}\right\|_1\right] \leq \sum_{w^n\in\mathcal{T}^{(n)}_\delta(W)}\frac{\lambda_{w^n}}{(1+\eta)}\left\|\hat{\rho}_{w^n} - \tilde{\rho}_{w^n}\right\|_1 + \sum_{w^n\notin\mathcal{T}^{(n)}_\delta(W)}\frac{\lambda_{w^n}}{(1+\eta)} \leq \epsilon_{\widetilde{S}_{12}},$$

where the inequality is based on the repeated usage of the average gentle measurement lemma by setting $\epsilon_{\tilde{S}_{12}} = \frac{(1-\varepsilon)}{(1+\eta)}(2\sqrt{\varepsilon'}+2\sqrt{\varepsilon''})$ with $\epsilon_{\tilde{S}_{12}} \searrow 0$ as $n \to \infty$ and $\varepsilon' = \varepsilon'_p + 2\sqrt{\varepsilon'_p}$ and $\varepsilon'' = 2\varepsilon'_p + 2\sqrt{\varepsilon'_p}$ for $\varepsilon'_p \triangleq 1 - \min\{\mathrm{Tr}\{\Pi_\rho \hat{\rho}_{w^n}\}, \mathrm{Tr}\{\Pi_{w^n}\hat{\rho}_{w^n}\}, 1-\varepsilon\}$ (see (35) in *Wilde et al.* (2012) for more details).

B.4 Proof of Lemma III.15

We begin by using the Hólder's inequality *Wilde* (2013a); *Carlen* (2010) for operator norm, i.e., ($\|AB\|_1 \leq \|A\|_\infty \|B\|_1$), and defining $\hat{\Lambda}_{w^n} = \sqrt{\rho^{\otimes n}}^{-1} \tilde{\rho}_{w^n} \sqrt{\rho^{\otimes n}}^{-1}$. This gives

$$\begin{aligned}
\sum_{w^n} \gamma_{w^n}^{(\mu)} \left\| \sqrt{\rho^{\otimes n}} \left(\bar{A}_{w^n}^{(\mu)} - A_{w^n}^{(\mu)} \right) \sqrt{\rho^{\otimes n}} \right\|_1 &= \sum_{w^n} \alpha_{w^n} \gamma_{w^n}^{(\mu)} \left\| \Pi_\rho \sqrt{\rho^{\otimes n}} \hat{\Lambda}_{w^n} \sqrt{\rho^{\otimes n}} \Pi_\rho - \Pi_\rho \sqrt{\rho^{\otimes n}} \Pi^\mu \hat{\Lambda}_{w^n} \Pi^\mu \sqrt{\rho^{\otimes n}} \Pi_\rho \right\|_1 \\
&\leq \sum_{w^n} \alpha_{w^n} \gamma_{w^n}^{(\mu)} \left\| \Pi_\rho \sqrt{\rho^{\otimes n}} \right\|_\infty^2 \left\| \hat{\Lambda}_{w^n} - \Pi^\mu \hat{\Lambda}_{w^n} \Pi^\mu \right\|_1 \\
&\leq 2^{-n(S(\rho)-\delta_\rho)} \sum_{w^n} \alpha_{w^n} \gamma_{w^n}^{(\mu)} 2 \sqrt{\mathrm{Tr}\left\{ (\Pi_{\rho_A} - \Pi^\mu) \hat{\Lambda}_{w^n} \right\} \mathrm{Tr}\left\{ \hat{\Lambda}_{w^n} \right\}},
\end{aligned}$$

where the equality follows from the fact that Π_ρ and Π^μ commute, the first inequality follows from the Hólder's inequality, and the second inequality uses the following bounds

$$\begin{aligned}
\left\| \hat{\Lambda}_{w^n} - \Pi^\mu \hat{\Lambda}_{w^n} \Pi^\mu \right\|_1 &\leq \left\| \hat{\Lambda}_{w^n} - \Pi^\mu \hat{\Lambda}_{w^n} \right\|_1 + \left\| \Pi^\mu \hat{\Lambda}_{w^n} - \Pi^\mu \hat{\Lambda}_{w^n} \Pi^\mu \right\|_1 \\
&\leq \sqrt{\mathrm{Tr}\left\{ (\Pi_\rho - \Pi^\mu)^2 \hat{\Lambda}_{w^n} \right\} \mathrm{Tr}\left\{ \hat{\Lambda}_{w^n} \right\}} + \sqrt{\mathrm{Tr}\left\{ \Pi^\mu \hat{\Lambda}_{w^n} \right\} \mathrm{Tr}\left\{ \hat{\Lambda}_{w^n} (\Pi_\rho - \Pi^\mu)^2 \right\}} \\
&\leq 2 \sqrt{\mathrm{Tr}\left\{ (\Pi_\rho - \Pi^\mu) \hat{\Lambda}_{w^n} \right\} \mathrm{Tr}\left\{ \hat{\Lambda}_{w^n} \right\}},
\end{aligned}$$

where the second inequality uses Cauchy-Schwarz inequality along with the polar decomposition (see the usage in (*Wilde*, 2013a, Lemma 9.4.2)) and the last inequality uses the arguments: (i) Π^μ is a projector onto a subspace of Π_ρ and (ii) $\mathrm{Tr}\left\{ \Pi^\mu \hat{\Lambda}_{w^n} \right\} \leq \mathrm{Tr}\left\{ \hat{\Lambda}_{w^n} \right\}$. Further, using the fact that for $w^n \in \mathcal{T}_\delta^{(n)}(W)$,

$$\mathrm{Tr}\{\hat{\Lambda}_{w^n}\} = \|\Pi_\rho \hat{\Lambda}_{w^n} \Pi_\rho\|_1 \leq \|\Pi_\rho \sqrt{\rho^{\otimes n}}^{-1}\|_\infty \underbrace{\|\tilde{\rho}_{w^n}\|_1}_{\leq 1} \|\Pi_\rho \sqrt{\rho^{\otimes n}}^{-1}\|_\infty \leq \|\Pi_\rho \sqrt{\rho^{\otimes n}}^{-1}\|_\infty^2 \leq 2^{n(S(\rho)+\delta_\rho)},$$

it follows that

$$\begin{aligned}
\sum_{w^n} \gamma_{w^n}^{(\mu)} \left\| \sqrt{\rho^{\otimes n}} \left(\bar{A}_{w^n}^{(\mu)} - A_{w^n}^{(\mu)} \right) \sqrt{\rho^{\otimes n}} \right\|_1 &\leq 2 \cdot 2^{-\frac{n}{2}(S(\rho)-4\delta_\rho)} \sum_{w^n} \alpha_{w^n} \gamma_{w^n}^{(\mu)} \sqrt{\operatorname{Tr}\left\{ (\Pi_\rho - \Pi^\mu) \hat{\Lambda}_{w^n} \right\}} \\
&\leq 2 \cdot 2^{3n\delta_\rho} \Delta^{(\mu)} \sqrt{\sum_{w^n} \frac{\alpha_{w^n} \gamma_{w^n}^{(\mu)}}{\Delta^{(\mu)}} \operatorname{Tr}\{(\Pi_\rho - \Pi^\mu) \tilde{\rho}_{w^n}\}} \\
&= 2 \cdot 2^{3n\delta_\rho} \left(\Delta^{(\mu)} - \mathbb{E}[\Delta^{(\mu)}] + \mathbb{E}[\Delta^{(\mu)}] \right) \sqrt{\operatorname{Tr}\left\{ (\Pi_\rho - \Pi^\mu) \sum_{w^n} \frac{\alpha_{w^n} \gamma_{w^n}^{(\mu)}}{\Delta^{(\mu)}} \tilde{\rho}_{w^n} \right\}} \\
&\leq 2 \cdot 2^{3n\delta_\rho} \mathbb{E}[\Delta^{(\mu)}] \sqrt{\operatorname{Tr}\left\{ (\Pi_\rho - \Pi^\mu) \sum_{w^n} \frac{\alpha_{w^n} \gamma_{w^n}^{(\mu)}}{\Delta^{(\mu)}} \tilde{\rho}_{w^n} \right\}} + 2 \cdot 2^{3n\delta_\rho} \underbrace{\left| \Delta^{(\mu)} - \mathbb{E}[\Delta^{(\mu)}] \right|}_{H_0} \\
&\leq 2 \cdot 2^{3n\delta_\rho} \left(H_0 + \frac{\sqrt{(1-\varepsilon)}}{(1+\eta)} \sqrt{H_1 + H_2 + H_3} \right),
\end{aligned}$$

where the second inequality above follows by defining $\Delta^{(\mu)} = \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \alpha_{w^n} \gamma_{w^n}^{(\mu)}$ and using the concavity of the square-root function, the third inequality follows by using the fact that

$$\sum_{w^n} \frac{\alpha_{w^n} \gamma_{w^n}^{(\mu)}}{\Delta^{(\mu)}} \operatorname{Tr}\{(\Pi_\rho - \Pi^\mu) \tilde{\rho}_{w^n}\} \leq \sum_{w^n} \frac{\alpha_{w^n} \gamma_{w^n}^{(\mu)}}{\Delta^{(\mu)}} \operatorname{Tr}\{\tilde{\rho}_{w^n}\} \leq 1, \tag{B.9}$$

and defining H_0 as above. and the last one follows by first using $\mathbb{E}[\Delta^{(\mu)}] = \frac{(1-\varepsilon)}{(1+\eta)}$ and then defining H_1, H_2 and H_3 as in the statement of the lemma and using the inequality $\operatorname{Tr}\{\Lambda(\omega - \sigma)\} \leq \|\Lambda(\omega - \sigma)\|_1 \leq \|\Lambda\|_\infty \|\omega - \sigma\|_1$. This completes the proof.

B.5 Proof of Proposition III.16

To provide a bound for $\widetilde{S}_2$, we individually bound the terms corresponding to H_0 and $\widetilde{H}$ in an expected sense. Let us first consider $\widetilde{H}$. To provide a bound for $\tilde{H}$ we use Lemma III.12 with the following identification: λ_x with $\frac{\lambda_{w^n}}{(1-\varepsilon)}$, σ_x with $\hat{\rho}_{w^n}$, $\mathcal{X}$ with $\mathcal{T}_\delta^{(n)}(W)$, $\bar{\mathcal{X}}$ with $\mathbb{F}_p^n$, Π with Π_ρ, Π_x with Π_{w^n}, and μ_x with $\frac{1}{p^n}$.

Firstly, we have $\frac{\lambda_{w^n}}{1/p^n} \leq 2^{-n(S(W)_\sigma - \log p - \delta_w)}$ for all $w^n \in \mathbb{F}_p^n$, where $\delta_w(\delta) \searrow 0$ as $\delta \searrow 0$, which gives $\kappa = 2^{-n(S(W)_\sigma - \log p - \delta_w)}$. With these, we check the hypotheses of Lemma III.12. As for the first hypothesis (3.6a), using the pinching arguments described in (*Wilde*, 2013a, Property 15.2.7), we have $\operatorname{Tr}\{\Pi_\rho \hat{\rho}_{w^n}\} \geq 1 - \epsilon$ for all $\epsilon \in (0,1), \delta > 0$ and sufficiently large n. Then the hypotheses (3.6b)

and (3.6e) are satisfied from the construction of Π_{w^n}. Next, consider the hypothesis (3.6c). We have

$$\left\| \Pi_\rho \sqrt{\left(\sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \frac{\lambda_{w^n}}{(1-\varepsilon)} \hat{\rho}_{w^n} \right)} \right\|_1 \leq \frac{1}{\sqrt{1-\varepsilon}} \operatorname{Tr}\left\{ \sqrt{\Pi_\rho \rho^{\otimes n} \Pi_\rho} \right\} \leq 2^{\frac{n}{2}(S(R)_\sigma + \delta'_\rho)},$$

where the first inequality above follows from using $\sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \frac{\lambda_{w^n}}{(1-\varepsilon)} \hat{\rho}_{w^n} \leq \frac{1}{(1-\varepsilon)} \rho^{\otimes n}$ and the operator monotonicty of the square-root function. The second inequality follows from the property of the typical projector for some δ'_ρ such that $\delta'_w \searrow 0$ as $\delta \searrow 0$. This gives $D = 2^{n(S(R)_\sigma + \delta'_\rho)}$, where σ is as defined in the statement of the theorem. Finally, the hypotheses (3.6d) is satisfied from the property of conditional typical projectors for $d = 2^{n(S(R|W)_\sigma - \delta''_w)}$, where $\delta''_w \searrow 0$ as $\delta \searrow 0$. Next we check the pairwise independence of $W^{n,(\mu)}(a,m)$ and $W^{n,(\mu)}(\tilde{a},\tilde{m})$. Since these are constructed using randomly and uniformly generated G and $h^{(\mu)}$, we have $\{W^{n,(\mu)}(a,m)\}_{a \in \mathbb{F}_p^k, m \in \mathbb{F}_p^l, \mu \in [1,N]}$ to be pairwise independent (see *Pradhan et al.* (2021) for details). Therefore, employing inequality (3.8) of Lemma III.12, we get

$$\mathbb{E}[\tilde{H}] \leq \sqrt{\frac{2^{n(S(R)_\sigma + \delta'_\rho)} 2^{-n(S(W)_\sigma - \log p - \delta_w)}}{N 2^{nS} 2^{n(S(R|W)_\sigma - \delta''_w)}}} \leq 2^{-\frac{n}{2}\left(\frac{k+l}{n} \log p + \frac{1}{n} \log N - I(R;W)_\sigma - \log p + S(W)_\sigma - \delta_w - \delta'_\rho - \delta''_w \right)}.$$

Next, consider H_0 and perform the following simplification

$$\begin{aligned}
\mathbb{E}[H_0] &= \frac{(1-\varepsilon)}{(1+\eta)} \mathbb{E} \left| \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \frac{\lambda_{w^n}}{(1-\varepsilon)} - \frac{p^n}{p^{k+l}} \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \sum_{a,i} \frac{\lambda_{w^n}}{(1-\varepsilon)} \mathbb{1}_{\{W^{n,(\mu)}(a,i) = w^n\}} \right| \\
&= \frac{(1-\varepsilon)}{(1+\eta)} \mathbb{E} \left\| \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \frac{\lambda_{w^n}}{(1-\varepsilon)} \omega_0^{\otimes n} - \frac{p^n}{p^{k+l}} \sum_{w^n \in \mathcal{T}_\delta^{(n)}(W)} \sum_{a,i} \frac{\lambda_{w^n}}{(1-\varepsilon)} \mathbb{1}_{\{W^{n,(\mu)}(a,i) = w^n\}} \omega_0^{\otimes n} \right\|_1,
\end{aligned} \tag{B.10}$$

where $\omega_0 \in \mathcal{D}(\mathcal{H})$ is any state independent of W. We again apply Lemma III.12 to the above term with the following identification: λ_x with $\frac{\lambda_{w^n}}{(1-\varepsilon)}$, σ_x with $\omega_0^{\otimes n}$, $\mathcal{X}$ with $\mathcal{T}_\delta^{(n)}(W)$, $\bar{\mathcal{X}}$ with $\mathbb{F}_p^n$, Π and Π_x with Identity operator I, and μ_x with $\frac{1}{p^n}$. With this identification, κ remains as above, $\kappa = 2^{-n(S(W)_\sigma - \log p - \delta_w)}$ and $D = d = 1$. Hence, using in inequality (3.8) of Lemma III.12, we

obtain

$$\mathbb{E}[H_0] \leq 2^{-\frac{n}{2}\left(\frac{k+l}{n}\log p - \log p + S(W)_\sigma - \delta_w\right)}.$$

This completes the proof.

B.6 Proof of Proposition III.18

We begin using the definition of $A^{(\mu)}_{w^n}$ and applying triangle inequality to S_2 to obtain

$$\begin{aligned}
S_2 &\leq \frac{1}{(1+\eta)}\frac{1}{N}\sum_{\mu}\sum_{a,i>0}\sum_{w^n,z^n}\frac{\lambda_{w^n}p^n}{p^{k+l}}\mathbb{1}_{\{aG+h^{(\mu)}(i)=w^n\}}\left\|\sqrt{\rho^{\otimes n}}\Pi^\mu\sqrt{\rho^{\otimes n}}^{-1}\tilde{\rho}_{w^n}\sqrt{\rho^{\otimes n}}^{-1}\Pi^\mu\sqrt{\rho^{\otimes n}}\right\|_1 \\
&\qquad\qquad \times\left|P^n_{Z|W}(z^n|w^n) - P^n_{Z|W}\left(z^n|F^{(\mu)}(i)\right)\right| \\
&\leq \frac{2^{2n\delta_\rho}}{(1+\eta)}\frac{1}{N}\sum_{\mu}\sum_{a,i>0}\sum_{w^n,z^n}\frac{\lambda_{w^n}p^n}{p^{k+l}}\mathbb{1}_{\{aG+h^{(\mu)}(i)=w^n\}}\left|P^n_{Z|W}(z^n|w^n) - P^n_{Z|W}\left(z^n|F^{(\mu)}(i)\right)\right| \\
&\leq \frac{2^{2n\delta_\rho}}{(1+\eta)}\frac{1}{N}\sum_{\mu}\sum_{a,i>0}\sum_{w^n}2\frac{\lambda_{w^n}p^n}{p^{k+l}}\mathbb{1}_{\{aG+h^{(\mu)}(i)=w^n\}}\mathbb{1}^{(\mu)}(w^n,i), \qquad \text{(B.11)}
\end{aligned}$$

where the second inequality above uses the following arguments

$$\begin{aligned}
\left\|\sqrt{\rho^{\otimes n}}\Pi^\mu\sqrt{\rho^{\otimes n}}^{-1}\tilde{\rho}_{w^n}\sqrt{\rho^{\otimes n}}^{-1}\Pi^\mu\sqrt{\rho^{\otimes n}}\right\|_1 &= \left\|\sqrt{\rho^{\otimes n}}\Pi_\rho\Pi^\mu\sqrt{\rho^{\otimes n}}^{-1}\Pi_\rho\tilde{\rho}_{w^n}\Pi_\rho\sqrt{\rho^{\otimes n}}^{-1}\Pi^\mu\Pi_\rho\sqrt{\rho^{\otimes n}}\right\|_1 \\
&\leq \left\|\sqrt{\rho^{\otimes n}}\Pi_\rho\right\|_\infty\left\|\Pi^\mu\sqrt{\rho^{\otimes n}}^{-1}\Pi_\rho\tilde{\rho}_{w^n}\Pi_\rho\sqrt{\rho^{\otimes n}}^{-1}\Pi^\mu\right\|_1\left\|\sqrt{\rho^{\otimes n}}\Pi_\rho\right\|_\infty \\
&\leq 2^{-n(S(\rho)-\delta_\rho)}\left\|\Pi^\mu\right\|^2_\infty\left\|\sqrt{\rho^{\otimes n}}^{-1}\Pi_\rho\tilde{\rho}_{w^n}\Pi_\rho\sqrt{\rho^{\otimes n}}^{-1}\right\|_1 \leq 2^{2n\delta_\rho}\left\|\tilde{\rho}_{w^n}\right\|_1 \leq 2^{2n\delta_\rho}, \qquad \text{(B.12)}
\end{aligned}$$

where the above inequalities follow from the Hólder's inequality. Finally, the last inequality in (B.11) follows by defining $\mathbb{1}^{(\mu)}(w^n,i)$ as

$$\mathbb{1}^{(\mu)}(w^n,i) \triangleq \mathbb{1}\left\{\exists(\tilde{w}^n,\tilde{a}^n) : \tilde{w}^n = \tilde{a}^nG + h^{(\mu)}(i), \tilde{w}^n \in \mathcal{T}^{(n)}_\delta(W), \tilde{w}^n \neq w^n\right\}.$$

Observe that

$$\mathbb{E}[\mathbb{1}^{(\mu)}(w^n,i)\mathbb{1}_{\{aG+h^{(\mu)}(i)=w^n\}}] \le \sum_{\tilde{a}\in\mathbb{F}_p^k} \sum_{\substack{\tilde{w}\in\mathcal{T}_\delta^{(n)}(W)\\ \tilde{w}\neq w^n}} \frac{1}{p^n p^n},$$

which follows from the pairwise independence of the codewords. Using this, we obtain

$$\mathbb{E}[S_2] \le \frac{2}{(1+\eta)} \frac{2^{2n\delta_\rho} 2^{-nR} p^{k+l}}{p^n} \sum_{\tilde{w}^n\in\mathcal{T}_\delta^{(n)}(W)} \sum_{w^n\in\mathcal{T}_\delta^{(n)}(W)} \lambda_{w^n} \le 2\cdot 2^{n(\frac{k+l}{n}\log p - R - \log p + S(W)_\sigma + \delta_{S_2})},$$

where $\delta_{S_2} \searrow 0$ as $\delta \searrow 0$, and σ is as defined in the statement of the theorem. This completes the proof.

B.7 Proof of Proposition III.30

Recalling S_2, we have $S_2 \le S_{21} + S_{22}$, where

$$S_{21} \triangleq \frac{2}{N_1 N_2} \sum_{\bar{\mu}_1,\bar{\mu}_2} \sum_{u^n,v^n} \alpha_{u^n}\beta_{v^n}\gamma_{u^n}^{(\bar{\mu}_1)}\zeta_{v^n}^{(\bar{\mu}_2)}\Omega_{u^n,v^n}\mathbb{1}_{\{(u^n,v^n)\notin\mathcal{T}_\delta^{(n)}(U,V)\}},$$
$$S_{22} \triangleq \frac{2}{\bar{N}_1 \bar{N}_2} \sum_{\bar{\mu}_1,\bar{\mu}_2} \sum_{u^n,v^n} \alpha_{u^n}\beta_{v^n}\gamma_{u^n}^{(\bar{\mu}_1)}\zeta_{v^n}^{(\bar{\mu}_2)}\Omega_{u^n,v^n}\mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n+v^n,i,j),$$

where Ω_{u^n,v^n} and $\mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(w^n,i,j)$ are defined as

$$\Omega_{u^n,v^n} \triangleq \mathrm{Tr}\left\{\left[\left(\Pi_A^{\bar{\mu}_1}\otimes\Pi_B^{\bar{\mu}_2}\right)\sqrt{\rho_A^{\otimes n}\otimes\rho_B^{\otimes n}}^{-1}(\tilde{\rho}_{u^n}^A\otimes\tilde{\rho}_{v^n}^B)\sqrt{\rho_A^{\otimes n}\otimes\rho_B^{\otimes n}}^{-1}\left(\Pi_A^{\bar{\mu}_1}\otimes\Pi_B^{\bar{\mu}_2}\right)\right]\rho_{AB}^{\otimes n}\right\},$$
$$\mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(w^n,i,j) \triangleq \mathbb{1}\left\{\exists(\tilde{w}^n,\tilde{a}^n) : \tilde{w}^n = \tilde{a}^nG + h_1^{(\bar{\mu}_1)}(i) + h_2^{(\bar{\mu}_2)}(j), \tilde{w}^n\in\mathcal{T}_{\hat{\delta}}^{(n)}(U+V), \tilde{w}^n\neq w^n\right\}.$$

We begin by bounding the term corresponding to S_{21}. Consider the following argument.

$$S_{21} \le \left|\frac{2}{\bar{N}_1\bar{N}_2}\sum_{\bar{\mu}_1,\bar{\mu}_2}\sum_{u^n,v^n}\alpha_{u^n}\beta_{v^n}\gamma_{u^n}^{(\bar{\mu}_1)}\zeta_{v^n}^{(\bar{\mu}_2)}\Omega_{u^n,v^n}\mathbb{1}_{\{(u^n,v^n)\notin\mathcal{T}_\delta^{(n)}(U,V)\}} - \sum_{\substack{(u^n,v^n)\notin\mathcal{T}_\delta^{(n)}(UV)\\ u^n\in\mathcal{T}_\delta^{(n)}(U),v^n\in\mathcal{T}_\delta^{(n)}(V)}} 2\lambda_{u^n,v^n}^{AB}\right|$$
$$+ \sum_{(u^n,v^n)\notin\mathcal{T}_\delta^{(n)}(U,V)} 2\lambda_{u^n,v^n}^{AB}$$

$$\overset{(a)}{\leq} 2 \sum_{u^n \in \mathcal{U}^n} \sum_{v^n \in \mathcal{V}^n} \left| \lambda^{AB}_{u^n,v^n} - \frac{1}{\bar{N}_1 \bar{N}_2} \sum_{\bar{\mu}_1, \bar{\mu}_2} \alpha_{u^n} \beta_{v^n} \gamma^{(\bar{\mu}_1)}_{u^n} \zeta^{(\bar{\mu}_2)}_{v^n} \Omega_{u^n,v^n} \right| + \sum_{(u^n,v^n) \notin \mathcal{T}^{(n)}_{\delta}(UV)} 2\lambda^{AB}_{u,v}$$

$$\overset{(b)}{\leq} 2\tilde{S}_1 + 2 \sum_{(u^n,v^n) \notin \mathcal{T}^{(n)}_{\delta}(UV)} \lambda^{AB}_{u^n,v^n},$$

where

$$\tilde{S}_1 \triangleq \left\| (\mathrm{id} \otimes \bar{M}_A^{\otimes n} \otimes \bar{M}_B^{\otimes n})(\Psi^{\rho}_{RAB})^{\otimes n} - \frac{1}{\bar{N}_1 \bar{N}_2} \sum_{\bar{\mu}_1, \bar{\mu}_2} (\mathrm{id} \otimes [M_1^{(\bar{\mu}_1)}] \otimes [M_2^{(\bar{\mu}_2)}])(\Psi^{\rho}_{RAB})^{\otimes n} \right\|_1,$$

(a) follows by applying the triangle inequality, and (b) follows from the Lemma B.1 given below. Note that in $\tilde{S}_1$, the average over the entire common information sequence $(\bar{\mu}_1, \bar{\mu}_2)$ is inside the norm.

Lemma B.1. *We have*

$$\sum_{u^n \in \mathcal{U}^n} \sum_{v^n \in \mathcal{V}^n} \left| \lambda^{AB}_{u^n,v^n} - \frac{1}{\bar{N}_1 \bar{N}_2} \sum_{\bar{\mu}_1, \bar{\mu}_2} \alpha_{u^n} \beta_{v^n} \gamma^{(\bar{\mu}_1)}_{u^n} \zeta^{(\bar{\mu}_2)}_{v^n} \Omega_{u^n,v^n} \right| \leq S_1. \tag{B.13}$$

Proof. The proof follows from Lemma 2 in *Wilde et al.* (2012). □

Next we use Theorem III.10 twice with (a) $\rho = \rho_A$, $M = \bar{M}_A$, $\mathcal{W} = \mathcal{U}$, $\mathcal{Z} = \mathcal{U}$ and $P_{Z|W}(z|w) = \mathbb{1}\{z = w\}$, and (b) $\rho = \rho_B$, $M = \bar{M}_B$, $\mathcal{W} = \mathcal{V}$, $\mathcal{Z} = \mathcal{V}$ and $P_{Z|W}(z|w) = \mathbb{1}\{z = w\}$, and the mutual covering lemma (Lemma II.12) developed in Chapter II to yield the following: for any $\epsilon \in (0,1)$, and any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large $\mathbb{E}[S_1] \leq 2\epsilon$ if $\frac{k+l_1}{n} \log p > I(U;RB)_{\sigma_1} - S(U)_{\sigma_3} + \log p$, $\frac{k+l_2}{n} \log p > I(V;RA)_{\sigma_2} - S(V)_{\sigma_3} + \log p$, $\frac{k+l_1}{n} \log p + \frac{1}{n} \log \bar{N}_1 > \log p$, $\frac{k+l_2}{n} \log p + \frac{1}{n} \log \bar{N}_2 > \log p$, where σ_1, σ_2 and σ_3 are defined as in the statement of the theorem. Consequently, we have $\mathbb{E}[S_{21}] \leq 4\epsilon$ for all sufficiently large n.

In regards to S_{22}, note that

$$\mathbb{E}\Big[\mathbb{1}^{(\bar{\mu}_1,\bar{\mu}_2)}(u^n + v^n, i, j) \mathbb{1}_{\{a_1 G + h_1^{(\bar{\mu}_1)}(i) = u^n\}} \mathbb{1}_{\{a_2 G + h_2^{(\bar{\mu}_2)}(j) = v^n\}}\Big] \leq \sum_{\substack{\tilde{a} \in \mathbb{F}_p^k \\ \tilde{a} \neq a}} \sum_{\substack{\tilde{w} \in \mathcal{T}^{(n)}_{\hat{\delta}}(U+V) \\ \tilde{w} \neq u^n + v^n}} \frac{1}{p^n p^n p^n}.$$

Using this, we obtain

$$\mathbb{E}[S_{22}] \leq \frac{2}{(1+\eta)^2}\frac{p^{k+l_1}2^{nR_1}}{p^n}\sum_{\tilde{w}^n\in\mathcal{T}^{(n)}_{\hat{\delta}}(U+V)}\sum_{u^n\in\mathcal{T}^{(n)}_{\delta}(U)}\sum_{v^n\in\mathcal{T}^{(n)}_{\delta}(V)}\lambda^A_{u^n}\lambda^B_{v^n}\Omega_{u^n,v^n}$$
$$\leq \frac{2\ 2^{n(\frac{k+l_1}{n}\log p - R_1 - \log p + S(U+V)_{\sigma_3}+\delta_{\rho_{AB}}+\hat{\delta}_W)}}{(1+\eta)^2},$$

where $\hat{\delta}_W \searrow 0$ as $\delta \searrow 0$ and the above inequality follows from the following lemma (Lemma B.2). Hence, $\mathbb{E}[S_{21}] \leq \epsilon$ if the conditions in the proposition are satisfied.

Lemma B.2. *For $\lambda^A_{u^n}, \lambda^B_{v^n}$ and Ω_{u^n,v^n} as defined above, we have*

$$\sum_{u^n\in\mathcal{T}^{(n)}_{\delta}(U)}\sum_{v^n\in\mathcal{T}^{(n)}_{\delta}(V)}\Omega_{u^n,v^n}\lambda^A_{u^n}\lambda^B_{v^n} \leq 2^{n\delta_{\rho_{AB}}},$$

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

Proof. Firstly, note that

$$\sum_{u^n,v^n}\Omega_{u^n,v^n}\lambda^A_{u^n}\lambda^B_{v^n} = \mathrm{Tr}\Bigg\{\Bigg[(\Pi^{\bar{\mu}_1}_A\otimes\Pi^{\bar{\mu}_2}_B)\Bigg(\sqrt{\rho^{\otimes n}_A}^{-1}\Big(\sum_{u^n}\lambda^A_{u^n}\tilde{\rho}^A_{u^n}\Big)\sqrt{\rho^{\otimes n}_A}^{-1}$$
$$\otimes\sqrt{\rho^{\otimes n}_B}^{-1}\Big(\sum_{v^n}\lambda^B_{v^n}\tilde{\rho}^B_{v^n}\Big)\sqrt{\rho^{\otimes n}_B}^{-1}\Bigg)(\Pi^{\bar{\mu}_1}_A\otimes\Pi^{\bar{\mu}_2}_B)\Bigg]\rho^{\otimes n}_{AB}\Bigg\}. \tag{B.14}$$

We know, $\sum_{u^n}\lambda^A_{u^n}\tilde{\rho}^A_{u^n} \leq 2^{-n(S(\rho_A)-\delta_{\rho_A})}\Pi_{\rho_A}$, where $\delta_{\rho_A}\searrow 0$ as $\delta\searrow 0$. This implies,

$$\Pi^{\bar{\mu}_1}_A\sqrt{\rho^{\otimes n}_A}^{-1}\left(\sum_{u^n}\lambda^A_{u^n}\tilde{\rho}^A_{u^n}\right)\sqrt{\rho^{\otimes n}_A}^{-1}\Pi^{\bar{\mu}_1}_A \leq 2^{-n(S(\rho_A)-\delta_{\rho_A})}\Pi^{\bar{\mu}_1}_A\sqrt{\rho^{\otimes n}_A}^{-1}\Pi_{\rho_A}\sqrt{\rho^{\otimes n}_A}^{-1}\Pi^{\bar{\mu}_1}_A$$
$$\leq 2^{2n\delta_{\rho_A}}\Pi^{\bar{\mu}_1}_A\Pi_{\rho_A}\Pi^{\bar{\mu}_1}_A \leq 2^{2n\delta_{\rho_A}}\Pi^{\bar{\mu}_1}_A, \tag{B.15}$$

where the second inequality appeals to the fact that $\sqrt{\rho^{\otimes n}_A}^{-1}\Pi_{\rho_A}\sqrt{\rho^{\otimes n}_A}^{-1} \leq 2^{n(S(\rho_A)+\delta_{\rho_A})}\Pi_{\rho_A}$. Similarly, using the same arguments above for the operators acting on $\mathcal{H}_B$, we have

$$\Pi^{\bar{\mu}_2}_B\sqrt{\rho^{\otimes n}_B}^{-1}\left(\sum_{v^n}\lambda^B_{v^n}\tilde{\rho}^B_{v^n}\right)\sqrt{\rho^{\otimes n}_B}^{-1}\Pi^{\bar{\mu}_2}_B \leq 2^{2n\delta_{\rho_B}}\Pi^{\bar{\mu}_2}_B, \tag{B.16}$$

where $\delta_{\rho_B}\searrow 0$ as $\delta\searrow 0$. Using (i) the simplifications in (B.15) and (B.16), and (ii) the fact that

for $A_1 \geq B_1 \geq 0$ and $A_2 \geq B_2 \geq 0$, $(A_1 \otimes A_2) \geq (B_1 \otimes B_2)$ in (B.14), gives

$$\sum_{u^n, v^n} \Omega_{u^n, v^n} \lambda^A_{u^n} \lambda^B_{v^n} \leq 2^{2n(\delta_{\rho_A} + \delta_{\rho_B})} \operatorname{Tr}\left\{ \left(\Pi_A^{\bar{\mu}_1} \otimes \Pi_B^{\bar{\mu}_2} \right) \rho_{AB}^{\otimes n} \right\} \leq 2^{2n(\delta_{\rho_A} + \delta_{\rho_B})} \operatorname{Tr}\left\{ \rho_{AB}^{\otimes n} \right\} = 2^{2n(\delta_{\rho_A} + \delta_{\rho_B})}.$$

Substituting $\delta_{\rho_{AB}} = 2(\delta_{\rho_A} + \delta_{\rho_B})$ gives the result. □

B.8 Proof of Proposition III.31

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

$$\begin{aligned}
\widetilde{S}_2 &\triangleq \left\| \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{i>0} \sqrt{\rho_{AB}^{\otimes n}} \left(\Gamma_i^{A,(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P^n_{Z|U+V}(z^n | w_0^n) \right\|_1, \\
\widetilde{S}_3 &\triangleq \left\| \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{j>0} \sqrt{\rho_{AB}^{\otimes n}} \left(\Gamma_0^{A,(\mu_1)} \otimes \Gamma_j^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P^n_{Z|U+V}(z^n | w_0^n) \right\|_1, \\
\widetilde{S}_4 &\triangleq \left\| \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sqrt{\rho_{AB}^{\otimes n}} \left(\Gamma_0^{A,(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} P^n_{Z|U+V}(z^n | w_0^n) \right\|_1.
\end{aligned}$$

Analysis of $\widetilde{S}_2$: We have

$$\begin{aligned}
\widetilde{S}_2 &\leq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \sum_{i>0} \sum_{z^n} P^n_{Z|U+V}(z^n | w_0^n) \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(\Gamma_i^{A,(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\
&\leq \frac{1}{N_1 N_2} \sum_{\mu_1, \mu_2} \left\| \sqrt{\rho_B^{\otimes n}} \Gamma_0^{B,(\mu_2)} \sqrt{\rho_B^{\otimes n}} \right\|_1 \\
&\leq \frac{1}{N_2} \sum_{\mu_2} \left\| \sum_{v^n} \lambda^B_{v^n} \hat{\rho}^B_{v^n} - \sum_{v^n} \sqrt{\rho_B^{\otimes n}} \zeta^{(\mu_2)}_{v^n} \bar{B}^{(\mu_2)}_{v^n} \sqrt{\rho_B^{\otimes n}} \right\|_1 \\
&\quad + \frac{1}{N_2} \sum_{\mu_2} \left\| \sum_{v^n} \sqrt{\rho_B^{\otimes n}} \zeta^{(\mu_2)}_{v^n} \left(\bar{B}^{(\mu_2)}_{v^n} - B^{(\mu_2)}_{v^n} \right) \sqrt{\rho_B^{\otimes n}} \right\|_1 \\
&\leq \underbrace{\frac{1}{N_2} \sum_{\mu_2} \left\| \sum_{v^n} \lambda^B_{v^n} \hat{\rho}^B_{v^n} - \frac{1}{(1+\eta)} \frac{p^n}{p^{k+l_2}} \sum_{v^n} \sum_{a_2, j} \lambda^B_{v^n} \hat{\rho}^B_{v^n} \mathbb{1}_{\{\mathtt{w}_2(a_2, m_2, \mu_2) = v^n\}} \right\|_1}_{\widetilde{S}_{21}}
\end{aligned}$$

$$+\underbrace{\frac{1}{N_2}\sum_{\mu_2}\sum_{v^n}\beta_{v^n}\zeta_{v^n}^{(\mu_2)}\left\|\hat{\rho}_{v^n}^B-\tilde{\rho}_{v^n}^B\right\|_1}_{\widetilde{S}_{22}}+\underbrace{\frac{1}{N_2}\sum_{\mu_2}\sum_{v^n}\left\|\sqrt{\rho_B^{\otimes n}}\zeta_{v^n}^{(\mu_2)}\left(\bar{B}_{v^n}^{(\mu_2)}-B_{v^n}^{(\mu_2)}\right)\sqrt{\rho_B^{\otimes n}}\right\|_1}_{\widetilde{S}_{23}}, \tag{B.17}$$

where the first inequality uses triangle inequality. The next inequality follows by using Lemma III.1 where we use the fact that $\sum_{i>0}\Gamma_i^{A,(\mu_1)}\leq I$. Finally, the last two inequalities follows again from triangle inequality.

Regarding the first term in (B.17), using Lemma III.13 we claim that for all $\epsilon>0$, and $\eta,\delta\in(0,1)$ sufficiently small, and any n sufficiently large, $\mathbb{E}[\tilde{S}_{21}]<\epsilon$, if $\frac{k+l_2}{n}\log p\geq I(V;RA)_{\sigma_2}-S(V)_{\sigma_3}+\log p$, where σ_2,σ_3 are as defined in the statement of the theorem. As for the second term, we use the gentle measurement lemma (as in (B.23)) and bound its expected value as

$$\begin{aligned}\mathbb{E}[\tilde{S}_{22}]&=\mathbb{E}\left[\frac{1}{N_2}\sum_{\mu_2}\sum_{v^n}\beta_{v^n}\zeta_{v^n}^{(\mu_2)}\left\|\hat{\rho}_{v^n}^B-\tilde{\rho}_{v^n}^B\right\|_1\right]\\&=\sum_{v^n\in\mathcal{T}_\delta^{(n)}(V)}\frac{\lambda_{v^n}^B}{(1+\eta)}\left\|\hat{\rho}_{v^n}^B-\tilde{\rho}_{v^n}^B\right\|_1+\sum_{v^n\notin\mathcal{T}_\delta^{(n)}(V)}\frac{\lambda_{v^n}^B}{(1+\eta)}\left\|\hat{\rho}_{v^n}^B\right\|_1\leq\epsilon_{\tilde{S}_{21}},\end{aligned}$$

where the inequality is based on the repeated usage of the Average Gentle Measurement Lemma and $\epsilon_{\tilde{S}_{21}}\searrow 0$ as $\delta\searrow 0$ (see (35) in *Wilde et al.* (2012) for more details). Finally, consider the last term. To simplify this term, we appeal to Lemma III.15 in Section 3.3.3. This gives us

$$\tilde{S}_{23}\leq\frac{2\,2^{3n\delta}}{N_2}\sum_{\mu_2=1}^{N_2}\left(H_0^B+\frac{\sqrt{(1-\varepsilon_B)}}{(1+\eta)}\sqrt{H_1^B+H_2^B+H_3^B}\right), \tag{B.18}$$

where

$$\begin{aligned}H_0^B&\triangleq\left|\Delta_B^{(\mu_2)}-\mathbb{E}[\Delta_B^{(\mu_2)}]\right|,\quad H_1^B\triangleq\mathrm{Tr}\left\{(\Pi_{\rho_B}-\Pi_B^{\mu_2})\sum_{v^n}\lambda_{v^n}^B\tilde{\rho}_{v^n}^B\right\},\\H_2^B&\triangleq\left\|\sum_{v^n}\lambda_{v^n}^B\tilde{\rho}_{v^n}^B-(1-\varepsilon_B)\sum_{v^n}\frac{\beta_{v^n}\zeta_{v^n}^{(\mu_2)}}{\mathbb{E}[\Delta_B^{(\mu)}]}\tilde{\rho}_{v^n}^B\right\|_1,\\H_3^B&\triangleq(1-\varepsilon_B)\left\|\sum_{v^n}\frac{\beta_{v^n}\zeta_{v^n}^{(\mu_2)}}{\Delta_B^{(\mu_2)}}\tilde{\rho}_{v^n}^B-\sum_{v^n}\frac{\beta_{v^n}\zeta_{v^n}^{(\mu_2)}}{\mathbb{E}[\Delta_B^{(\mu_2)}]}\tilde{\rho}_{v^n}^B\right\|_1,\end{aligned} \tag{B.19}$$

and $\Delta_B^{(\mu)}\triangleq\sum_{v^n\in\mathcal{T}_\delta^{(n)}(V)}\beta_{v^n}\zeta_{v^n}^{(\mu_2)}$ and $\varepsilon_B=\sum_{v^n\notin\mathcal{T}_\delta^{(n)}(V)}\lambda_{v^n}^B$. Further, using the simplification

performed in (3.19), (3.21), and (3.22), and the concavity of the square-root function, we obtain,

$$\mathbb{E}[\tilde{S}_{23}] \leq \frac{2}{N_2} 2^{3n\delta_{\rho_B}} \sum_{\mu_2=1}^{N_2} \left(\mathbb{E}[H_0^B] + \frac{(1-\varepsilon_B)}{(1+\eta)} \sqrt{\left(\frac{2^{2n\delta_{\rho_B}}}{\eta} + 1 \right) \mathbb{E}[\widetilde{H}^B]} + \sqrt{\frac{(1-\varepsilon_B)}{(1+\eta)}} \sqrt{\mathbb{E}[H_0^B]} \right),$$

$$\text{where } \widetilde{H}^B \triangleq \left\| \frac{1}{(1-\varepsilon_B)} \sum_{v^n} \lambda_{v^n}^B \tilde{\rho}_{v^n}^B - \frac{p^n}{p^{k+l_2}} \sum_{v^n} \sum_{a_2, j>0} \frac{\lambda_{v^n}^B \tilde{\rho}_{v^n}^B}{(1-\varepsilon_B)} \mathbb{1}_{\{\mathtt{w}_2(a_2,m_2,\mu_2)=v^n\}} \right\|_1. \tag{B.20}$$

Using Proposition III.16, for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}\left[\tilde{S}_{23}\right] \leq \epsilon$ if $\frac{k+l_2}{n} \log p > I(V;RA)_{\sigma_2} + \log p - S(V)_{\sigma_3}$, where σ_2, σ_3 are the auxiliary state defined in the statement of the theorem.

Analysis of $\widetilde{S}_3$: Due to the symmetry in $\widetilde{S}_2$ and $\widetilde{S}_3$, the analysis of $\widetilde{S}_3$ follows very similar arguments as that of $\widetilde{S}_2$ and hence we obtain the following, for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}\left[\tilde{S}_3\right] \leq \epsilon$ if $S_1 > I(U;RB)_{\sigma_1} + \log p - S(U)_{\sigma_3}$, where σ_1, σ_3 are the auxiliary state defined in the statement of the theorem.

Analysis of $\widetilde{S}_4$: We have

$$\begin{aligned}
\widetilde{S}_4 &\leq \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{z^n} P_{Z|U+V}^n(z^n|w_0^n) \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(\Gamma_0^{A,(\mu_1)} \otimes \Gamma_0^{B,(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 \\
&\leq \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(\Gamma_0^{A,(\mu_1)} \otimes I \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1 + \frac{1}{N_1 N_2} \sum_{\mu_1,\mu_2} \sum_{v^n} \left\| \sqrt{\rho_{AB}^{\otimes n}} \left(\Gamma_0^{A,(\mu_1)} \otimes B_{v^n}^{(\mu_2)} \right) \sqrt{\rho_{AB}^{\otimes n}} \right\|_1,
\end{aligned} \tag{B.21}$$

where the inequalities above are obtained by a straight forward substitution and use of triangle inequality. Further, since $0 \leq \Gamma_0^{A,(\mu_1)} \leq I$ and $0 \leq \Gamma_0^{B,(\mu_2)} \leq I$, this simplifies the first term in (B.21) 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 (B.21) simplifies using Lemma III.1 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

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

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

B.9 Proof of Proposition III.32

We start by applying triangle inequality to obtain $J_1 \leq J_{11} + J_{12}$, where

$$J_{11} \triangleq \sum_{z^n, v^n} \left\| \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \frac{\alpha_{u^n} \gamma_{u^n}^{(\mu_1)}}{\lambda_{u^n}^A} \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|W}^n(z^n | u^n + v^n) \right\|_1 ,$$

$$J_{12} \triangleq \sum_{z^n, v^n} \left\| \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sum_{u^n} \sqrt{\rho_{AB}^{\otimes n}} \left(\frac{\alpha_{u^n} \gamma_{u^n}^{(\mu_1)}}{\lambda_{u^n}^A} \bar{\Lambda}_{u^n}^A \otimes \bar{\Lambda}_{v^n}^B - \gamma_{u^n}^{(\mu_1)} \bar{A}_{u^n}^{(\mu_1)} \otimes \bar{\Lambda}_{v^n}^B \right) \sqrt{\rho_{AB}^{\otimes n}} P_{Z|W}^n(z^n | u^n + v^n) \right\|_1 ,$$

Now with the intention of employing Lemma III.13, we express J_{11} as

$$J_{11} = \Bigg\| \sum_{u^n, v^n, z^n} \lambda_{u^n, v^n}^{AB} \hat{\rho}_{u^n, v^n}^{AB} \otimes \phi_{u^n, v^n, z^n} - \frac{1}{(1+\eta)} \frac{p^n}{p^{k+l_1} N_1} \sum_{\mu_1} \sum_{u^n, v^n, z^n} \sum_{a_1, i>0} \lambda_{u^n}^A \mathbb{1}_{\{\mathtt{w}_1(a_1, m_1, \mu_1) = u^n\}} \frac{\lambda_{u^n, v^n}^{AB}}{\lambda_{u^n}^A} \hat{\rho}_{u^n, v^n}^{AB} \otimes \phi_{u^n, v^n, z^n} \Bigg\|_1 ,$$

where the equality above is obtained by defining $\phi_{u^n, v^n, z^n} = P_{Z|W}^n(z^n | u^n + v^n) \, |v^n\rangle\langle v^n| \otimes |z^n\rangle\langle z^n|$ and using the definitions of $\alpha_{u^n}, \gamma_{u^n}^{(\mu_1)}$ and $\hat{\rho}_{u^n, v^n}^{AB}$, followed by using the triangle inequality for the block diagonal operators. Note that the triangle inequality in this case becomes an equality.
Let us define $\mathcal{T}_{u^n}$ as

$$\mathcal{T}_{u^n} \triangleq \sum_{v^n, z^n} \frac{\lambda_{u^n, v^n}^{AB}}{\lambda_{u^n}^A} \hat{\rho}_{u^n, v^n}^{AB} \otimes \phi_{u^n, v^n, z^n} .$$

Note that in the above definition of $\mathcal{T}_{u^n}$ we have $\mathcal{T}_{u^n} \geq 0$ and $\mathrm{Tr}\{\mathcal{T}_{u^n}\} = 1$ for all $u^n \in \mathbb{F}_p^n$. Further, it contains all the elements in product form, and thus can be written as $\mathcal{T}_{u^n} = \bigotimes_{i=1}^n \mathcal{T}_{u_i}$. This

simplifies J_{11} as

$$J_{11} = \left\| \sum_{u^n} \lambda^A_{u^n} \mathcal{T}_{u^n} - \frac{1}{(1+\eta)} \frac{p^n}{p^{k+l_1}} \frac{1}{N_1} \sum_{\mu_1} \sum_{u^n} \sum_{a_1, i>0} \lambda^A_{u^n} \mathcal{T}_{u^n} \mathbb{1}_{\{\mathtt{w}_1(a_1, m_1, \mu_1) = u^n\}} \right\|_1 .$$

Using Lemma III.13, we claim the following: for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[J_{11}] \leq \epsilon$, if $\frac{k+l_1}{n} \log p + \frac{1}{n} \log N_1 > I(U;RZV)_{\sigma_3} - S(U)_{\sigma_3} + \log p$, where σ_3 is the auxiliary state defined in the statement of the theorem.

Now we consider the term corresponding to J_{12} and prove that its expectation with respect to the Alice's codebook is small. Recalling J_{12}, we get

$$\begin{aligned} J_{12} &\leq \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sum_{u^n, v^n} \sum_{z^n} P^n_{Z|W}(z^n | u^n + v^n) \left\| \sqrt{\rho^{\otimes n}_{AB}} \left(\frac{\alpha_{u^n} \gamma^{(\mu_1)}_{u^n}}{\lambda^A_{u^n}} \bar{\Lambda}^A_{u^n} \otimes \bar{\Lambda}^B_{v^n} - \gamma^{(\mu_1)}_{u^n} \bar{A}^{(\mu_1)}_{u^n} \otimes \bar{\Lambda}^B_{v^n} \right) \sqrt{\rho^{\otimes n}_{AB}} \right\|_1, \\ &= \frac{1}{N_1} \sum_{\mu_1=1}^{N_1} \sum_{u^n, v^n} \alpha_{u^n} \gamma^{(\mu_1)}_{u^n} \left\| \sqrt{\rho^{\otimes n}_{AB}} \left(\left(\frac{1}{\lambda^A_{u^n}} \bar{\Lambda}^A_{u^n} - \overline{\sqrt{\rho^{\otimes n}_A}^{-1}} \tilde{\rho}^A_{u^n} \overline{\sqrt{\rho^{\otimes n}_A}^{-1}} \right) \otimes \bar{\Lambda}^B_{v^n} \right) \sqrt{\rho^{\otimes n}_{AB}} \right\|_1, \end{aligned}$$

where the inequality is obtained by using triangle and the next equality follows from the fact that $\sum_{z^n} P^n_{Z|W}(z^n | u^n + v^n) = 1$ for all $u^n \in \mathcal{U}^n$ and $v^n \in \mathcal{V}^n$ and using the definition of $A^{(\mu_1)}_{u^n}$. By applying expectation of J_{12} over the Alice's codebook, we get

$$\mathbb{E}[J_{12}] \leq \frac{1}{(1+\eta)} \sum_{u^n} \lambda^A_{u^n} \sum_{v^n} \left\| \sqrt{\rho^{\otimes n}_{AB}} \left(\left(\frac{1}{\lambda^A_{u^n}} \bar{\Lambda}^A_{u^n} - \overline{\sqrt{\rho^{\otimes n}_A}^{-1}} \tilde{\rho}^A_{u^n} \overline{\sqrt{\rho^{\otimes n}_A}^{-1}} \right) \otimes \bar{\Lambda}^B_{v^n} \right) \sqrt{\rho^{\otimes n}_{AB}} \right\|_1,$$

where we have used the fact that $\mathbb{E}[\alpha_{u^n} \gamma^{(\mu_1)}_{u^n}] = \frac{\lambda^A_{u^n}}{(1+\eta)}$. To simplify the above equation, we employ Lemma III.1 which completely discards the effect of Bob's measurement. Since $\sum_{v^n} \bar{\Lambda}^B_{v^n} = I$, from Lemma III.1 we have for every u^n,

$$\begin{aligned} &\sum_{v^n} \left\| \sqrt{\rho^{\otimes n}_{AB}} \left(\left(\frac{1}{\lambda^A_{u^n}} \bar{\Lambda}^A_{u^n} - \overline{\sqrt{\rho^{\otimes n}_A}^{-1}} \tilde{\rho}^A_{u^n} \overline{\sqrt{\rho^{\otimes n}_A}^{-1}} \right) \otimes \bar{\Lambda}^B_{v^n} \right) \sqrt{\rho^{\otimes n}_{AB}} \right\|_1 \\ &\qquad = \left\| \sqrt{\rho^{\otimes n}_A} \left(\frac{1}{\lambda^A_{u^n}} \bar{\Lambda}^A_{u^n} - \overline{\sqrt{\rho^{\otimes n}_A}^{-1}} \tilde{\rho}^A_{u^n} \overline{\sqrt{\rho^{\otimes n}_A}^{-1}} \right) \sqrt{\rho^{\otimes n}_A} \right\|_1 . \end{aligned}$$

This simplifies $\mathbb{E}[J_{12}]$ as

$$\mathbb{E}[J_{12}] \leq \frac{1}{(1+\eta)} \sum_{u^n} \lambda^A_{u^n} \left\| \sqrt{\rho^{\otimes n}_A} \left(\frac{1}{\lambda^A_{u^n}} \bar{\Lambda}^A_{u^n} - \overline{\sqrt{\rho^{\otimes n}_A}^{-1}} \tilde{\rho}^A_{u^n} \overline{\sqrt{\rho^{\otimes n}_A}^{-1}} \right) \sqrt{\rho^{\otimes n}_A} \right\|_1$$

$$\leq \frac{1}{(1+\eta)} \sum_{u^n \notin \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A \left\| \hat{\rho}_{u^n}^A \right\|_1 + \frac{1}{(1+\eta)} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A \left\| \left(\hat{\rho}_{u^n}^A - \tilde{\rho}_{u^n}^A \right) \right\|_1 \leq \varepsilon_A + \epsilon'_{J_{12}} \quad \text{(B.22)}$$

where the last inequality is obtained by repeated usage of the Average Gentle Measurement Lemma and $\epsilon'_{J_{12}} \searrow 0$ as $\delta \searrow 0$ (see (35) in *Wilde et al.* (2012) for details). This completes the proof.

B.10 Proof of Proposition III.33

Noting the similarity between J_2 and the term $\tilde{S}_2$ defined in the proof of Theorem III.10 (see Section 3.3.3), we begin by further simplifying J_2 using Lemma III.15. This gives us

$$J_2 \leq \frac{22^{3n\delta_{\rho_A}}}{N_1} \sum_{\mu_1=1}^{N_1} \left(H_0^A + \frac{\sqrt{(1-\varepsilon_A)}}{(1+\eta)} \sqrt{H_1^A + H_2^A + H_3^A} \right), \quad \text{where}$$

$$H_0^A \triangleq \left| \Delta_A^{(\mu_1)} - \mathbb{E}[\Delta_A^{(\mu_1)}] \right|, \quad H_1^A \triangleq \operatorname{Tr}\left\{ (\Pi_{\rho_A} - \Pi_A^{\mu_1}) \sum_{w^n} \lambda_{u^n}^A \tilde{\rho}_{u^n}^A \right\},$$
$$H_2^A \triangleq \left\| \sum_{u^n} \lambda_{u^n}^A \tilde{\rho}_{u^n}^A - (1-\varepsilon_A) \sum_{u^n} \frac{\alpha_{u^n} \gamma_{u^n}^{(\mu_1)}}{\mathbb{E}[\Delta_A^{(\mu_1)}]} \tilde{\rho}_{u^n}^A \right\|_1,$$
$$H_3^A \triangleq (1-\varepsilon_A) \left\| \sum_{u^n} \frac{\alpha_{u^n} \gamma_{u^n}^{(\mu_1)}}{\Delta_A^{(\mu_1)}} \tilde{\rho}_{u^n}^A - \sum_{u^n} \frac{\alpha_{u^n} \gamma_{u^n}^{(\mu_1)}}{\mathbb{E}[\Delta_A^{(\mu_1)}]} \tilde{\rho}_{u^n}^A \right\|_1,$$

and $\Delta_A^{(\mu_1)} \triangleq \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \alpha_{u^n} \gamma_{u^n}^{(\mu_1)}$, $\varepsilon_A \triangleq \sum_{u^n \notin \mathcal{T}_\delta^{(n)}(U)} \lambda_{u^n}^A$, and $\delta_{\rho_A}(\delta) \searrow 0$ as $\delta \searrow 0$. Further, using the simplification performed in (3.19), (3.21), and (3.22), and the concavity of the square-root function, we obtain,

$$\mathbb{E}[J_2] \leq \frac{2}{N_1} 2^{3n\delta_{\rho_A}} \sum_{\mu_1=1}^{N_1} \left(\mathbb{E}[H_0^A] + \frac{(1-\varepsilon_A)}{(1+\eta)} \sqrt{\left(\frac{2^{2n\delta_{\rho_A}}}{\eta} + 1 \right) \mathbb{E}[\widetilde{H}^A]} + \sqrt{\frac{(1-\varepsilon_A)}{(1+\eta)}} \sqrt{\mathbb{E}[H_0^A]} \right),$$

where

$$\widetilde{H}^A \triangleq \left\| \sum_{u^n} \frac{\lambda_{u^n}^A}{(1-\varepsilon_A)} \tilde{\rho}_{u^n}^A - \frac{p^n}{2^{nS_1}} \sum_{u^n} \sum_{a_1, i>0} \frac{\lambda_{u^n}^A}{(1-\varepsilon_A)} \tilde{\rho}_{u^n}^A \mathbb{1}_{\{\mathtt{w}_1(a_1, m_1, \mu_1) = u^n\}} \right\|_1.$$

The proof from here follows from Proposition III.16.

B.11 Proof of Proposition III.34

We start by adding and subtracting the following terms within Q_2

$$(i)\ \sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\bar{\Lambda}_{u^n}^{A}\otimes\bar{\Lambda}_{v^n}^{B}\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^{n}(z^n|u^n+v^n),$$

$$(ii)\ \sum_{u^n,v^n}\frac{1}{N_2}\sum_{\mu_2=1}^{N_2}\sqrt{\rho_{AB}^{\otimes n}}\left(\bar{\Lambda}_{u^n}^{A}\otimes\frac{\beta_{v^n}\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^{B}}\bar{\Lambda}_{v^n}^{B}\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^{n}(z^n|u^n+v^n),$$

$$(iii)\ \sum_{u^n,v^n}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho_{AB}^{\otimes n}}\left(\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\otimes\frac{\beta_{v^n}\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^{B}}\bar{\Lambda}_{v^n}^{B}\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^{n}(z^n|u^n+v^n),$$

$$(iv)\ \sum_{u^n,v^n}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho_{AB}^{\otimes n}}\left(\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\otimes\zeta_{v^n}^{(\mu_2)}\bar{B}_{v^n}^{(\mu_2)}\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^{n}(z^n|u^n+v^n).$$

This gives us $Q_2\le Q_{21}+Q_{22}+Q_{23}+Q_{24}+Q_{25}$, where

$$Q_{21}\triangleq\sum_{z^n}\left\|\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\left(\frac{1}{N_1}\sum_{\mu_1=1}^{N_1}\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\right)\otimes\bar{\Lambda}_{v^n}^{B}-\bar{\Lambda}_{u^n}^{A}\otimes\bar{\Lambda}_{v^n}^{B}\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^{n}(z^n|u^n+v^n)\right\|_1,$$

$$Q_{22}\triangleq\sum_{z^n}\left\|\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\bar{\Lambda}_{u^n}^{A}\otimes\bar{\Lambda}_{v^n}^{B}-\bar{\Lambda}_{u^n}^{A}\otimes\left(\frac{1}{N_2}\sum_{\mu_2=1}^{N_2}\frac{\beta_{v^n}\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^{B}}\bar{\Lambda}_{v^n}^{B}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^{n}(z^n|u^n+v^n)\right\|_1,$$

$$Q_{23}\triangleq\sum_{z^n}\left\|\sum_{u^n,v^n}\sqrt{\rho_{AB}^{\otimes n}}\left(\left(\bar{\Lambda}_{u^n}^{A}-\frac{1}{N_1}\sum_{\mu_1}\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\right)\otimes\left(\frac{1}{N_2}\sum_{\mu_2}\frac{\beta_{v^n}\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^{B}}\bar{\Lambda}_{v^n}^{B}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^{n}(z^n|u^n+v^n)\right\|_1,$$

$$Q_{24}\triangleq\sum_{z^n}\left\|\sum_{u^n,v^n}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho_{AB}^{\otimes n}}\left(\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\otimes\left(\frac{\beta_{v^n}\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^{B}}\bar{\Lambda}_{v^n}^{B}-\zeta_{v^n}^{(\mu_2)}\bar{B}_{v^n}^{(\mu_2)}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^{n}(z^n|u^n+v^n)\right\|_1,$$

$$Q_{25}\triangleq\sum_{z^n}\left\|\sum_{u^n,v^n}\frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho_{AB}^{\otimes n}}\left(\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\otimes\left(\zeta_{v^n}^{(\mu_2)}\bar{B}_{v^n}^{(\mu_2)}-\zeta_{v^n}^{(\mu_2)}B_{v^n}^{(\mu_2)}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}P_{Z|W}^{n}(z^n|u^n+v^n)\right\|_1.$$

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 $Q_{22}=$

$$\left\|\sum_{u^n,v^n,z^n}\lambda_{u^n,v^n}^{AB}\hat{\rho}_{u^n,v^n}^{AB}\otimes\psi_{u^n,v^n,z^n}-\frac{1}{N_2}\sum_{\mu_2}\sum_{u^n,v^n,z^n}\beta_{v^n}\sum_{a_2,j>0}\mathbb{1}_{\{\mathtt{w}_2(a_2,m_2,\mu_2)=v^n\}}\frac{\lambda_{u^n,v^n}^{AB}}{\lambda_{v^n}^{B}}\hat{\rho}_{u^n,v^n}^{AB}\otimes\psi_{u^n,v^n,z^n}\right\|_1,$$

where ψ_{u^n,v^n,z^n} is defined as $\psi_{u^n,v^n,z^n} = P^n_{Z|W}(z^n|u^n+v^n)\,|z^n\rangle\langle z^n|$, and the equality uses the triangle inequality for block operators. Now we use Lemma III.13 to bound Q_{22}. Let

$$\mathcal{T}_{v^n} \triangleq \sum_{u^n,z^n} \frac{\lambda^{AB}_{u^n,v^n}}{\lambda^B_{v^n}} \hat{\rho}^{AB}_{u^n,v^n} \otimes \psi_{u^n,v^n,z^n}.$$

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

$$Q_{22} = \left\| \sum_{v^n} \lambda^B_{v^n} \mathcal{T}_{v^n} - \frac{1}{(1+\eta)} \frac{p^n}{2^{nS_2} N_2} \sum_{\mu_2} \sum_{v^n} \sum_{a_2,j>0} \lambda^B_{v^n} \mathcal{T}_{v^n} \mathbb{1}_{\{\mathtt{w}_2(a_2,m_2,\mu_2)=v^n\}} \right\|_1.$$

Application of Lemma III.13 gives the following: for any $\epsilon \in (0,1)$, any $\eta, \delta \in (0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[Q_{22}] \leq \epsilon$ if

$$\frac{k+l_2}{n} \log p + \frac{1}{n} \log N_2 > I(V;RZ)_{\sigma_3} - S(V)_{\sigma_3} + \log p.$$

Now, we move on to consider Q_{23}. Taking expectation with respect $G, h_1^{(\mu_1)}, h_2^{(\mu_2)}$ gives

$$\begin{aligned}
\mathbb{E}\left[Q_{23}\right] \leq \mathbb{E}\Bigg[& \sum_{z^n,v^n} \frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \frac{\beta_{v^n} \zeta^{(\mu_2)}_{v^n}}{\lambda^B_{v^n}} \Bigg\| \sum_{u^n} \sqrt{\rho^{\otimes n}_{AB}} \left(\bar{\Lambda}^A_{u^n} \otimes \bar{\Lambda}^B_{v^n}\right) \sqrt{\rho^{\otimes n}_{AB}} P^n_{Z|W}(z^n|u^n+v^n) \\
& - \sum_{u^n} \sqrt{\rho^{\otimes n}_{AB}} \left(\frac{1}{N_1} \sum_{\mu_1} \gamma^{(\mu_1)}_{u^n} A^{(\mu_1)}_{u^n} \otimes \bar{\Lambda}^B_{v^n} \right) \sqrt{\rho^{\otimes n}_{AB}} P^n_{Z|W}(z^n|u^n+v^n) \Bigg\|_1 \Bigg] \\
= \mathbb{E}_{G,h_1} \Bigg[& \sum_{z^n,v^n} \frac{1}{N_2} \sum_{\mu_2=1}^{N_2} \frac{\mathbb{E}_{h_2|G}\left[\beta_{v^n} \zeta^{(\mu_2)}_{v^n} | G\right]}{\lambda^B_{v^n}} \Bigg\| \sum_{u^n} \sqrt{\rho^{\otimes n}_{AB}} \left(\bar{\Lambda}^A_{u^n} \otimes \bar{\Lambda}^B_{v^n}\right) \sqrt{\rho^{\otimes n}_{AB}} P^n_{Z|W}(z^n|u^n+v^n) \\
& - \sum_{u^n} \sqrt{\rho^{\otimes n}_{AB}} \left(\frac{1}{N_1} \sum_{\mu_1} \gamma^{(\mu_1)}_{u^n} A^{(\mu_1)}_{u^n} \otimes \bar{\Lambda}^B_{v^n} \right) \sqrt{\rho^{\otimes n}_{AB}} P^n_{Z|W}(z^n|u^n+v^n) \Bigg\|_1 \Bigg] \\
= \mathbb{E}_{G,h_1} \Bigg[& \sum_{z^n,v^n} \frac{1}{(1+\eta)} \Bigg\| \sum_{u^n} \sqrt{\rho^{\otimes n}_{AB}} \left(\bar{\Lambda}^A_{u^n} \otimes \bar{\Lambda}^B_{v^n}\right) \sqrt{\rho^{\otimes n}_{AB}} P^n_{Z|W}(z^n|u^n+v^n) \\
& - \sum_{u^n} \sqrt{\rho^{\otimes n}_{AB}} \left(\frac{1}{N_1} \sum_{\mu_1} \gamma^{(\mu_1)}_{u^n} A^{(\mu_1)}_{u^n} \otimes \bar{\Lambda}^B_{v^n} \right) \sqrt{\rho^{\otimes n}_{AB}} P^n_{Z|W}(z^n|u^n+v^n) \Bigg\|_1 \Bigg] = \mathbb{E}\left[\frac{J}{(1+\eta)}\right],
\end{aligned}$$

where the inequality above is obtained by using the triangle inequality, and the first equality follows from $h_1^{(\mu_1)}$ and $h_2^{(\mu_2)}$ being generated independently. The last equality follows from the definition

of J as in (3.38). Hence, we use the result obtained in bounding $\mathbb{E}[J]$. Next, we consider Q_{24}.

$$\begin{aligned}
Q_{24} \leq \sum_{u^n,v^n}\sum_{z^n} P^n_{Z|W}(z^n|u^n+v^n)\Bigg\| \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho_{AB}^{\otimes n}}\left(\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\otimes\frac{\beta_{v^n}\zeta_{v^n}^{(\mu_2)}}{\lambda_{v^n}^B}\bar{\Lambda}_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}} \\
- \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sqrt{\rho_{AB}^{\otimes n}}\left(\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\otimes\beta_{v^n}\zeta_{v^n}^{(\mu_2)}\left(\sqrt{\rho_B}^{-1}\tilde{\rho}_{v^n}^B\sqrt{\rho_B}^{-1}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}\Bigg\|_1 \\
\leq \frac{1}{N_2}\sum_{\mu_2}\sum_{u^n,v^n}\beta_{v^n}\zeta_{v^n}^{(\mu_2)}\Bigg\|\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\otimes\frac{1}{\lambda_{v^n}^B}\bar{\Lambda}_{v^n}^B\right)\sqrt{\rho_{AB}^{\otimes n}} \\
- \sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\otimes\left(\sqrt{\rho_B}^{-1}\tilde{\rho}_{v^n}^B\sqrt{\rho_B}^{-1}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}\Bigg\|_1,
\end{aligned}$$

where the inequalities follow from the definition of $\bar{B}_{v^n}^{(\mu_2)}$ and using multiple triangle inequalities. Taking expectation of Q_{24} with respect to $h_2^{(\mu_2)}$, we get

$$\begin{aligned}
\mathbb{E}\left[Q_{24}\right] &\leq \mathbb{E}_{G,h_1}\left[\sum_{u^n,v^n}\frac{\lambda_{v^n}^B}{(1+\eta)}\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\frac{1}{N_1}\sum_{\mu_1}\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\otimes\left(\frac{1}{\lambda_{v^n}^B}\bar{\Lambda}_{v^n}^B-\sqrt{\rho_B}^{-1}\tilde{\rho}_{v^n}^B\sqrt{\rho_B}^{-1}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|\right] \\
&\leq \mathbb{E}_{G,h_1}\left[\sum_{v^n}\frac{\lambda_{v^n}^B}{(1+\eta)}\left\|\sqrt{\rho_B^{\otimes n}}\left(\frac{1}{\lambda_{v^n}^B}\bar{\Lambda}_{v^n}^B-\sqrt{\rho_B}^{-1}\tilde{\rho}_{v^n}^B\sqrt{\rho_B}^{-1}\right)\sqrt{\rho_B^{\otimes n}}\right\|_1\right] \\
&= \sum_{v^n\notin\mathcal{T}_\delta^{(n)}(V)}\frac{\lambda_{v^n}^B}{(1+\eta)}\left\|\hat{\rho}_{v^n}^B\right\|_1+\sum_{v^n\in\mathcal{T}_\delta^{(n)}(V)}\frac{\lambda_{v^n}^B}{(1+\eta)}\left\|\hat{\rho}_{v^n}^B-\tilde{\rho}_{v^n}^B\right\|_1\leq\varepsilon_B+\epsilon'_{Q_{24}},
\end{aligned}\tag{B.23}$$

where the second inequality follows by using Lemma III.1 and the fact that $\frac{1}{N_1}\sum_{\mu_1}\sum_{u^n}\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\leq I$, and the last inequality follows by applying the Average Gentle Measurement Lemma repeated and $\epsilon'_{Q_{24}}\searrow 0$ as $\delta\searrow 0$ (see (35) in *Wilde et al.* (2012) for more details). This completes the proof for the term Q_{24}. Finally, we move onto considering Q_{25}. Simplifying Q_{25} gives

$$\begin{aligned}
Q_{25} &\leq \frac{1}{N_1N_2}\sum_{\mu_1,\mu_2}\sum_{z^n}\sum_{u^n,v^n}P^n_{Z|W}(z^n|u^n+v^n)\left\|\sqrt{\rho_{AB}^{\otimes n}}\left(\gamma_{u^n}^{(\mu_1)}A_{u^n}^{(\mu_1)}\otimes\left(\zeta_{v^n}^{(\mu_2)}\bar{B}_{v^n}^{(\mu_2)}-\zeta_{v^n}^{(\mu_2)}B_{v^n}^{(\mu_2)}\right)\right)\sqrt{\rho_{AB}^{\otimes n}}\right\|_1 \\
&\leq \frac{1}{N_2}\sum_{\mu_2}\sum_{v^n}\left\|\sqrt{\rho_B^{\otimes n}}\left(\zeta_{v^n}^{(\mu_2)}\bar{B}_{v^n}^{(\mu_2)}-\zeta_{v^n}^{(\mu_2)}B_{v^n}^{(\mu_2)}\right)\sqrt{\rho_B^{\otimes n}}\right\|_1=\tilde{S}_{23},
\end{aligned}$$

where the first inequality uses traingle inequality and the second inequality uses Lemma III.1 to remove the affect of approximating Alice's POVM on Bob's approximation, and $\tilde{S}_{23}$ is defined in (B.17) in the proof of Proposition III.31. Therefore, we have the following: for any $\epsilon\in(0,1)$, any $\eta,\delta\in(0,1)$ sufficiently small, and any n sufficiently large, we have $\mathbb{E}[Q_{25}]\leq\epsilon$, if $S_2\geq I(V;RA)_{\sigma_2}-$

$S(V)_{\sigma_3} + \log p$. This completes the proof for Q_{25} and hence for all the terms corresponding to Q_2.

APPENDIX C

Proofs of Chapter IV

C.1 Proof of Lemma IV.8

We begin by observing the fact that $|\hat{\sigma}_{u^n}\rangle^{AE}$ and $|\tilde{\sigma}_{u^n}\rangle^{AE}$ are purification of the isometric versions of states $\hat{\rho}^A_{u^n}$ and $\frac{\tilde{\rho}^A_{u^n}}{\mathrm{Tr}\{\hat{\rho}^A_{u^n}\}}$, respectively. More precisely,

$$\mathrm{Tr}_A\,|\hat{\sigma}_{u^n}\rangle\langle\hat{\sigma}_{u^n}|^{AE} = V^E\hat{\rho}^A_{u^n}(V^E)^\dagger \quad \text{and} \quad \mathrm{Tr}_A\,|\tilde{\sigma}_{u^n}\rangle\langle\hat{\sigma}_{u^n}|^{AE} = \frac{V^E\tilde{\rho}^A_{u^n}(V^E)^\dagger}{\mathrm{Tr}\{\tilde{\rho}^A_{u^n}\}},$$

where we use the fact that

$$\left|\Psi_{\rho^{\otimes n}}\right\rangle^{ABCR} \triangleq V^E\left|\Psi_{\rho^{\otimes n}}\right\rangle^{AA'},$$

for some isometry V^E, and $\left|\Psi_{\rho^{\otimes n}}\right\rangle^{AA'}$ is the canonical purification of $\rho_A^{\otimes n}$. Recall that $\hat{\rho}^A_{u^n}$ is the post-measurement state of the canonical reference obtained after observing the classical outcome u^n, and thus $\hat{\rho}^A_{u^n} \in \mathcal{D}(\mathcal{H}_E^{\otimes n})$. Also, note that the Hilbert space $\mathcal{H}_A^{\otimes}$, corresponding to the subsystem A^n purifies the states $\hat{\rho}^A_{u^n}$ and $\tilde{\rho}^A_{u^n}$. This implies, from Uhlmann's theorem, there exists a unitary $U_r(l)$ acting on the subsystem A^n, such that, for $u^n = U^n(l)$, we have

$$F(\hat{\rho}^A_{u^n}, \frac{\tilde{\rho}^A_{u^n}}{\mathrm{Tr}\{\tilde{\rho}^A_{u^n}\}}) = F\left(V^E\hat{\rho}^A_{u^n}(V^E)^\dagger, \frac{V^E\tilde{\rho}^A_{u^n}(V^E)^\dagger}{\mathrm{Tr}\{\tilde{\rho}^A_{u^n}\}}\right) = F(|\hat{\sigma}_{u^n}\rangle^{AE}, U_r(l)\,|\tilde{\sigma}_{u^n}\rangle^{AE}).$$

Finally, using the relation (*Wilde*, 2013a, Theorem 9.3.1), and the inequality

$$\|\hat{\rho}^A_{u^n} - \frac{\tilde{\rho}^A_{u^n}}{\mathrm{Tr}\{\tilde{\rho}^A_{u^n}\}}\|_1 \leq \|\hat{\rho}^A_{u^n} - \tilde{\rho}^A_{u^n}\|_1 + (1 - \mathrm{Tr}\{\tilde{\rho}^A_{u^n}\}),$$

we obtain the desired result. An identical analysis for $|\hat{\sigma}_{v^n}\rangle^{BF}$ and $|\tilde{\sigma}_{v^n}\rangle^{BF}$ produces the second statement of the lemma.

C.2 Proof of Lemma IV.9

Observe that

$$\mathcal{N}_A(\rho_A) = \sum_{l=1}^{2^{n\tilde{R}_1}} \gamma \, \mathrm{Tr}_{EA_g} \left\{ (I^E \otimes U_p^A(l) U_r^A(l) \sqrt{A_l}) \Psi^{ABCR}_{\rho^{\otimes n}} (I^E \otimes U_p^A(l) U_r^A(l) \sqrt{A_l})^\dagger \right\},$$

where E denotes the subsystems corresponding to B, C, and R. Using the triangle inequality and monotonicity of trace distance, we obtain

$$\left\| \mathcal{N}_A(\rho_A) - |0\rangle\langle 0|_{A_p} \right\|_1 \mathbb{1}_{\{sP\}} \leq T_1 + T_2 + T_3,$$

where

$$\begin{aligned}
T_1 &\triangleq \sum_l \gamma \left\| (I \otimes U_p^A(l)) \left[\mathrm{Tr}(\tilde{\rho}_l^A)(I \otimes U_r^A(l)) \tilde{\sigma}_l^{AE} (I \otimes U_r^A(l))^\dagger - \hat{\sigma}_l^{AE} \right] (I \otimes U_p^A(l))^\dagger \right\|_1, \\
T_2 &\triangleq 1 - \sum_l \gamma \, \mathrm{Tr}\{\tilde{\rho}_l^A\}, \quad T_3 \triangleq \sum_l \gamma \left\| \mathrm{Tr}_{A_g} \left\{ U_p^A(l) \hat{\sigma}_l^A (U_p^A(l))^\dagger \right\} - |0\rangle\langle 0|_{A_p} \right\|_1
\end{aligned} \tag{C.1}$$

Further, using the trace distance and fidelity relation (*Wilde*, 2013a, Theorem 9.3.1)

$$\begin{aligned}
T_1 &\leq T_2 + \sum_l \gamma \left\| (I \otimes U_r^A(l)) \tilde{\sigma}_l^{AE} (I \otimes U_r^A(l))^\dagger - \hat{\sigma}_l^{AE} \right\|_1 \\
&\leq T_2 + 2\sqrt{1 - F((I \otimes U_r^A(l))|\tilde{\sigma}_l^{AE}\rangle, |\hat{\sigma}_l^{AE}\rangle)}.
\end{aligned} \tag{C.2}$$

Using the result from Lemma IV.8 and (4.5), and taking expectation, we obtain

$$\mathbb{E}\left[T_1 + T_2\right] \leq 2\mathbb{E}[T_2] + 2\sqrt{\mathbb{E}\left[\sum_l \gamma \|\hat{\rho}_l^A - \tilde{\rho}_l^A\|_1\right] + \mathbb{E}[T_2]} \leq 2\epsilon + 2\sqrt{2\epsilon} \tag{C.3}$$

for all sufficiently large n and all sufficiently small $\eta, \delta > 0$. Similarly, $\mathbb{E}[T_3]$ can be upper bounded from (*Devetak*, 2005a, Lemma 1).

C.3 Proof of Lemma IV.10

Let $\Phi^{A_g A_p A_R L} \triangleq (I \otimes V)\Psi_{\rho_A^{\otimes n}}^{A_R A}(I \otimes V)^\dagger$. Note that,

$$\mathbb{E}_A\left[[F(\Phi^{A_p}, |0\rangle\langle 0|_{A_p})]\right] \geq \left(1 - \frac{1}{2}\mathbb{E}_A\left[\|\mathcal{N}_A(\rho_A) - |0\rangle\langle 0|_{A_p}\|_1\right]\right)^2 \geq 1 - \mathbb{E}_A\left[\|\mathcal{N}_A(\rho_A) - |0\rangle\langle 0|_{A_p}\|_1\right].$$

This means that there exists an $\omega^{A_g A_R L} \in \mathcal{D}(\mathcal{H}_{A_g} \otimes \mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_L)$ such that

$$F(\Phi^{A_p}, |0\rangle\langle 0|_{A_p}) = F(\Phi^{A_g A_p A_R L}, \omega^{A_g A_R L} \otimes |0\rangle\langle 0|_{A_p}),$$

which gives

$$\mathbb{E}_A\left[\|\Phi^{A_g A_p A_R L} - \omega^{A_g A_R L} \otimes |0\rangle\langle 0|_{A_p}\|_1\right] \leq 2\sqrt{\mathbb{E}_A\left[\|\mathcal{N}_A(\rho_A) - |0\rangle\langle 0|_{A_p}\|_1\right]}.$$

Using monotonicity of trace distance, we have

$$\mathbb{E}_A\left[\|\Phi^{A_g A_R L} - \omega^{A_g A_R L}\|_1\right] \leq 2\sqrt{\mathbb{E}_A\left[\|\mathcal{N}_A(\rho_A) - |0\rangle\langle 0|_{A_p}\|_1\right]} \tag{C.4}$$

Therefore, we have the result

$$\begin{aligned}
&\mathbb{E}_A\left[\|\Phi^{A_p A_g A_R L} - \Phi^{A_R A_g L} \otimes |0\rangle\langle 0|_{A_p}\|_1\right] \\
&\qquad \leq \mathbb{E}_A\left[\|\Phi^{A_p A_g A_R L} - \omega^{A_g A_R L} \otimes |0\rangle\langle 0|_{A_p}\|_1\right] + \mathbb{E}_A\left[\|\omega^{A_g A_R L} - \Phi^{A_g A_R L}\|_1\right] \\
&\qquad \leq 4\sqrt{\mathbb{E}_A\left[\|\mathcal{N}_A(\rho_A) - |0\rangle\langle 0|_{A_p}\|_1\right]}.
\end{aligned} \tag{C.5}$$

C.4 Proof of Proposition IV.14

We begin by defining $\mathcal{J}$ as

$$\mathcal{J} \triangleq \Big\{\exists(\tilde{l}, \tilde{k}, i, j) : (U^n(l), V^n(k)) \in \mathcal{B}_1(i) \times \mathcal{B}_2(j), (U^n(\tilde{l}), V^n(\tilde{k})) \in \mathcal{B}_1(i) \times \mathcal{B}_2(j),$$

$$(U^n(\tilde{l}), V^n(\tilde{k})) \in \mathcal{T}_\delta^{(n)}(U,V)\Big\}.$$

Using this, consider the following simplification:

$$\begin{aligned}
\|\xi^{T_p} - \xi_b^{T_p}\|_1 &\leq \sum_{l,k} \gamma_l \zeta_k \mathbb{1}_{\{(l,k)\neq d(l,k)\}} 2 \underbrace{\|\tilde{\Psi}_{\rho^{\otimes n}}\|_1}_{\leq 1} \\
&\leq 2 \sum_{l,k} \gamma_l \zeta_k \mathbb{1}_{\{\mathcal{J}\}} = \frac{2}{2^{n(\tilde{R}_1+\tilde{R}_2)}} \sum_{u^n, v^n} \sum_{l,k} \mathbb{1}_{\{U^n(l)=u^n, V^n(k)=v^n\}} \mathbb{1}_{\{\mathcal{J}\}}.
\end{aligned}$$

Note that, for every (u^n, v^n, l, k), we have

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

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,

$$\mathbb{E}\big[\|\xi^{T_p} - \xi_b^{T_p}\|_1\big] \leq \frac{10}{(1-\varepsilon)^2(1-\varepsilon')^2} 2^{-n(I(U;V)-2\delta_1)} 2^{n(\tilde{R}_1-R_1)} 2^{n(\tilde{R}_2-R_2)},$$

which completes the proof.

APPENDIX D

Proofs of Chapter V

D.1 Proof of Lemma V.4

Note that

$$|\psi_\rho\rangle \triangleq (I_R \otimes \sqrt{\rho^B})\,|\Gamma\rangle^{RB}, \quad |\psi_\sigma\rangle \triangleq (I_R \otimes \sqrt{\sigma^B})\,|\Gamma\rangle^{RB},$$

where $|\Gamma\rangle_{RB}$ is the unnormalized maximally entangled pure state: $|\Gamma\rangle_{RB} = \sum_i |i\rangle_R\,|i\rangle_B$. Consider the fidelity between the canonical purification states $|\psi_\rho\rangle$ and $|\psi_\sigma\rangle$:

$$\begin{aligned} F(|\psi_\rho\rangle, |\psi_\sigma\rangle) &\overset{a}{=} |\langle\psi_\rho|\psi_\sigma\rangle|^2 = |\langle\Gamma|^{RB}(I_R \otimes \sqrt{\rho^B}\sqrt{\sigma^B})\,|\Gamma\rangle^{RB}|^2, \\ &\overset{b}{=} |\operatorname{Tr}\Big(\sqrt{\rho^B}\sqrt{\sigma^B}\Big)|^2 \overset{c}{\geq} \left(1 - \frac{1}{2}\big\|\rho^B - \sigma^B\big\|_1\right)^2 \geq 1 - \big\|\rho^B - \sigma^B\big\|_1, \end{aligned}$$

where (a) follows from the definition of fidelity for a pure state, (b) follows from the definition of trace, (c) follows from the Power-Størmer inequality (*Powers and Størmer*, 1970, Lemma 4.1), i.e., for any positive semi-definite matrices A and B, we have

$$\operatorname{Tr}(A) + \operatorname{Tr}(B) - \|A - B\|_1 \leq 2\operatorname{Tr}\Big(\sqrt{A}\sqrt{B}\Big).$$

D.2 Proof of Lemma V.19

We first provide the following lemma.

Lemma D.1 (Covering superposition states)**.** *Consider a finite set $\mathcal{U}$, and a pair of collections $\{\rho_u\}_{u\in\mathcal{U}}$ and $\{\sigma_u\}_{u\in\mathcal{U}}$ where $\rho_u, \sigma_u \in \mathcal{D}(\mathcal{H}_A)$ for all $u \in \mathcal{U}$. Let $\{\Psi_u^\rho\}_{u\in\mathcal{U}}$ and $\{\Psi_u^\sigma\}_{u\in\mathcal{U}}$ acting on $\mathcal{D}(\mathcal{H}_{R_1}\otimes\mathcal{H}_A)$ and $\mathcal{D}(\mathcal{H}_{R_2}\otimes\mathcal{H}_A)$ be some purifications of $\{\rho_u\}_{u\in\mathcal{U}}$ and $\{\sigma_u\}_{u\in\mathcal{U}}$, respectively, with $\dim(\mathcal{H}_{R_1}) \le \dim(\mathcal{H}_{R_2})$. Then there exists a collection of isometric operators $\{U_r(u)\}_{u\in\mathcal{U}}$ acting on $\mathcal{H}_{R_1} \to \mathcal{H}_{R_2}$ and phases $\{\delta_u\}$ such that*

$$F((U_R\otimes I_A)\left|\tau_\rho\right\rangle, \left|\tau_\sigma\right\rangle) = F(\left|\tau_\rho\right\rangle, (U_R\otimes I_A)^\dagger\left|\tau_\sigma\right\rangle) \ge 1 - \sum_{u\in\mathcal{U}} \frac{1}{|\mathcal{U}|}\|\rho_u - \sigma_u\|_1, \tag{D.1}$$

where

$$U_R \triangleq \sum_{u\in\mathcal{U}} e^{-i\delta_u} U_r(u)\otimes|u\rangle\langle u|, \quad \left|\tau_\rho\right\rangle \triangleq \sum_{u\in\mathcal{U}} \frac{1}{\sqrt{|\mathcal{U}|}}\left|\psi_u^\rho\right\rangle\otimes|u\rangle, \quad \left|\tau_\sigma\right\rangle \triangleq \sum_{u\in\mathcal{U}} \frac{1}{\sqrt{|\mathcal{U}|}}\left|\psi_u^\sigma\right\rangle\otimes|u\rangle.$$

Proof. We provide a proof in Appendix D.3. □

Now, with the intention of employing the above lemma we perform the following identification. Identify $\mathcal{U}$ with $\mathcal{M}\times\mathcal{K}$, ρ_u with $\hat{\rho}_{m,k}^{B_R}$, σ_u with $\tilde{\rho}_{m,k}^{B_R}$, $\left|\psi_u^\rho\right\rangle$ with $\frac{(I\otimes M_{m,k})}{\sqrt{\lambda_{m,k}}}|\psi_\rho^{\otimes n}\rangle^{B_R^n B^n}$, and $\left|\psi_u^\sigma\right\rangle$ with $\frac{(I\otimes\sqrt{A_{m,k}})}{\sqrt{\delta_{m,k}}}|\psi_\rho^{\otimes n}\rangle^{B_R^n B^n}$. Note that the last two identifications are, in fact, the purifications of $\hat{\rho}_{m,k}^{B_R}$ and $\tilde{\rho}_{m,k}^{B_R}/\operatorname{Tr}\left(\tilde{\rho}_{m,k}^{B_R}\right)$, respectively as

$$\operatorname{Tr}_E\left(\frac{(I\otimes M_{m,k})}{\sqrt{\lambda_{m,k}}}\Psi_{\rho_B}^{\otimes n}\frac{(I\otimes M_{m,k}^\dagger)}{\sqrt{\lambda_{m,k}}}\right) = \hat{\rho}_{m,k}^{B_R}, \quad \operatorname{Tr}_B\left(\frac{(I\otimes\sqrt{A_{m,k}})}{\sqrt{\delta_{m,k}}}\Psi_{\rho_B}^{\otimes n}\frac{(I\otimes\sqrt{A_{m,k}})}{\sqrt{\delta_{m,k}}}\right) = \frac{\tilde{\rho}_{m,k}^{B_R}}{\operatorname{Tr}\left(\tilde{\rho}_{m,k}^{B_R}\right)}.$$

Using Lemma D.1, we obtain

$$\begin{aligned}
&F(|\hat{\sigma}\rangle^{B_R EMK}, (I_{B_R}\otimes U_{\mathcal{R}})|\tilde{\sigma}\rangle^{B_R BMK}) \\
&\quad\ge 1 - \frac{1}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}\sum_{m,k}\left\|\hat{\rho}_{m,k}^{B_R} - \frac{\tilde{\rho}_{m,k}^{B_R}}{\operatorname{Tr}\left(\tilde{\rho}_{m,k}^{B_R}\right)}\right\|_1 \\
&\quad\ge \frac{1}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}\sum_{m,k}\operatorname{Tr}\left\{\tilde{\rho}_{m,k}^{B_R}\right\} - \frac{1}{(1-\sqrt{\epsilon})|\mathcal{M}||\mathcal{K}|}\sum_{m,k}\left\|\hat{\rho}_{m,k}^{B_R} - \tilde{\rho}_{m,k}^{B_R}\right\|_1 \ge 1 - 4\sqrt{\epsilon},
\end{aligned} \tag{D.2}$$

where the last inequality follows from using the bounds in (5.29) and (5.30).

D.3 Proof of Lemma D.1

Consider the following:

$$\begin{aligned} F(|\tau_\rho\rangle, (U_R \otimes I)^\dagger |\tau_\sigma\rangle) &= \left| \sum_{u \in \mathcal{U}} \frac{1}{|\mathcal{U}|} e^{-i\delta_u} \langle \psi_u^\rho | U_r^\dagger(u) | \psi_u^\sigma \rangle \right|^2 \\ &\overset{a}{=} \left(\sum_{u \in \mathcal{U}} \frac{1}{|\mathcal{U}|} \left| \langle \psi_u^\rho | U_r^\dagger(u) | \psi_u^\sigma \rangle \right| \right)^2 \\ &\overset{b}{=} \left(\sum_{u \in \mathcal{U}} \frac{1}{|\mathcal{U}|} \sqrt{F(\rho_u, \sigma_u)} \right)^2 \\ &\overset{c}{\geq} \left(1 - \frac{1}{2} \sum_{u \in \mathcal{U}} \frac{1}{|\mathcal{U}|} \|\rho_u - \sigma_u\|_1 \right)^2 \geq 1 - \sum_{u \in \mathcal{U}} \frac{1}{|\mathcal{U}|} \|\rho_u - \sigma_u\|_1, \end{aligned}$$

where (a) follows by choosing δ_u such that $e^{-i\delta_u} \langle \psi_u^\rho | U_r^\dagger(u) | \psi_u^\sigma \rangle = | \langle \psi_u^\rho | U_r^\dagger(u) | \psi_u^\sigma \rangle |$, (b) follows from Uhlmann's theorem (*Wilde*, 2011, Theorem 9.2.1), i.e., there exists some isometry $U_r(u)$ such that $F(\rho_u, \sigma_u) = F(U_r(u) |\psi^\rho\rangle, |\psi^\sigma\rangle)$, and (c) follows from Lemma V.3.

D.4 Proof of Proposition V.20

We begin by defining indexing functions $f^{(m)} : [0, K'_m - 1] \to \mathcal{I}_{\mathcal{E}}^{(m)}$, for each $m \in \mathcal{M}'$, that uniquely map each element of the $[0, K'_m - 1]$ to the set $\mathcal{I}_{\mathcal{E}}^{(m)}$ in a monotonic fashion. Let $g^{(m)} : \mathcal{I}_{\mathcal{E}}^{(m)} \to [0, K'_m - 1]$ be the inverse of $f^{(m)}$, for each $m \in \mathcal{M}'$. Define the transformed vectors corresponding to the collections $\{\chi_k^{(m)}\}$ and $\{\phi_k^{(m)}\}$ as

$$|\hat{\chi}_s^{(m)}\rangle \triangleq c \sum_{j=0}^{K'_m - 1} e^{\frac{2\pi i j s}{K'_m}} |\chi_{f^{(m)}(j)}^{(m)}\rangle \quad \text{and} \quad |\hat{\phi}_s^{(m)}\rangle \triangleq c \sum_{j=0}^{K'_m - 1} e^{\frac{2\pi i j s}{K'_m}} |\phi_{f^{(m)}(j)}^{(m)}\rangle,$$

for $s \in [0, K'_m - 1]$. It follows from basic algebra that, for all $m \in \mathcal{M}'$,

$$\frac{1}{K'_m} \sum_{s=0}^{K'_m - 1} \langle \hat{\phi}_s^{(m)} | \hat{\chi}_s^{(m)} \rangle = c^2 \sum_{j=0}^{K'_m - 1} \langle \phi_{f^{(m)}(j)}^{(m)} | \chi_{f^{(m)}(j)}^{(m)} \rangle = c^2 \sum_{k \in \mathcal{I}_{\mathcal{E}}^{(m)}} \langle \phi_k^{(m)} | \chi_k^{(m)} \rangle = c^2 \sum_{k \in \mathcal{I}_{\mathcal{E}}^{(m)}} \operatorname{Tr}\left\{ \Xi_k^{(m)} \tau_k^{(m)} \right\},$$

where the last inequality follows from (5.47). Noting that the right hand side is a real number, we infer, for all $m \in \mathcal{M}'$, there exists at least one value of $s_m \in [0, K'_m - 1]$ that follows the inequality:

$$e^{i\hat{\theta}_m} \langle \hat{\phi}^{(m)}_{s_m} | \hat{\chi}^{(m)}_{s_m} \rangle \geq c^2 \sum_{k \in \mathcal{I}^{(m)}_{\mathcal{E}}} \langle \phi^{(m)}_k | \chi^{(m)}_k \rangle, \quad \text{for some phase } \hat{\theta}_m.$$

Observe that,

$$\langle \hat{\phi}^{(m)}_{s_m} | \hat{\chi}^{(m)}_{s_m} \rangle = c^2 \sum_{k \in \mathcal{I}^{(m)}_{\mathcal{E}}} \sum_{k' \in \mathcal{I}^{(m)}_{\mathcal{E}}} e^{\frac{2\pi i (g^{(m)}(k) - g^{(m)}(k'))s_m}{K'_m}} \langle \phi^{(m)}_{k'} | \chi^{(m)}_k \rangle,$$

for all $m \in \mathcal{M}'$. Choosing $\alpha^{(m)}_k = \frac{2\pi g^{(m)}(k) s_m}{K'_m}$ and $\beta^{(m)}_k = \frac{2\pi g^{(m)}(k) s_m}{K'_m} + \hat{\theta}_m$, we obtain

$$\frac{1}{M'} \sum_{m \in \mathcal{M}'} \langle \phi_m | \chi_m \rangle \geq \frac{c^2}{M'} \sum_{m \in \mathcal{M}'} \sum_{k \in \mathcal{I}^{(m)}_{\mathcal{E}}} \langle \phi^{(m)}_k | \chi^{(m)}_k \rangle = \frac{c^2}{M'} \sum_{m \in \mathcal{M}'} \sum_{k \in \mathcal{I}^{(m)}_{\mathcal{E}}} \mathrm{Tr}\left\{ \Xi^{(m)}_k \tau^{(m)}_k \right\} \geq 1 - 2\sqrt{\epsilon},$$

where the last inequality uses (5.31) and substitutes the value of c, which completes the proof.

D.5 Proof of Lemma V.22

Here we follow arguments similar to the proof of (*Cuff*, 2013, Lemma VI.5). We begin by defining $\mathcal{I}'_\epsilon$ (removing the relaxation in the rate) as, for all $\epsilon \geq 0$,

$$\mathcal{I}'_\epsilon \triangleq \left\{ R : \exists\, \rho^{A_R} \in \mathcal{S}_\epsilon(\rho^B, \mathcal{N}_W) \text{ such that } R \geq I_c(\mathcal{N}_W, \rho^{A_R}) \right\}, \tag{D.3}$$

and note from *Cuff* (2013) that

$$\bigcap_{\epsilon > 0} \mathcal{I}_\epsilon \subseteq \text{Closure} \left(\bigcap_{\epsilon > 0} \mathcal{I}'_\epsilon \right).$$

Now we prove the following:

$$\mathcal{S}_0(\rho^B, \mathcal{N}_W) = \bigcap_{\epsilon > 0} \mathcal{S}_\epsilon(\rho^B, \mathcal{N}_W). \tag{D.4}$$

$\mathcal{S}_0(\rho^B, \mathcal{N}_W) \subseteq \bigcap_{\epsilon>0} \mathcal{S}_\epsilon(\rho^B, \mathcal{N}_W)$ is straightforward. To show the other direction, consider any $\rho_1^{A_R} \in \bigcap_{\epsilon>0} \mathcal{S}_\epsilon(\rho^B, \mathcal{N}_W)$. This means, for all $\epsilon > 0$,

$$\|\mathcal{N}_W(\rho_1^{A_R}) - \rho^{B_R}\|_1 \leq \epsilon \implies \|\mathcal{N}_W(\rho_1^{A_R}) - \rho^{B_R}\|_1 = 0 \implies \mathcal{N}_W(\rho_1^{A_R}) = \rho^{B_R}, \tag{D.5}$$

where the second implication follows from the definition of a metric, and hence $\rho_1^{A_R} \in \mathcal{S}_0(\rho^B, \mathcal{N}_W)$ and (D.4) is true. Observe that, since the intersection of decreasing sequence of non-empty closed and bounded sets of a compact (finite-dimensional) metric space is non-empty, $\mathcal{S}_0(\rho^B, \mathcal{N}_W)$ is non-empty. Therefore, using the continuity of $f(\rho^{A_R}) = I_c(\mathcal{N}_W, \rho^{A_R})$, and the fact that $\mathcal{S}_\epsilon$ are decreasing non-empty closed and bounded subsets of a compact (finite-dimensional) metric space gives

$$f(\mathcal{S}_0(\rho^B, \mathcal{N}_W)) = \bigcap_{\epsilon>0} f(\mathcal{S}_\epsilon(\rho^B, \mathcal{N}_W)).$$

Noting that the images $f(\mathcal{S}_\epsilon(\rho^B, \mathcal{N}_W))$ and $f(\mathcal{S}_0(\rho^B, \mathcal{N}_W))$ characterize the rate regions $\mathcal{I}'_\epsilon$ and $\mathcal{I}_0$, respectively, and the fact that $\mathcal{I}_0$ is closed completes the proof.

APPENDIX E

Proofs of Chapter VI

E.1 Characterization of Certain High Probable Subspaces

In this appendix, we characterize certain high probability subspaces of tensor product quantum states. The statements we prove here are colloquially referred to as 'pinching' *Wilde* (2013a) in the literature. We prove statements in a form that can be used for use in the proof of both Theorems VI.3, Lemma E.1. We begin with definitions of typical and conditional typical projectors. We adopt strong (frequency) typicality. All statements hold for most of the variants of notion of typicality. For concreteness, the reader may refer to (*Pradhan et al.*, 2021, App. A).

Lemma E.1. *Suppose (i) $\mathcal{A}, \mathcal{B}$ are finite sets, (ii) p_{AB} is a PMF on $\mathcal{A}\times\mathcal{B}$, (iii) $(\rho_b \in \mathcal{D}(\mathcal{H}) : b \in \mathcal{B})$ is a collection of density operators, $\rho_a \triangleq \sum_{b\in\mathcal{B}} p_{B|A}(b|a)\rho_b$ for $a \in \mathcal{A}$ and $\rho = \sum_{a\in\mathcal{A}} p_A(a)\rho_a = \sum_{b\in\mathcal{B}} p_B(b)\rho_b$. There exists a strictly positive $\mu > 0$, whose value depends only on p_{AB}, such that for every $\delta > 0$, there exists a $N(\delta) \in \mathbb{N}$ such that for all $n \geq N(\delta)$, we have*

$$\mathrm{Tr}\Big(\Pi^{\delta}_{\rho}\Pi^{\delta}_{a^n}\Pi^{\delta}_{\rho}\rho_{b^n}\Big) \geq 1 - \exp\{-n\lambda\delta^2\}$$

whenever $(a^n, b^n) \in T^n_{\frac{\delta}{4}}(p_{AB})$ where $\Pi^{\delta}_{a^n}$ is the conditional typical projector of $\rho_{a^n} = \otimes_{t=1}^n \rho_{a_t}$ (Wilde, 2013a, Defn. 15.2.4) and Π^{δ}_{ρ} is the unconditional typical projector (Wilde, 2013a, Defn. 15.1.3) of $\rho^{\otimes n}$.

Proof. We rename $\mathcal{A} = \mathcal{V}$, $\mathcal{B} = \mathcal{X}$, $p_{AB} = p_{VX}$, a as v and b as x. We have

$$\begin{aligned} \mathrm{Tr}\Big(\Pi_\rho^\delta \Pi_{v^n}^\delta \Pi_\rho^\delta \rho_{x^n}\Big) &= \mathrm{Tr}\Big(\Pi_\rho^\delta \Pi_{v^n}^\delta \rho_{x^n} \Pi_\rho^\delta\Big) && \text{(E.1)} \\ &\geq \mathrm{Tr}\Big(\Pi_{v^n}^\delta \rho_{x^n}\Big) - \frac{1}{2}\left\|\rho_{x^n} - \Pi_\rho^\delta \rho_{x^n} \Pi_\rho^\delta\right\|. && \text{(E.2)} \end{aligned}$$

In the following we derive a lower bound on $\mathrm{Tr}\big(\Pi_{v^n}^\delta \rho_{x^n}\big)$ and derive an upper bound on $\left\|\rho_{x^n} - \Pi_\rho^\delta \rho_{x^n} \Pi_\rho^\delta\right\|$. Toward the deriving the former, we recall that we have $(v^n, x^n) \in {T_{\delta/2}}^n(p_{VX})$. Let us define:

$$p_{Y|XV}(y|x,v) := \left\langle e_{y|v}|\rho_x|e_{y|v}\middle|e_{y|v}|\rho_x|e_{y|v}\right\rangle, \tag{E.3}$$

for all $(x, v, y) \in \mathcal{X} \times \mathcal{V} \times \mathcal{Y}$.

Clearly, we have $p_{Y|XV}(y|x,v) \geq 0$, and

$$\sum_{y\in\mathcal{Y}} p_{Y|XV}(y|x,v) = \sum_{y\in\mathcal{Y}} \left\langle e_{y|v}|\rho_x|e_{y|v}\middle|e_{y|v}|\rho_x|e_{y|v}\right\rangle = \mathrm{Tr}(\rho_x) = 1.$$

Hence we see that $p_{Y|XV}$ is a stochastic matrix.

Next we note that

$$\begin{aligned} \sum_{x\in\mathcal{X}} p_{Y|XV}(y|x,v) p_{XV}(x,v) &= \sum_{x\in\mathcal{X}} p_{XV}(x,v) \left\langle e_{y|v}|\rho_x|e_{y|v}\middle|e_{y|v}|\rho_x|e_{y|v}\right\rangle \\ &= p_V(v) \left\langle e_{y|v}| \sum_{x\in\mathcal{X}} p_{X|V}(x|v)\rho_x|e_{y|v}\middle| e_{y|v}| \sum_{x\in\mathcal{X}} p_{X|V}(x|v)\rho_x|e_{y|v}\right\rangle \\ &= p_V(v) \left\langle e_{y|v}|\rho_v|e_{y|v}\middle|e_{y|v}|\rho_v|e_{y|v}\right\rangle = p_V(v) q_{Y|V}(y|v), \end{aligned} \tag{E.4}$$

where we have used the spectral decomposition of ρ_v.

Observe that if $(x^n, v^n) \in T_{\delta/4}^n(p_{XV} p_{Y|XV})$, and $y^n \in T_\delta^n(p_{XV} p_{Y|XV}|x^n, v^n)$, then we have $(x^n, v^n, y^n) \in T_\delta^n(p_{XV} p_{Y|XV})$. This implies that we have $(v^n, y^n) \in T_\delta^n(p_{VY})$, where p_{VY} is the marginal of $p_{XV} p_{Y|XV}$. Using this and (E.4), we see that $(v^n, y^n) \in T_\delta^n(p_V q_{Y|V})$. In summary, we see that if $(x^n, v^n) \in T_{\delta/4}(p_{XV})$, then we have

$$T_\delta^n(p_{XV} p_{Y|XV}|x^n, v^n) \subseteq \left\{y^n : (v^n, y^n) \in T_\delta^n(p_V q_{Y|V})\right\}.$$

We are now set to provide the promised lower bound. Consider

$$\operatorname{Tr}(\Pi_{v^n}\rho_{x^n}) = \operatorname{Tr}\left(\left[\sum_{y^n:(v^n,y^n)\in T^n_\delta(p_V q_{Y|V})} \bigotimes_{t=1}^n |e_{y_t|v_t}\rangle\langle e_{y_t|v_t}|\right]\left[\bigotimes_{j=1}^n \rho_{x_j}\right]\right) \tag{E.5}$$

$$= \operatorname{Tr}\left(\left[\sum_{y^n:(v^n,y^n)\in T^n_\delta(p_V q_{Y|V})} \bigotimes_{t=1}^n |e_{y_t|v_t}\rangle\langle e_{y_t|v_t}| \rho_{x_t}\right]\right) \tag{E.6}$$

$$= \sum_{y^n:(v^n,y^n)\in T^n_\delta(p_V q_{Y|V})} \prod_{t=1}^n \langle e_{y_t|v_t}|\rho_{x_t}|e_{y_t|v_t}|e_{y_t|v_t}|\rho_{x_t}|e_{y_t|v_t}\rangle \tag{E.7}$$

$$\geq \sum_{y^n\in T^n_\delta(p_{XV}p_{Y|XV}|x^n,v^n))} \prod_{t=1}^n p_{Y|XV}(y_t|x_t,v_t) \tag{E.8}$$

$$\geq 1 - 2|\mathcal{X}||\mathcal{Y}||\mathcal{V}|\exp\left\{-\frac{2n\delta^2 p_{XVY}(x^*,v^*,y^*)}{4(\log(|\mathcal{X}||\mathcal{Y}||\mathcal{V}|))^2}\right\}, \tag{E.9}$$

where we used the definition (E.3) in the last equality.

We next provide the upper bound. Note from the Gentle measurements lemma (*Wilde*, 2013a, Lemma 9.4.2), we have $\|\rho_{x^n} - \Pi^\delta_\rho \rho_{x^n} \Pi^\delta_\rho\| \leq 3\sqrt{\epsilon}$ if $\operatorname{Tr}\left(\Pi^\delta_\rho \rho_{x^n}\right) \geq 1-\epsilon$. In the following we provide a lower bound on $\operatorname{Tr}\left(\Pi^\delta_\rho \rho_{x^n}\right)$. Recall that $\Pi^\delta_\rho = \sum_{y^n\in T^n_\delta(s_Y)} \bigotimes_{t=1}^n |g_{y_t}\rangle\langle g_{y_t}|$, where

$$\rho = \sum_{y\in\mathcal{Y}} s_Y(y)\,|g_y\rangle\langle g_y|\,,$$

is the spectral decomposition of ρ, and $\rho = \sum_{x\in\mathcal{X}} p_X(x)\rho_x$. Let $\hat{p}_{Y|X}(y|x) := \langle g_y|\rho_x|g_y|g_y|\rho_x|g_y\rangle$, for all $(x,y)\in\mathcal{X}\times\mathcal{Y}$. Note that $\hat{p}_{Y|X}$ is not related to $p_{Y|X}$ defined previously. We note that $\hat{p}_{Y|X}(y|x)\geq 0$, and $\sum_{y\in\mathcal{Y}}\hat{p}_{Y|X}(y|x) = \sum_{y\in\mathcal{Y}}\langle g_y|\rho_x|g_y|g_y|\rho_x|g_y\rangle = \operatorname{Tr}(\rho_x) = 1$ for all $x\in\mathcal{X}$. Thus we see that $\hat{p}_{Y|X}$ is a stochastic matrix. It can also be noted that

$$\sum_{x\in\mathcal{X}} \hat{p}_{Y|X}(y|x)p_X(x) = \left\langle g_y|\sum_{x\in\mathcal{X}} p_X(x)\rho_x|g_y \middle| g_y|\sum_{x\in\mathcal{X}} p_X(x)\rho_x|g_y\right\rangle = \langle g_y|\rho|g_y|g_y|\rho|g_y\rangle = s_Y(y),$$

for all $y\in\mathcal{Y}$. This implies that the condition $y^n\in T^n_\delta(s_Y)$ is equivalent to the condition $y^n\in T^n_\delta(\hat{p_Y})$, where $\hat{p_Y}(y) = \sum_{x\in\mathcal{X}}\hat{p}_{Y|X}(y|x)p_X(x)$. Moreover, if $x^n\in T^n_{\delta/2}(p_X)$, and $y^n\in T^n_\delta(p_X\hat{p}_{Y|X}|x^n)$, then we have $(x^n,y^n)\in T^n_\delta(p_X\hat{p}_{Y|X})$. Consequently, we have $y^n\in T^n_\delta(\hat{p_Y})$, which in turn implies that $y^n\in T^n_\delta(s_Y)$. In essence, we have that if $x^n\in T^n_{\delta/2}(p_X)$ then $T^n_\delta(p_X\hat{p}_{Y|X}|x^n)\subseteq$

$T_\delta^n(s_Y)$. Now we are set to provide the lower bound on $\mathrm{Tr}\left(\Pi_\rho^\delta \rho_{x^n}\right)$ as follows:

$$\mathrm{Tr}\left(\Pi_\rho^\delta \rho_{x^n}\right) = \mathrm{Tr}\left(\sum_{y^n \in T_\delta(s_Y)} \bigotimes_{t=1}^n |g_{y_t}\rangle \langle g_{y_t}| \rho_{x_t}\right) = \sum_{y^n \in T_\delta(s_Y)} \prod_{t=1}^n \langle g_{y_t}|\rho_{x_t}|g_{y_t}|g_{y_t}|\rho_{x_t}|g_{y_t}\rangle \tag{E.10}$$

$$= \sum_{y^n \in T_\delta(s_Y)} \prod_{t=1}^n \hat{p}_{Y|X}(y_t|x_t) \geq \sum_{y^n \in T_\delta(\hat{p}_{Y|X} p_X|x^n)} \prod_{t=1}^n \hat{p}_{Y|X}(y_t|x_t) \tag{E.11}$$

$$\geq 1 - 2|\mathcal{X}||\mathcal{Y}| \exp\left\{-\frac{2n\delta^2 p_X^2(x^*)\hat{p}_{Y|X}^2(y|x)}{4(\log(|\mathcal{X}||\mathcal{Y}|))^2}\right\}. \tag{E.12}$$

We therefore have

$$\|\rho_{x^n} - \Pi_\rho^\delta \rho_{x^n} \Pi_\rho^\delta\| \leq 6|\mathcal{X}||\mathcal{Y}| \exp\left\{-\frac{2n\delta^2 p_X^2(x^*)\hat{p}_{Y|X}^2(y|x)}{4(\log(|\mathcal{X}||\mathcal{Y}|))^2}\right\},$$

and

$$\mathrm{Tr}(\Pi_{v^n} \rho_{x^n}) \geq 1 - 2|\mathcal{X}||\mathcal{Y}||\mathcal{V}| \frac{2n\delta^2 p_X^2(x^*)\hat{p}_{Y|X}^2(y|x)}{4(\log(|\mathcal{X}||\mathcal{Y}|))^2},$$

thereby permitting us to conclude that

$$\mathrm{Tr}\left(\Pi_\rho^\delta \Pi_{v^n}^\delta \Pi_\rho^\delta \rho_{x^n}\right) \geq \mathrm{Tr}\left(\Pi_{v^n}^\delta \rho_{x^n}\right) - \frac{1}{2}\|\rho_{x^n} - \Pi_\rho^\delta \rho_{x^n} \Pi_\rho^\delta\| \geq 1 - \frac{2n\delta^2 p_X^2(x^*)\hat{p}_{Y|X}^2(y|x)}{4(\log(|\mathcal{X}||\mathcal{Y}|))^2},$$

if $(x^n, v^n) \in T_{\delta/2}^n(p_{XV})$. □

E.2 Proof of Proposition VI.8

We begin by defining the following events:

$$\mathcal{J} \triangleq \left\{\left(V_1^n(A_{1.m_1}, m_1), X_1^n(m_1), V_2^n(A_{2.m_2}, m_2), X_2^n(m_2)\right) = (v_1^n, x_1^n, v_2, x_2) \in T_{8\delta}(p_{V_1X_1V_2X_2})\right\},$$

$$\mathcal{V} \triangleq \{V_j^n(a_j, m_j) = v_j^n : j \in [2]\}, \quad \hat{\mathcal{V}} \triangleq \{V^n(\hat{a}, m_1 \oplus m_2) = \hat{v}^n\}, \quad \mathcal{A} \triangleq \{A_{j,m_j} = a_j : j \in [2]\},$$

This gives,

$$\mathbb{E}_{\mathcal{P}}[T_{22}] = \mathbb{E}_{\mathcal{P}}\left[4 \sum_{\underline{m}} \sum_{a_1,a_2} \sum_{\hat{a} \neq \underline{a}^\oplus} \sum_{\substack{(\underline{v}^n, \underline{x}) \in \\ T_{8\delta}(p_{\underline{V}\,\underline{X}})}} p_{\underline{M}}(\underline{m}) \,\mathrm{Tr}\left(\Gamma_{\hat{a},\underline{m}^\oplus} \rho_{m_1m_2}^{\otimes n}\right) \mathbb{1}_{\mathcal{J}} \mathbb{1}_{\mathcal{A}}\right],$$

$$
\begin{aligned}
&\overset{(a)}{=} 4 \sum_{\underline{m}} \sum_{a_1,a_2} \sum_{\hat{a}\neq \underline{a}^{\oplus}} \sum_{\substack{(\underline{v}^n,\underline{x})\in \\ T_{8\delta}(p_{\underline{V}\underline{X}})}} \sum_{\hat{v}^n\in\mathcal{V}^n} p_{\underline{M}}(\underline{m}) \operatorname{Tr}\Big(\pi_{\hat{v}^n}\pi_\rho \rho^{\otimes n}_{x_1^n x_2^n}\pi_\rho\Big)\mathbb{E}_{\mathcal{P}}\left\{\mathbb{1}_{\mathcal{J}}\mathbb{1}_{\mathcal{A}}\mathbb{1}_{\hat{\mathcal{V}}}\right\} \\
&\overset{(b)}{\leq} 4 \sum_{\underline{m}} \sum_{a_1,a_2} \sum_{\hat{a}\neq \underline{a}^{\oplus}} \sum_{\substack{(\underline{v}^n)\in \\ T_{8\delta}(p_{\underline{V}})}} \sum_{\underline{x}^n\in\underline{\mathcal{X}}^n} \sum_{\hat{v}^n\in\mathcal{V}^n} p_{\underline{M}}(\underline{m}) \left[\prod_{j=1}^{2} p_{X_j|V_j}(x_j^n|v_j^n)\right] \operatorname{Tr}\Big(\pi_{\hat{v}^n}\pi_\rho \rho^{\otimes n}_{x_1^n x_2^n}\pi_\rho\Big)\mathcal{P}(\mathcal{V},\mathcal{A},\hat{\mathcal{V}}) \\
&\overset{(c)}{=} 4 \sum_{\underline{m}} \sum_{a_1,a_2} \sum_{\hat{a}\neq \underline{a}^{\oplus}} \sum_{\substack{(\underline{v}^n)\in \\ T_{8\delta}(p_{\underline{V}})}} \sum_{\hat{v}^n\in\mathcal{V}^n} p_{\underline{M}}(\underline{m}) \operatorname{Tr}\Big(\pi_{\hat{v}^n}\pi_\rho \rho^{\otimes n}_{v_1^n v_2^n}\pi_\rho\Big)\mathcal{P}(\mathcal{V},\mathcal{A},\hat{\mathcal{V}}) \\
&\overset{(d)}{\leq} 4 \sum_{\underline{m}} \sum_{a_1,a_2} \sum_{\hat{a}\neq \underline{a}^{\oplus}} \sum_{\substack{(\underline{v}^n)\in \\ T_{8\delta}(p_{\underline{V}})}} \sum_{\hat{v}^n\in\mathcal{V}^n} p_{\underline{M}}(\underline{m}) \operatorname{Tr}\Big(\pi_{\hat{v}^n}\pi_\rho \rho^{\otimes n}_{v_1^n v_2^n}\pi_\rho\Big)\mathcal{P}(\mathcal{V},\hat{\mathcal{V}}) \\
&\leq 4 \sum_{\underline{m}} \sum_{a_1,a_2} \sum_{\hat{a}\neq \underline{a}^{\oplus}} \sum_{\substack{(\underline{v}^n)\in \\ T_{8\delta}(p_{\underline{V}})}} \sum_{\hat{v}^n\in\mathcal{V}^n} p_{\underline{M}}(\underline{m}) \operatorname{Tr}\Big(\pi_{\hat{v}^n}\pi_\rho \rho^{\otimes n}_{v_1^n v_2^n}\pi_\rho\Big)\frac{1}{q^{3n}} \\
&\overset{(e)}{\leq} 4 \sum_{\substack{\underline{m}, \\ a_1,a_2}} \sum_{\hat{a}} \sum_{\hat{v}^n\in T_{8\delta}(V)} p_{\underline{M}}(\underline{m}) \operatorname{Tr}(\pi_{\hat{v}^n}\pi_\rho)\frac{2^{-n(S(\rho)-H(V_1,V_2)-\delta_1)}}{q^{3n}} \\
&\overset{(f)}{\leq} 4 \sum_{\substack{\underline{m}, \\ a_1,a_2}} \sum_{\hat{a}} p_{\underline{M}}(\underline{m}) \frac{\exp\left\{-n\left[S(\rho)-H(V_1,V_2)-\delta_1-\sum_v p_V(v)S(\rho_v)-H(V)\right]\right\}}{q^{3n}} \\
&\overset{(g)}{\leq} 4\exp\left\{-n\left\{\left[\log q-\left(\sum_v p_V(v)S(\rho_v)+H(V)-S(\rho)\right)\right]+2\log q-H(V_1,V_2)-\frac{3k}{n}\log q-\delta_1\right\}\right\},
\end{aligned}
\tag{E.13}
$$

and (a) follows from a summing over possible choices for $V^n(\hat{a}, m_1 \oplus m_2)$, (b) follows from evaluating the expectation, enlarging the summation range of x_1^n, x_2^n and substituting the distribution of the random code, (c) follows from the definitions of $\rho_{v_1 v_2} : \underline{v} \in \underline{\mathcal{V}}$, (d) follows as an upper bound since one of the events has been enlarged, (e) follows from (*Padakandla*, 2014, Lemma N.0.21c) and the operator inequality $\sum_{x^n\in T_\delta(p_V)} \pi_\rho \rho_{x^n} \pi_\rho \leq 2^{n(H(p_V)+\delta_1(\delta))}\pi_\rho \rho^{\otimes n}\pi_\rho \leq 2^{n(H(p_V)+\delta_1(\delta)-S(\rho))}\pi_\rho$ found in (*Wilde*, 2013a, Eqn. 20.34, 15.20), and from the definition of $\pi_{\hat{v}^n}$ which is the 0 projector if $\hat{v}^n$ is not typical with respect to p_V, (f) follows from $\pi_\rho \leq I$ and (*Wilde*, 2013a, Eqn. 15.77), and finally (g) follows by collating all the bounds.

We now analyze T_{23}.

$$\mathbb{E}_{\mathcal{P}}[T_{23}] = \mathbb{E}_{\mathcal{P}}\left\{ 4 \sum_{\substack{\underline{m}\in\mathcal{V}^{2l} \\ \hat{m}\neq m_1\oplus m_2}} \sum_{\substack{\underline{a}\in\mathcal{V}^{2k} \\ \hat{a}\in\mathcal{V}^k}} \sum_{\substack{(\underline{v}^n,\underline{x})\in \\ T_{8\delta}(p_{\underline{V}\underline{X}})}} p_{\underline{M}}(\underline{m}) \operatorname{Tr}\big(\Gamma_{\hat{a},\underline{m}^{\oplus}} \rho_{m_1m_2}^{\otimes n}\big) \mathbb{1}_{\mathcal{J}} \mathbb{1}_{\mathcal{A}} \right\}$$

$$= 4 \sum_{\substack{\underline{m}\in\mathcal{V}^{2l} \\ \hat{m}\neq m_1\oplus m_2}} \sum_{\substack{\underline{a}\in\mathcal{V}^{2k} \\ \hat{a}\in\mathcal{V}^k}} \sum_{\substack{(\underline{v}^n,\underline{x})\in \\ T_{8\delta}(p_{\underline{V}\underline{X}})}} \sum_{\hat{v}^n\in\mathcal{V}^n} p_{\underline{M}}(\underline{m}) \operatorname{Tr}\Big(\pi_{\hat{v}^n} \pi_\rho \rho_{x_1^n x_2^n}^{\otimes n} \pi_\rho\Big) \mathbb{E}_{\mathcal{P}}\left\{\mathbb{1}_{\mathcal{J}} \mathbb{1}_{\mathcal{A}} \mathbb{1}_{\hat{\mathcal{V}}}\right\}$$

$$\leq 4 \sum_{\substack{\underline{m}\in\mathcal{V}^{2l} \\ \hat{m}\neq m_1\oplus m_2}} \sum_{\substack{\underline{a}\in\mathcal{V}^{2k} \\ \hat{a}\in\mathcal{V}^k}} \sum_{\substack{(\underline{v}^n)\in \\ T_{8\delta}(p_{\underline{V}})}} \sum_{\underline{x}^n\in\underline{\mathcal{X}}^n} \sum_{\hat{v}^n\in\mathcal{V}^n} p_{\underline{M}}(\underline{m}) \left[\prod_{j=1}^{2} p_{X_j|V_j}(x_j^n|v_j^n)\right] \operatorname{Tr}\Big(\pi_{\hat{v}^n} \pi_\rho \rho_{x_1^n x_2^n}^{\otimes n} \pi_\rho\Big) \mathcal{P}(\mathcal{V},\mathcal{A},\hat{\mathcal{V}})$$

$$= 4 \sum_{\substack{\underline{m}\in\mathcal{V}^{2l} \\ \hat{m}\neq m_1\oplus m_2}} \sum_{\substack{\underline{a}\in\mathcal{V}^{2k} \\ \hat{a}\in\mathcal{V}^k}} \sum_{\substack{(\underline{v}^n)\in \\ T_{8\delta}(p_{\underline{V}})}} \sum_{\hat{v}^n\in\mathcal{V}^n} p_{\underline{M}}(\underline{m}) \operatorname{Tr}\Big(\pi_{\hat{v}^n} \pi_\rho \rho_{v_1^n v_2^n}^{\otimes n} \pi_\rho\Big) \mathcal{P}(\mathcal{V},\mathcal{A},\hat{\mathcal{V}})$$

$$\leq 4 \sum_{\substack{\underline{m}\in\mathcal{V}^{2l} \\ \hat{m}\neq m_1\oplus m_2}} \sum_{\substack{\underline{a}\in\mathcal{V}^{2k} \\ \hat{a}\in\mathcal{V}^k}} \sum_{\substack{(\underline{v}^n)\in \\ T_{8\delta}(p_{\underline{V}})}} \sum_{\hat{v}^n\in\mathcal{V}^n} p_{\underline{M}}(\underline{m}) \operatorname{Tr}\Big(\pi_{\hat{v}^n} \pi_\rho \rho_{v_1^n v_2^n}^{\otimes n} \pi_\rho\Big) \mathcal{P}(\mathcal{V},\mathcal{A},\hat{\mathcal{V}})$$

$$\leq 4 \sum_{\substack{\underline{m}\in\mathcal{V}^{2l} \\ \hat{m}\neq m_1\oplus m_2}} \sum_{\substack{\underline{a}\in\mathcal{V}^{2k} \\ \hat{a}\in\mathcal{V}^k}} \sum_{\substack{(\underline{v}^n)\in \\ T_{8\delta}(p_{\underline{V}})}} \sum_{\hat{v}^n\in\mathcal{V}^n} p_{\underline{M}}(\underline{m}) \operatorname{Tr}\Big(\pi_{\hat{v}^n} \pi_\rho \rho_{v_1^n v_2^n}^{\otimes n} \pi_\rho\Big) \frac{1}{q^{3n}}$$

$$\leq 4 \sum_{\substack{\underline{m}\in\mathcal{V}^{2l} \\ \hat{m}\neq m_1\oplus m_2}} \sum_{\substack{\underline{a}\in\mathcal{V}^{2k} \\ \hat{a}\in\mathcal{V}^k}} \sum_{\hat{v}^n\in\mathcal{V}^n} p_{\underline{M}}(\underline{m}) \operatorname{Tr}(\pi_{\hat{v}^n}\pi_\rho) \frac{2^{-n(S(\rho)-H(V_1,V_2)-\delta_1)}}{q^{3n}}$$

$$= 4 \sum_{\substack{\underline{m}, \\ a_1,a_2}} \sum_{\hat{a}} \sum_{\hat{v}^n\in T_\delta(V_1\oplus V_2)} p_{\underline{M}}(\underline{m}) \operatorname{Tr}(\pi_{\hat{v}^n}\pi_\rho) \frac{2^{-n(S(\rho)-H(V_1,V_2)-\delta_1)}}{q^{3n}}$$

$$\leq 4 \sum_{\substack{\underline{m}, \\ a_1,a_2}} \sum_{\hat{a}} p_{\underline{M}}(\underline{m}) \frac{\exp\left\{-n\left[S(\rho) - H(V_1,V_2) - \delta_1 - \sum_v p_V(v) S(\rho_v) - H(V)\right]\right\}}{q^{3n}}$$

$$\leq 4 \exp\left\{-n\left\{\log q - \sum_v p_V(v) S(\rho_v) - H(V) + S(\rho) + 2\log q - H(V_1,V_2) - \frac{(3k+l)\log q}{n} - \delta_1\right\}\right\}.$$

The above sequence of steps are analogous to those used in deriving an upper bound on T_{22} and follow from the same set of arguments as provided for the bounds in (E.13). This completes the proof of the claimed statement.

APPENDIX F

Proofs of Chapter VII

F.1 Proof of Proposition VII.12

We begin by defining the sets $\mathcal{J}$ and $\mathcal{K}$ as

$$\mathcal{J} \triangleq \left\{ \begin{array}{l} V_2^n(a_2, m_2) = v_2^n, \alpha_2(m_2) = a_2, \Theta_1(m_1) > 0 \\ V_3^n(a_3, m_3) = v_3^n, \alpha_3(m_3) = a_3, \Theta_2(m_2) > 0 \end{array} \right\} \subseteq \mathcal{K} \triangleq \{V_2^n(a_2, m_2) = v_2^n, V_3^n(a_3, m_3) = v_3^n\} \tag{F.1}$$

Now we simplify $\rho_{c,\underline{m}}^{Y_1}$ as

$$\begin{aligned} \rho_{c,\underline{m}}^{Y_1} &= \sum_{v_2^n, v_3^n \in \mathcal{F}_q^n} \sum_{x_1^n \in \mathcal{X}_1^n} \sum_{x_2^n, x_3^n \in \mathcal{X}_2^n \otimes \mathcal{X}_2^n} p_{X_2|V_2}(x_2^n|v_2^n) p_{X_3|V_3}^n(x_3^n|v_3^n) \rho_{x_1^n x_2^n x_3^n}^{Y_1} \\ &\qquad \times \mathbb{1}_{\{x_1^n(m_1)=x_1^n, v_2^n(\alpha_2(m_2),m_2)=v_2^n, v_3^n(\alpha_3(m_3),m_3)=v_3^n\}} \\ &= \sum_{v_2^n, v_3^n \in \mathcal{F}_q^n} \sum_{x_1^n \in \mathcal{X}_1^n} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \mathbb{1}_{\{x_1^n(m_1)=x_1^n, v_2^n(\alpha_2(m_2),m_2)=v_2^n, v_3^n(\alpha_3(m_3),m_3)=v_3^n\}} \\ &= \sum_{x_1^n \in \mathcal{X}_1^n} \sum_{v_2^n, v_3^n \in \mathcal{F}_q^n} \sum_{a_2, a_3 \in \mathcal{F}_q^k} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \mathbb{1}_{\{x_1^n(m_1)=x_1^n\}} \mathbb{1}_{\mathcal{J}}, \end{aligned} \tag{F.2}$$

where the above two equalities are based on the encoding rules employed by the encoders and the last one follows from the definition of $\mathcal{J}$. Using the above simplification in the term T_{21}, we get

$$\begin{aligned} T_{21}(\underline{m}) &= \sum_{m_1' \neq m_1} \sum_{x_1^n \in \mathcal{X}_1^n} \sum_{v_2^n, v_3^n \in \mathcal{F}_q^n} \sum_{a_2, a_3 \in \mathcal{F}_q^k} \mathrm{Tr}\Big(\gamma_{m_1'}^{a,l} \pi_l^a \rho_{x_1^n v_2^n v_3^n}^{Y_1} \pi_l^a\Big) \mathbb{1}_{\{x_1^n(m_1)=x_1^n\}} \mathbb{1}_{\mathcal{J}}, \\ &= \sum_{m_1' \neq m_1} \sum_{x_1^n, \hat{x}_1^n \in \mathcal{X}_1^n} \sum_{v_2^n, v_3^n \in \mathcal{F}_q^n} \sum_{a_2, a_3 \in \mathcal{F}_q^k} \sum_{u^n \in \mathcal{F}_q^n} \mathrm{Tr}\Big(\gamma_{\hat{x}_1^n}^{u^n} \pi_{u^n} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \pi_{u^n}\Big) \\ &\qquad \times \mathbb{1}_{\{x_1^n(m_1)=x_1^n, x_1^n(m_1')=\hat{x}_1^n\}} \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}} \mathbb{1}_{\mathcal{J}}, \end{aligned}$$

where $\gamma_{\hat{x}_1^n}^{u^n}$ is defined as $\gamma_{\hat{x}_1^n}^{u^n} \triangleq \pi_\rho \pi_{\hat{x}_1^n} \pi_{\hat{x}_1^n, u^n} \pi_{\hat{x}_1^n} \pi_\rho$ Taking expectation of the above term, we obtain

$$\begin{aligned} \mathbb{E}[T_{21}] &= \sum_{m_1' \neq m_1} \sum_{x_1^n, \hat{x}_1^n \in \mathcal{X}_1^n} \sum_{\substack{v_2^n, v_3^n \in \mathcal{F}_q^n \\ a_2, a_3 \in \mathcal{F}_q^k}} \sum_{u^n \in \mathcal{F}_q^n} \mathrm{Tr}\Big(\gamma_{\hat{x}_1^n}^{u^n} \pi_{u^n} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \pi_{u^n}\Big) \\ &\qquad \times \mathbb{P}(x_1^n(m_1) = x_1^n, x_1^n(m_1') = \hat{x}_1^n) \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}} \mathbb{P}(\mathcal{J}) \\ &\overset{(a)}{\leq} \sum_{m_1' \neq m_1} \sum_{x_1^n, \hat{x}_1^n \in \mathcal{X}_1^n} \sum_{a_2, a_3 \in \mathcal{F}_q^k} \sum_{\substack{v_2^n \in \mathcal{T}_\delta^{(n)}(V_2), \\ v_3^n \in \mathcal{T}_\delta^{(n)}(V_3)}} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \mathrm{Tr}\Big(\gamma_{\hat{x}_1^n}^{u^n} \pi_{u^n} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \pi_{u^n}\Big) \\ &\qquad \times p_{X_1}^n(x_1^n) p_{X_1}^n(\hat{x}_1^n) \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}} \mathbb{P}(\mathcal{K}) \\ &\overset{(b)}{\leq} \frac{2^{nR_1} q^{2k}}{q^{2n}} \sum_{x_1^n, \hat{x}_1^n \in \mathcal{X}_1^n} \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \mathrm{Tr}\left(\gamma_{\hat{x}_1^n}^{u^n} \pi_{u^n} \left(\sum_{v_2^n \in \mathcal{T}_\delta^{(n)}(V_2), v_3^n \in \mathcal{T}_\delta^{(n)}(V_3)} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \right) \pi_{u^n} \right) \\ &\qquad \times p_{X_1}^n(x_1^n) p_{X_1}^n(\hat{x}_1^n) \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}} \\ &\overset{(c)}{\leq} \frac{2^{nR_1} q^{2k}}{q^{2n}} 2^{n(H(V_2,V_3|U)+\delta_{u_2})} \sum_{\hat{x}_1^n \in \mathcal{X}_1^n} p_{X_1}^n(\hat{x}_1^n) \\ &\qquad \times \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \mathrm{Tr}\left(\gamma_{\hat{x}_1^n}^{u^n} \pi_{u^n} \left(\sum_{x_1^n \in \mathcal{X}_1^n} p_{X_1}^n(x_1^n) \rho_{x_1 u^n}^{Y_1} \right) \pi_{u^n} \right) \\ &\overset{(d)}{=} \frac{2^{nR_1} q^{2k}}{q^{2n}} 2^{n(H(V_2,V_3|U)+\delta_{u_2})} \sum_{\hat{x}_1^n \in \mathcal{X}_1^n} p_{X_1}^n(\hat{x}_1^n) \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \mathrm{Tr}\Big(\gamma_{\hat{x}_1^n}^{u^n} \pi_{u^n} \rho_{u^n}^{Y_1} \pi_{u^n}\Big) \\ &\overset{(e)}{\leq} \frac{2^{nR_1} q^{2k}}{q^{2n}} 2^{n(H(V_2,V_3|U)+\delta_{u_2})} 2^{-n(S(Y_1|U)_{\sigma_1}-\delta_{u_3})} \sum_{\hat{x}_1^n \in \mathcal{X}_1^n} p_{X_1}^n(\hat{x}_1^n) \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \\ &\qquad \times \mathrm{Tr}\big(\pi_\rho \pi_{\hat{x}_1^n} \pi_{\hat{x}_1^n, u^n} \pi_{\hat{x}_1^n} \pi_\rho \pi_{u^n}\big) \end{aligned}$$

$$\overset{\text{(f)}}{\leq} \frac{2^{nR_1}q^{2k}}{q^{2n}} 2^{n(H(V_2,V_3|U)+\delta_{u_2})} 2^{-n(S(Y_1|U)_{\sigma_1}-\delta_{u_3})} \sum_{\hat{x}_1^n \in \mathcal{X}_1^n} p_{X_1}^n(\hat{x}_1^n) \sum_{u^n \in \mathcal{T}_\delta^{(n)}(U)} \mathrm{Tr}\big(\pi_{\hat{x}_1^n,u^n}\big), \quad \text{(F.3)}$$

where (a) follows from the presence of indicators $\theta_1(m_1) > 0$, $\theta_2(m_2) > 0$, and definition of $\mathcal{K}$, (b) follows by observing that $\mathbb{P}(\mathcal{K}) = \frac{1}{q^n q^n}$, and (c) follows by using the following arguments for any $u^n \in \mathcal{T}_\delta^{(n)}(U)$,

$$\begin{aligned}
&\sum_{\substack{v_2^n \in \mathcal{T}_\delta^{(n)}(V_2),\\ v_3^n \in \mathcal{T}_\delta^{(n)}(V_3)}} \rho^{Y_1}_{x_1^n v_2^n v_3^n} \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}} \\
&\qquad \leq 2^{n(H(V_2,V_3|U)+\delta_{u_2})} \sum_{\substack{v_2^n \in \mathcal{T}_\delta^{(n)}(V_2),\\ v_3^n \in \mathcal{T}_\delta^{(n)}(V_3)}} p^n_{V_2V_3|U}(v_2^n, v_3^n|u^n) \rho^{Y_1}_{x_1^n v_2^n v_3^n} \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}} \\
&\qquad \leq 2^{n(H(V_2,V_3|U)+\delta_{u_2})} \sum_{v_2^n, v_3^n} p^n_{V_2V_3|U}(v_2^n, v_3^n|u^n) \rho^{Y_1}_{x_1^n v_2^n v_3^n} \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}} \\
&\qquad = 2^{n(H(V_2,V_3|U)+\delta_{u_2})} \rho^{Y_1}_{x_1^n u^n}. \qquad \text{(F.4)}
\end{aligned}$$

The equality in (d) follows from the definition of $\rho^{Y_1}_{u^n} \triangleq \sum_{x^n \in \mathcal{X}^n} p^n_{X_1}(x_1^n) \rho^{Y_1}_{x_1^n u^n}$, and the inequality in (e) uses the property of conditional projector i.e., $\pi_{u^n} \rho^{Y_1}_{u^n} \pi_{u^n} \leq 2^{-n(S(Y_1|U)_{\sigma_1}-\delta_{u_3})} \pi_{u^n}$, for σ_1 as defined in the statement of the theorem. Finally (f) follows using the cyclicity of trace and the fact that $\pi_{\hat{x}_1^n} \pi_\rho \pi_{u^n} \pi_\rho \pi_{\hat{x}_1^n} \leq \pi_{\hat{x}_1^n} \pi_\rho \pi_{\hat{x}_1^n} \leq \pi_{\hat{x}_1^n} \leq I$.

Using the dimensional bound of a conditional typical projector i.e., $\mathrm{Tr}\big(\pi_{\hat{x}_1^n,u^n}\big) \leq 2^{n(S(Y_1|X_1,U)+\delta_{u_4})}$ in (F.3) we obtain

$$\begin{aligned}
\mathbb{E}[T_{21}] &\leq \frac{2^{nR_1}q^{2k}}{q^{2n}} 2^{n(H(V_2,V_3|U)+\delta_{u_2})} 2^{-n(S(Y_1|U)_{\sigma_1}-\delta_{u_3})} 2^{n(S(Y_1|X_1,U)+\delta_{u_4})} 2^{n(H(U)+\delta_{u_1})} \\
&= \frac{2^{nR_1}q^{2k}}{q^{2n}} 2^{n(H(V_2,V_3)+\delta_{u_2})} 2^{-n(I(Y_1;X_1|U)_{\sigma_1}-\delta_{u_3})} 2^{n\delta_{u_4}} 2^{n\delta_{u_1}}.
\end{aligned}$$

This completes the proof.

F.2 Proof of Proposition VII.13

We begin by substituting the simplification performed in (F.2) into the expression corresponding to T_{22}. This gives

$$\begin{aligned}
T_{22} &= \sum_{a'\neq a, l'\neq l} \mathrm{Tr}\Big(\gamma_{m_1}^{a',l'} \pi_l^a \rho_{c,\underline{m}}^{Y_1} \pi_l^a\Big) \\
&= \sum_{a'\neq a, l'\neq l} \sum_{x_1^n\in\mathcal{X}_1^n} \sum_{v_2^n, v_3^n \in \mathcal{F}_q^n} \sum_{a_2,a_3\in\mathcal{F}_q^k} \mathrm{Tr}\Big(\gamma_{m_1}^{a',l'} \pi_l^a \rho_{x_1^n v_2^n v_3^n}^{Y_1} \pi_l^a\Big) \mathbb{1}_{\{x_1^n(m_1)=x_1^n\}} \mathbb{1}_{\mathcal{J}} \\
&\leq \sum_{\substack{a'\neq a,\\ l'\neq l}} \sum_{x_1^n\in\mathcal{X}_1^n} \sum_{\substack{v_2^n\in\mathcal{T}_\delta^{(n)}(V_2),\\ v_3^n\in\mathcal{T}_\delta^{(n)}(V_3)}} \sum_{a_2,a_3\in\mathcal{F}_q^k} \sum_{u^n,\hat{u}^n\in\mathcal{F}_q^n} \mathrm{Tr}\Big(\gamma_{x_1^n}^{\hat{u}^n} \pi_{u^n} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \pi_{u^n}\Big) \mathbb{1}_{\{x_1^n(m_1)=x_1^n\}} \\
&\qquad\qquad \mathbb{1}_{\{u^n=v_2^n\oplus v_3^n\}} \mathbb{1}_{\{\hat{u}^n=a'g_I+l'g_{O/I}+b_2^n+b_3^n\}} \mathbb{1}_{\mathcal{K}}
\end{aligned}$$

Taking expectation over the codebook generation distribution gives

$$\begin{aligned}
&\mathbb{E}[T_{22}] \\
&\leq \sum_{a'\neq a,l'\neq l} \sum_{x_1^n\in\mathcal{X}_1^n} \sum_{\substack{v_2^n\in\mathcal{T}_\delta^{(n)}(V_2),\\ v_3^n\in\mathcal{T}_\delta^{(n)}(V_3)}} \sum_{a_2,a_3\in\mathcal{F}_q^k} \sum_{u^n,\hat{u}^n\in\mathcal{F}_q^n} \mathrm{Tr}\Big(\gamma_{x_1^n}^{\hat{u}^n} \pi_{u^n} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \pi_{u^n}\Big) p_{X_1}^n(x_1^n) \mathbb{1}_{\{u^n=v_2^n\oplus v_3^n\}} \\
&\quad \times \mathbb{P}(\hat{u}^n = a'G_I + l'G_{O/I} + B_2^n + B_3^n, v_2^n = a_2G_I + m_2G_{O/I} + B_2^n, v_3^n = a_3G_I + m_3G_{O/I} + B_3^n) \\
&\overset{(a)}{\leq} \frac{q^{3k}q^l}{q^{3n}} \sum_{x_1^n\in\mathcal{X}_1^n} p_{X_1}^n(x_1^n) \sum_{\substack{v_2^n\in\mathcal{T}_\delta^{(n)}(V_2),\\ v_3^n\in\mathcal{T}_\delta^{(n)}(V_3)}} \sum_{u^n\in\mathcal{F}_q^n} \mathrm{Tr}\left(\Big(\sum_{\hat{u}^n\in\mathcal{F}_q^n} \gamma_{x_1^n}^{\hat{u}^n}\Big) \pi_{u^n} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \pi_{u^n}\right) \mathbb{1}_{\{u^n=v_2^n\oplus v_3^n\}}, \qquad \text{(F.5)}
\end{aligned}$$

where the inequality (a) is obtained by noting that for $a' \neq a_2 \oplus a_3$ and $l' \neq m_1 \oplus m_2$, we have

$$\mathbb{P}(\hat{u}^n = a'G_I + l'G_{O/I} + B_2^n + B_3^n, v_2^n = a_2G_I + m_2G_{O/I} + B_2^n, v_3^n = a_3G_I + m_3G_{O/I} + B_3^n) = \frac{1}{q^{3n}}. \qquad \text{(F.6)}$$

Now consider the following simplification of $\sum_{\hat{u}^n\in\mathcal{F}_q^n} \gamma_{x_1^n}^{\hat{u}^n}$. We have

$$\sum_{\hat{u}^n\in\mathcal{F}_q^n} \gamma_{x_1^n}^{\hat{u}^n} = \sum_{\hat{u}^n\in\mathcal{T}_\delta^{(n)}(U)} \pi_\rho \pi_{x_1^n} \pi_{x_1^n,u^n} \pi_{x_1^n} \pi_\rho$$

$$
\begin{aligned}
&\overset{(a)}{\leq} 2^{n(S(Y_1|X_1,U)_{\sigma_1}+\delta_{u_4})} \sum_{\hat{u}^n \in \mathcal{T}_\delta^{(n)}(U)} \pi_\rho \pi_{x_1^n} \rho_{x_1^n u^n} \pi_{x_1^n} \pi_\rho \\
&\overset{(b)}{\leq} 2^{n(S(Y_1|X_1,U)_{\sigma_1}+\delta_{u_4})} 2^{n(H(U)+\delta_{u_1})} \sum_{\hat{u}^n \in \mathcal{T}_\delta^{(n)}(U)} p_U^n(u^n) \pi_\rho \pi_{x_1^n} \rho_{x_1^n u^n} \pi_{x_1^n} \pi_\rho \\
&\leq 2^{n(S(Y_1|X_1,U)_{\sigma_1}+\delta_{u_4})} 2^{n(H(U)+\delta_{u_1})} \pi_\rho \pi_{x_1^n} \left(\sum_{\hat{u}^n \in \mathcal{F}_q^n} p_U^n(u^n) \rho_{x_1^n u^n} \right) \pi_{x_1^n} \pi_\rho \\
&= 2^{n(S(Y_1|X_1,U)_{\sigma_1}+\delta_{u_4})} 2^{n(H(U)+\delta_{u_1})} \pi_\rho \pi_{x_1^n} \rho_{x_1^n} \pi_{x_1^n} \pi_\rho \\
&\overset{(c)}{\leq} 2^{n(S(Y_1|X_1,U)_{\sigma_1}+\delta_{u_4})} 2^{n(H(U)+\delta_{u_1})} 2^{-n(S(Y_1|X_1)-\delta_{x_1})} \pi_\rho \pi_{x_1^n} \pi_\rho \overset{(d)}{=} c_1 \cdot I,
\end{aligned}
$$

where (a) follows using the following arguments

$$
\begin{aligned}
\pi_{x_1^n,u^n} \leq 2^{n(S(Y_1|X_1,U)_{\sigma_1}+\delta_{u_4})} \pi_{x_1^n,u^n} \rho_{x_1^n u^n} \pi_{x_1^n,u^n} &= 2^{n(S(Y_1|X_1,U)_{\sigma_1}+\delta_{u_4})} \sqrt{\pi_{x_1^n,u^n}} \rho_{x_1^n u^n} \sqrt{\pi_{x_1^n,u^n}} \\
&= 2^{n(S(Y_1|X_1,U)_{\sigma_1}+\delta_{u_4})} \rho_{x_1^n u^n}.
\end{aligned}
$$

The inequalities (b), (c) uses the typicality arguments. Lastly, the inequality (d) follows by using the fact that $\pi_\rho \pi_{x_1^n} \pi_\rho \leq I$ and by defining $c_1 \triangleq 2^{n(S(Y_1|X_1,U)_{\sigma_1}+H(U)-S(Y_1|X_1)+\delta_{u_1}+\delta_{u_4}+\delta_{x_1})}$.

Substituting the above simplification in (F.5), we obtain

$$
\begin{aligned}
\mathbb{E}[T_{22}] &\leq \frac{c_1 q^{3k} q^l}{q^{3n}} \sum_{x_1^n \in \mathcal{X}_1^n} p_{X_1}^n(x_1^n) \sum_{\substack{v_2^n \in \mathcal{T}_\delta^{(n)}(V_2), \\ v_3^n \in \mathcal{T}_\delta^{(n)}(V_3)}} \sum_{u^n \in \mathcal{F}_q^n} \mathrm{Tr}\left(\pi_{u^n} \rho^{Y_1}_{x_1^n v_2^n v_3^n} \pi_{u^n} \right) \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}} \\
&\overset{(a)}{\leq} \frac{c_1 q^{3k} q^l}{q^{3n}} \sum_{\substack{v_2^n \in \mathcal{T}_\delta^{(n)}(V_2), \\ v_3^n \in \mathcal{T}_\delta^{(n)}(V_3)}} \sum_{u^n \in \mathcal{F}_q^n} \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}} \\
&\overset{(b)}{\leq} \frac{q^{3k} q^l}{q^{3n}} 2^{n(+H(U)-I(Y_1;U|X_1)_{\sigma_1}+\delta_{u_1}+\delta_{u_4}+\delta_{x_1})} 2^{n(H(V_1)+H(V_2)+\delta_{v_1}+\delta_{v_2})},
\end{aligned}
$$

where (a) uses $\mathrm{Tr}\left(\pi_{u^n} \rho^{Y_1}_{x_1^n v_2^n v_3^n} \pi_{u^n} \right) \leq 1$, and (a) follows from the definition of c_1 for σ_1 as defined in the statement of the theorem. This gives the desired rate to bound $\mathbb{E}[T_{22}]$.

F.3 Proof of Proposition VII.14

Using the simplification performed in (F.2), we obtain

$$T_{23} = \sum_{\substack{m_1' \neq m_1,\\ a' \neq a, l' \neq l}} \sum_{x_1^n \in \mathcal{X}_1^n} \sum_{v_2^n, v_3^n \in \mathcal{F}_q^n} \sum_{a_2, a_3 \in \mathcal{F}_q^k} \mathrm{Tr}\left(\gamma_{m_1'}^{a',l'} \pi_l^a \rho_{x_1^n v_2^n v_3^n}^{Y_1} \pi_l^a\right) \mathbb{1}_{\{x_1^n(m_1)=x_1^n\}} \mathbb{1}_{\mathcal{J}}$$

$$\leq \sum_{\substack{m_1' \neq m_1,\\ a' \neq a, l' \neq l}} \sum_{x_1^n, \hat{x}_1^n \in \mathcal{X}_1^n} \sum_{\substack{v_2^n \in \mathcal{T}_\delta^{(n)}(V_2),\\ v_3^n \in \mathcal{T}_\delta^{(n)}(V_3)}} \sum_{a_2, a_3 \in \mathcal{F}_q^k} \sum_{u^n, \hat{u}^n \in \mathcal{F}_q^n} \mathrm{Tr}\left(\gamma_{\hat{x}_1^n}^{\hat{u}^n} \pi_{u^n} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \pi_{u^n}\right) \mathbb{1}_{\{x_1^n(m_1)=x_1^n\}}$$

$$\times \mathbb{1}_{\{x_1^n(m_1')=\hat{x}_1^n\}} \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}} \mathbb{1}_{\{\hat{u}^n = a' g_I + l' g_{O/I} + b_2^n + b_3^n\}} \mathbb{1}_{\mathcal{K}}, \quad \text{(F.7)}$$

where the above inequality follows by noting that $\mathcal{J} \subseteq \mathcal{K}$. By taking expectation of the above term with respect to the codebook generating distributions, we get

$$\mathbb{E}[T_{23}]$$

$$\leq \sum_{\substack{m_1' \neq m_1,\\ a' \neq a, l' \neq l}} \sum_{x_1^n, \hat{x}_1^n \in \mathcal{X}_1^n} \sum_{\substack{v_2^n \in \mathcal{T}_\delta^{(n)}(V_2),\\ v_3^n \in \mathcal{T}_\delta^{(n)}(V_3)}} \sum_{a_2, a_3 \in \mathcal{F}_q^k} \sum_{u^n, \hat{u}^n \in \mathcal{F}_q^n} \mathrm{Tr}\left(\gamma_{\hat{x}_1^n}^{\hat{u}^n} \pi_{u^n} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \pi_{u^n}\right) p_{X_1}^n(x_1^n) p_{X_1}^n(\hat{x}_1^n) \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}}$$

$$\mathbb{P}(\hat{u}^n = a' G_I + l' G_{O/I} + B_2^n + B_3^n, v_2^n = a_2 G_I + m_2 G_{O/I} + B_2^n, v_3^n = a_3 G_I + m_3 G_{O/I} + B_3^n)$$

$$\overset{(a)}{\leq} \frac{2^{nR_1} q^{3k} q^l}{q^{3n}} \sum_{x_1^n, \hat{x}_1^n \in \mathcal{X}_1^n} p_{X_1}^n(\hat{x}_1^n) p_{X_1}^n(x_1^n) \sum_{u^n, \hat{u}^n \in \mathcal{F}_q^n} \mathrm{Tr}\left(\gamma_{\hat{x}_1^n}^{\hat{u}^n} \pi_{u^n} \left(\sum_{\substack{v_2^n \in \mathcal{T}_\delta^{(n)}(V_2),\\ v_3^n \in \mathcal{T}_\delta^{(n)}(V_3)}} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \right) \pi_{u^n} \right) \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}},$$

where the second inequality above uses the claim from (F.6).

Consider the following simplifications.

$$\sum_{x_1^n \in \mathcal{X}_1^n} p_{X_1}^n(\hat{x}_1^n) \sum_{u^n, \hat{u}^n \in \mathcal{F}_q^n} \mathrm{Tr}\left(\gamma_{\hat{x}_1^n}^{\hat{u}^n} \pi_{u^n} \left(\sum_{\substack{v_2^n \in \mathcal{T}_\delta^{(n)}(V_2),\\ v_3^n \in \mathcal{T}_\delta^{(n)}(V_3)}} \rho_{x_1^n v_2^n v_3^n}^{Y_1} \mathbb{1}_{\{u^n = v_2^n \oplus v_3^n\}} \right) \pi_{u^n} \right)$$

$$\overset{(a)}{\leq} 2^{n(H(V_2,V_3|U)+\delta_{u_2})} \sum_{u^n, \hat{u}^n \in \mathcal{T}_\delta^{(n)}(U)} \mathrm{Tr}\left(\gamma_{\hat{x}_1^n}^{\hat{u}^n} \pi_{u^n} \left(\sum_{x_1^n \in \mathcal{X}_1^n} p_{X_1}^n(\hat{x}_1^n) \rho_{x_1^n u^n}^{Y_1} \right) \pi_{u^n} \right)$$

$$\overset{(b)}{\leq} 2^{n(H(V_2,V_3|U)+\delta_{u_2})} \sum_{u^n, \hat{u}^n \in \mathcal{T}_\delta^{(n)}(U)} \mathrm{Tr}\left(\gamma_{\hat{x}_1^n}^{\hat{u}^n} \rho_{u^n}^{Y_1}\right)$$

$$\leq 2^{n(H(V_2,V_3|U)+\delta_{u_2})}2^{n(H(U)+\delta_{u_1})}\sum_{\hat{u}^n\in\mathcal{T}_\delta^{(n)}(U)}\mathrm{Tr}\left(\pi_{\hat{x}_1^n}\pi_{\hat{x}_1^n,\hat{u}^n}\pi_{\hat{x}_1^n}\pi_\rho\Big(\sum_{u^n\in\mathcal{T}_\delta^{(n)}(U)}p_U^n(u^n)\rho_{u^n}^{Y_1}\Big)\pi_\rho\right)$$
$$\overset{(c)}{\leq} 2^{n(H(V_2,V_3|U)+\delta_{u_2})}2^{n(H(U)+\delta_{u_1})}2^{-n(S(Y_1)_\sigma-\delta_\rho)}\sum_{\hat{u}^n\in\mathcal{T}_\delta^{(n)}(U)}\mathrm{Tr}\big(\pi_{\hat{x}_1^n,\hat{u}^n}\big), \tag{F.8}$$

where (a) holds from the above bounds obtained in (F.4), (b) follows by the definition of $\rho_{u^n}^{Y_1}$ and by using the inequality $\pi_{u^n}\rho_{u^n}^{Y_1}\pi_{u^n}\leq\rho_{u^n}^{Y_1}$, (c) follows using (i) $\pi_\rho\Big(\sum_{u^n\in\mathcal{T}_\delta^{(n)}(U)}p_U^n(u^n)\rho_{u^n}^{Y_1}\Big)\pi_\rho\leq\pi_\rho\rho\pi_\rho\leq 2^{-n(S(Y_1)_\sigma-\delta_\rho)}\pi_\rho$ and (ii) $\mathrm{Tr}\big(\pi_{\hat{x}_1^n}\pi_{\hat{x}_1^n,\hat{u}^n}\pi_{\hat{x}_1^n}\pi_\rho\big)\leq\mathrm{Tr}\big(\pi_{\hat{x}_1^n,\hat{u}^n}\big)$. Using the above simplification in (F.8) we obtain

$$\begin{aligned}\mathbb{E}[T_{23}] &\leq \frac{2^{nR_1}q^{3k}q^l}{q^{3n}}2^{n(H(V_2,V_3|U)+\delta_{u_2})}2^{n(H(U)+\delta_{u_1})}2^{-n(S(Y_1)_\sigma-\delta_\rho)}\sum_{\hat{x}_1^n\in\mathcal{X}_1^n}p_{X_1}^n(\hat{x}_1^n)\sum_{\hat{u}^n\in\mathcal{T}_\delta^{(n)}(U)}\mathrm{Tr}\big(\pi_{\hat{x}_1^n,\hat{u}^n}\big)\\ &\leq \frac{2^{nR_1}q^{3k}q^l}{q^{3n}}2^{n(H(V_2,V_3|U)+\delta_{u_2})}2^{n(H(U)+\delta_{u_1})}2^{-n(S(Y_1)_\sigma-\delta_\rho)}2^{n(H(U)+\delta_{u_1})}2^{n(S(Y_1|X_1,U)_{\sigma_1}+\delta_{u_4})}\\ &= \frac{2^{nR_1}q^{3k}q^l}{q^{3n}}2^{n(H(V_2,V_3)+H(U)-I(Y_1;X_1,U)_{\sigma_1}+\delta_\rho+2\delta_{u_1}+\delta_{u_2}+\delta_{u_4})}.\end{aligned} \tag{F.9}$$

This completes the proof.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abeyesinghe, A., I. Devetak, P. Hayden, and A. Winter (2009), The mother of all protocols: Restructuring quantum information's family tree, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *465*(2108), 2537–2563.

Ahlswede, R., and N. Cai (2006), A strong converse theorem for quantum multiple access channels, in *General Theory of Information Transfer and Combinatorics*, pp. 460–485, Springer.

Ahlswede, R., and A. Winter (2002), Strong converse for identification via quantum channels, *IEEE Transactions on Information Theory*, *48*(3), 569–579.

Alicki, R., M. Horodecki, P. Horodecki, and R. Horodecki (2004), Thermodynamics of quantum information systems—hamiltonian description, *Open Systems & Information Dynamics*, *11*(3), 205–217.

Anshu, A., V. K. Devabathini, and R. Jain (2014), Quantum message compression with applications, *arXiv preprint arXiv:1410.3031*.

Anshu, A., V. K. Devabathini, and R. Jain (2017a), Quantum communication using coherent rejection sampling, *Physical review letters*, *119*(12), 120,506.

Anshu, A., R. Jain, and N. Warsi (2017b), Measurement compression with quantum side information using shared randomness, *arXiv preprint arXiv:1703.02342*.

Anshu, A., R. Jain, and N. A. Warsi (2018a), A generalized quantum Slepian–Wolf, *IEEE Transactions on Information Theory*, *64*(3), 1436–1453, doi:10.1109/TIT.2017.2786348.

Anshu, A., R. Jain, and N. A. Warsi (2018b), Building blocks for communication over noisy quantum networks, *IEEE Transactions on Information Theory*, *65*(2), 1287–1306.

Anshu, A., R. Jain, and N. A. Warsi (2019), Convex-split and hypothesis testing approach to one-shot quantum measurement compression and randomness extraction, *IEEE Transactions on Information Theory*, *65*(9), 5905–5924, doi:10.1109/TIT.2019.2915242.

Atif, T. A., and S. S. Pradhan (2021), Distributed quantum faithful simulation and function computation using algebraic structured measurements, in *2021 IEEE International Symposium on Information Theory (ISIT)*, pp. 124–129, IEEE.

Atif, T. A., and S. S. Pradhan (2022), Multi-party quantum purity distillation with bounded classical communication, in *2022 IEEE International Symposium on Information Theory (ISIT)*, pp. 2974–2979, doi:10.1109/ISIT50566.2022.9834596.

Atif, T. A., M. Heidari, and S. S. Pradhan (2021a), Faithful simulation of distributed quantum measurements with applications in distributed rate-distortion theory, *IEEE Transactions on Information Theory*, pp. 1–1, doi:10.1109/TIT.2021.3124976.

Atif, T. A., A. Padakandla, and S. S. Pradhan (2021b), Achievable rate-region for 3—user classical-quantum interference channel using structured codes, in *2021 IEEE International Symposium on Information Theory (ISIT)*, pp. 760–765, doi:10.1109/ISIT45174.2021.9518095.

Atif, T. A., A. Padakandla, and S. S. Pradhan (2021c), Computing sum of sources over a classical-quantum mac, in *2021 IEEE International Symposium on Information Theory (ISIT)*, pp. 414–419, doi:10.1109/ISIT45174.2021.9518261.

Atif, T. A., M. Heidari, and S. S. Pradhan (2022a), Faithful simulation of distributed quantum measurements with applications in distributed rate-distortion theory, *IEEE Transactions on Information Theory*, *68*(2), 1085–1118, doi:10.1109/TIT.2021.3124976.

Atif, T. A., A. Padakandla, and S. Sandeep Pradhan (2022b), Source coding for synthesizing correlated randomness, *IEEE Transactions on Information Theory*, pp. 1–1, doi:10.1109/TIT.2022.3196186.

Aubrun, G., and I. Nechita (2008), Catalytic majorization and ℓ_p norms, *Communications in Mathematical Physics*, *278*(1), 133–144.

Baghali Khanian, Z., K. Kuroiwa, and D. Leung (2022), Rate-distortion theory for mixed states, *arXiv e-prints*, pp. arXiv–2208.

Barnum, H. (2000), Quantum rate-distortion coding, *Physical Review A*, *62*(4), 042,309.

Barnum, H., and E. Knill (2002), Reversing quantum dynamics with near-optimal quantum and classical fidelity, *Journal of Mathematical Physics*, *43*(5), 2097–2106.

Beals, R., S. Brierley, O. Gray, A. W. Harrow, S. Kutin, N. Linden, D. Shepherd, and M. Stather (2013), Efficient distributed quantum computing, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *469*(2153), 20120,686.

Bennett, C. H., P. W. Shor, J. A. Smolin, and A. V. Thapliyal (2002), Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem, *IEEE Transactions on Information Theory*, *48*(10), 2637–2655.

Bennett, C. H., I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter (2014), The quantum reverse Shannon theorem and resource tradeoffs for simulating quantum channels, *IEEE Transactions on Information Theory*, *60*(5), 2926–2959.

Berger, T. (1971), *Rate Distortion Theory: A Mathematical Basis for Data Compression*, Prentice-Hall electrical engineering series, Prentice-Hall.

Berger, T. (1975), Rate distortion theory and data compression, in *Advances in Source Coding*, pp. 1–39, Springer.

Berger, T. (1977), Multiterminal source coding, *The Inform. Theory Approach to Communications, G. Longo, Ed., New York: Springer-Verlag.*

Berta, M., M. Christandl, and R. Renner (2011), The quantum reverse shannon theorem based on one-shot information theory, *Communications in Mathematical Physics*, *306*(3), 579.

Berta, M., J. M. Renes, and M. M. Wilde (2014), Identifying the information gain of a quantum measurement, *IEEE Transactions on Information Theory*, *60*(12), 7987–8006.

Brandao, F., M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner (2015), The second laws of quantum thermodynamics, *Proceedings of the National Academy of Sciences*, *112*(11), 3275–3279.

Bu, K., U. Singh, and J. Wu (2016), Catalytic coherence transformations, *Physical Review A*, *93*(4), 042,326.

Buscemi, F., M. Hayashi, and M. Horodecki (2008), Global information balance in quantum measurements, *Physical review letters*, *100*(21), 210,504.

Carlen, E. (2010), Trace inequalities and quantum entropy: an introductory course, *Entropy and the quantum*, *529*, 73–140.

Cheng, H.-C., E. P. Hanson, N. Datta, and M.-H. Hsieh (2019), Duality between source coding with quantum side information and cq channel coding, in *2019 IEEE International Symposium on Information Theory (ISIT)*, pp. 1142–1146, IEEE.

Christandl, M., R. König, and R. Renner (2009), Postselection technique for quantum channels with applications to quantum cryptography, *Physical review letters*, *102*(2), 020,504.

Conway, J. B. (1985), *A Course in Functional Analysis*, Springer New York, doi:10.1007/978-1-4757-3828-5.

Csiszár, I., and J. Körner (2011), *Information theory: coding theorems for discrete memoryless systems*, Cambridge University Press.

Cuff, P. (2013), Distributed channel synthesis, *IEEE Transactions on Information Theory*, *59*(11), 7071–7096.

Cuff, P. W. (2009), *Communication in networks for coordinating behavior*, Stanford University.

Cuff, P. W., H. H. Permuter, and T. M. Cover (2010), Coordination capacity, *IEEE Transactions on Information Theory*, *56*(9), 4181–4206.

Daftuar, S., and M. Klimesh (2001), Mathematical structure of entanglement catalysis, *Physical Review A*, *64*(4), 042,314.

Datta, N., M.-H. Hsieh, and M. M. Wilde (2013a), Quantum rate distortion, reverse shannon theorems, and source-channel separation, *IEEE Transactions on Information Theory*, *59*(1), 615–630, doi:10.1109/TIT.2012.2215575.

Datta, N., M.-H. Hsieh, M. M. Wilde, and A. Winter (2013b), Quantum-to-classical rate distortion coding, *Journal of Mathematical Physics*, *54*(4), 042,201.

Datta, N., J. M. Renes, R. Renner, and M. M. Wilde (2013c), One-shot lossy quantum data compression, *IEEE Transactions on Information Theory*, *59*(12), 8057–8076.

Denchev, V. S., and G. Pandurangan (2008), Distributed quantum computing: A new frontier in distributed systems or science fiction?, *ACM SIGACT News*, *39*(3), 77–95.

Devetak, I. (2005a), Distillation of local purity from quantum states, *Physical Review A*, *71*(6), 062,303.

Devetak, I. (2005b), The private classical capacity and quantum capacity of a quantum channel, *IEEE Transactions on Information Theory*, *51*(1), 44–55.

Devetak, I., and T. Berger (2002), Quantum rate-distortion theory for memoryless sources, *IEEE Transactions on Information Theory*, *48*(6), 1580–1589.

Devetak, I., and A. Winter (2003), Classical data compression with quantum side information, *Physical Review A*, *68*(4), 042,301.

Devetak, I., and A. Winter (2004), Distilling common randomness from bipartite quantum states, *IEEE Transactions on Information Theory*, *50*(12), 3183–3196.

Devetak, I., and J. Yard (2008), Exact cost of redistributing multipartite quantum states, *Physical Review Letters*, *100*(23), 230,501.

Devetak, I., A. W. Harrow, and A. J. Winter (2008), A resource framework for quantum shannon theory, *IEEE Transactions on Information Theory*, *54*(10), 4587–4618.

Ding, F., X. Hu, and H. Fan (2021), Amplifying asymmetry with correlating catalysts, *Physical Review A*, *103*(2), 022,403.

Duarte, C., R. C. Drumond, and M. T. Cunha (2016), Self-catalytic conversion of pure quantum states, *Journal of Physics A: mathematical and theoretical*, *49*(14), 145,303.

Dueck, G. (1981), The strong converse to the coding theorem for the multiple–access channel, *J. Comb. Inform. Syst. Sci*, *6*(3), 187–196.

El Gamal, A., and Y.-H. Kim (2011), *Network information theory*, Cambridge university press.

Fawzi, O., P. Hayden, I. Savov, P. Sen, and M. M. Wilde (2012), Classical communication over a quantum interference channel, *IEEE Transactions on Information Theory*, *58*(6), 3670–3691, doi:10.1109/TIT.2012.2188620.

Fuchs, C. A., and J. Van De Graaf (1999), Cryptographic distinguishability measures for quantum-mechanical states, *IEEE Transactions on Information Theory*, *45*(4), 1216–1227.

Gallager, R. G. (1968), *Information Theory and Reliable Communication*, John Wiley & Sons, New York.

Gerrish, A. M. (1963), Estimation of information rates, Ph.D. thesis, Yale University.

Groenewold, H. J. (1971), A problem of information gain by quantal measurements, *International Journal of Theoretical Physics*, *4*(5), 327–338.

Han, T., and K. Kobayashi (1981), A new achievable rate region for the interference channel, *IEEE Transactions on Information Theory*, *27*(1), 49–60, doi:10.1109/TIT.1981.1056307.

Hayashi, M., and H. Nagaoka (2003), General formulas for capacity of classical-quantum channels, *IEEE Transactions on Information Theory*, *49*(7), 1753–1768.

Hayashi, M., and Á. Vázquez-Castro (2021), Computation-aided classical-quantum multiple access to boost network communication speeds, *Physical Review Applied*, *16*(5), 054,021.

Hayden, P., R. Jozsa, D. Petz, and A. Winter (2004), Structure of states which satisfy strong subadditivity of quantum entropy with equality, *Communications in mathematical physics*, *246*(2), 359–374.

Hayden, P., M. Horodecki, A. Winter, and J. Yard (2008), A decoupling approach to the quantum capacity, *Open Systems & Information Dynamics*, *15*(01), 7–19.

Hirche, C., C. Morgan, and M. M. Wilde (2016), Polar codes in network quantum information theory, *IEEE Transactions on Information Theory*, *62*(2), 915–924.

Holevo, A. S. (1998), The capacity of the quantum channel with general signal states, *IEEE Transactions on Information Theory*, *44*(1), 269–273, doi:10.1109/18.651037.

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

Holevo, A. S. (2019), Quantum systems, channels, information, in *Quantum Systems, Channels, Information*, de Gruyter.

Horodecki, M., K. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen, U. Sen, et al. (2003a), Local information as a resource in distributed quantum systems, *Physical review letters*, *90*(10), 100,402.

Horodecki, M., P. Horodecki, and J. Oppenheim (2003b), Reversible transformations from pure to mixed states and the unique measure of information, *Physical Review A*, *67*(6), 062,104.

Horodecki, M., P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen, U. Sen, B. Synak-Radtke, et al. (2005a), Local versus nonlocal information in quantum-information theory: formalism and phenomena, *Physical Review A*, *71*(6), 062,307.

Horodecki, M., J. Oppenheim, and A. Winter (2005b), Partial quantum information, *Nature*, *436*(7051), 673–676.

Horodecki, M., J. Oppenheim, and A. Winter (2007), Quantum state merging and negative information, *Communications in Mathematical Physics*, *269*(1), 107–136.

Hsieh, M.-H., and S. Watanabe (2016), Channel simulation and coded source compression, *IEEE Transactions on Information Theory*, *62*(11), 6609–6619.

Jafarian, A., and S. Vishwanath (2012), Achievable rates for k-user Gaussian interference channels, *IEEE Transactions on information theory*, *58*(7), 4367–4380.

Jonathan, D., and M. B. Plenio (1999), Entanglement-assisted local manipulation of pure quantum states, *Phys. Rev. Lett.*, *83*, 3566–3569, doi:10.1103/PhysRevLett.83.3566.

Jozsa, R., and B. Schumacher (1994), A new proof of the quantum noiseless coding theorem, *Journal of Modern Optics*, *41*(12), 2343–2349.

Khanian, Z. B., and A. Winter (2021), A rate-distortion perspective on quantum state redistribution, *arXiv preprint arXiv:2112.11952*.

Khanian, Z. B., and A. Winter (2022), General mixed-state quantum data compression with and without entanglement assistance, *IEEE Transactions on Information Theory*, *68*(5), 3130–3138.

Khatri, S., and M. M. Wilde (2020), Principles of quantum communication theory: A modern approach, *arXiv preprint arXiv:2011.04672*.

Koashi, M., and N. Imoto (2001), Compressibility of quantum mixed-state signals, *Physical Review Letters*, *87*(1), 017,902.

Korner, J., and K. Marton (1979), How to encode the modulo-two sum of binary sources (corresp.), *IEEE Transactions on Information Theory*, *25*(2), 219–221.

Kostina, V., and S. Verdú (2015), The output distribution of good lossy source codes, in *2015 Information Theory and Applications Workshop (ITA)*, pp. 308–312, IEEE.

Krithivasan, D., and S. S. Pradhan (2011), Distributed source coding using abelian group codes: A new achievable rate-distortion region, *IEEE Transactions on Information Theory*, *57*(3), 1495–1519.

Krovi, H., and I. Devetak (2007), Local purity distillation with bounded classical communication, *Physical Review A*, *76*(1), 012,321.

Lindblad, G. (1972), An entropy inequality for quantum measurements, *Communications in Mathematical Physics*, *28*(3), 245–249.

Lipka-Bartosik, P., and P. Skrzypczyk (2021), Catalytic quantum teleportation and beyond, *arXiv preprint arXiv:2102.11846*.

Lloyd, S. (1997), Capacity of the noisy quantum channel, *Physical Review A*, *55*(3), 1613.

Luo, S. (2010), Information conservation and entropy change in quantum measurements, *Physical Review A*, *82*(5), 052,103.

Luo, Z., and I. Devetak (2009), Channel simulation with quantum side information, *IEEE Transactions on Information Theory*, *55*(3), 1331–1342.

Mattson Jr, H. (2012), An upper bound on covering radius, in *Combinatorial Mathematics Proceedings of the International Colloquium on Graph Theory and Combinatorics*, vol. 75, pp. 453–458.

Meyer, P. A. (2006), *Quantum probability for probabilists*, Springer.

Naimark, M. (1940), Spectral functions of a symmetric operator, *Bull. Acad. Sci. URSS. Sér. Math.[Izvestia Akad. Nauk SSSR]*, *4*, 277–318.

Nazer, B., and M. Gastpar (2007), Computation over multiple-access channels, *IEEE Transactions on information theory*, *53*(10), 3498–3516.

Nielsen, M. A., and I. Chuang (2002), Quantum computation and quantum information.

Oppenheim, J., M. Horodecki, P. Horodecki, and R. Horodecki (2002), Thermodynamical approach to quantifying quantum correlations, *Physical review letters*, *89*(18), 180,402.

Ozawa, M. (1986), On information gain by quantum measurements of continuous observables, *Journal of mathematical physics*, *27*(3), 759–763.

Padakandla, A. (2014), An Algebraic Framework for Multi-terminal Communication, Ph.D. thesis, Univ. of Michigan, Ann Arbor, USA, available at http://deepblue.lib.umich.edu/handle/2027.42/107264.

Padakandla, A., and S. Pradhan (), Computing sum of sources over an arbitrary multiple access channel, available at `http://arxiv.org/abs/1301.5684`.

Padakandla, A., and S. S. Pradhan (2013), Computing sum of sources over an arbitrary multiple access channel, in *2013 IEEE International Symposium on Information Theory*, pp. 2144–2148, IEEE.

Padakandla, A., and S. S. Pradhan (2017), An achievable rate region based on coset codes for multiple access channel with states, *IEEE Transactions on Information Theory*, *63*(10), 6393–6415, doi:10.1109/TIT.2017.2726069.

Padakandla, A., and S. S. Pradhan (2018), Achievable Rate Region for Three User Discrete Broadcast Channel Based on Coset Codes, *IEEE Transactions on Information Theory*, *64*(4), 2267–2297, doi:10.1109/TIT.2018.2798669.

Padakandla, A., A. G. Sahebi, and S. S. Pradhan (2016), An Achievable Rate Region for the Three-User Interference Channel Based on Coset Codes, *IEEE Transactions on Information Theory*, *62*(3), 1250–1279, doi:10.1109/TIT.2016.2518171.

Petz, D. (1986), Sufficient subalgebras and the relative entropy of states of a von neumann algebra, *Communications in mathematical physics*, *105*(1), 123–131.

Philosof, T., and R. Zamir (2009), On the loss of single-letter characterization: The dirty multiple access channel, *IEEE Trans. on Info. Th.*, *55*, 2442–2454.

Poor, H. V. (1998), *An introduction to signal detection and estimation*, Springer Science & Business Media.

Powers, R. T., and E. Størmer (1970), Free states of the canonical anticommutation relations, *Communications in Mathematical Physics*, *16*(1), 1–33.

Pradhan, S. S. (2004), Approximation of test channels in source coding, in *Proc. Conf. Inform. Syst. Sci.(CISS)*.

Pradhan, S. S., A. Padakandla, and F. Shirani (2021), *An Algebraic and Probabilistic Framework for Network Information Theory*, vol. 18, 173–376 pp., Foundations and Trends in Communications and Information Theory.

Renes, J. M., and R. Renner (2012), One-shot classical data compression with quantum side information and the distillation of common randomness or secret keys, *IEEE Transactions on Information Theory*, *58*(3), 1985–1991.

Salek, S., D. Cadamuro, P. Kammerlander, and K. Wiesner (2018), Quantum rate-distortion coding of relevant information, *IEEE Transactions on Information Theory*, *65*(4), 2603–2613.

Sanders, Y. R., and G. Gour (2009), Necessary conditions for entanglement catalysts, *Physical Review A*, *79*(5), 054,302.

Savov, I. (2012), Network information theory for classical-quantum channels, *arXiv preprint arXiv:1208.4188*.

Schieler, C., and P. Cuff (2013), A connection between good rate-distortion codes and backward dmcs, in *2013 IEEE Information Theory Workshop (ITW)*, pp. 1–5, IEEE.

Schumacher, B. (1995), Quantum coding, *Physical Review A*, *51*(4), 2738.

Schumacher, B., and M. D. Westmoreland (1997), Sending classical information via noisy quantum channels, *Phys. Rev. A*, *56*, 131–138, doi:10.1103/PhysRevA.56.131.

Sen, P. (2012), Achieving the han-kobayashi inner bound for the quantum interference channel, in *2012 IEEE International Symposium on Information Theory Proceedings*, pp. 736–740, IEEE.

Sen, P. (2018a), Inner bounds via simultaneous decoding in quantum network information theory, *arXiv preprint arXiv:1806.07276*.

Sen, P. (2018b), A one-shot quantum joint typicality lemma, *arXiv preprint arXiv:1806.07278*.

Sen, P. (2021), Unions, intersections and a one-shot quantum joint typicality lemma, *Sādhanā*, *46*(1), 57.

Shamai, S., and S. Verdú (1997), The empirical distribution of good codes, *IEEE Transactions on Information Theory*, *43*(3), 836–846.

Shannon, C. E., et al. (1959), Coding theorems for a discrete source with a fidelity criterion, *IRE Nat. Conv. Rec*, *4*(142-163), 1.

Shiraishi, N., and T. Sagawa (2021), Quantum thermodynamics of correlated-catalytic state conversion at small scale, *Physical Review Letters*, *126*(15), 150,502.

Shirani, F., and S. S. Pradhan (2014), Finite block-length gains in distributed source coding, in *2014 IEEE International Symposium on Information Theory*, pp. 1702–1706, doi:10.1109/ISIT.2014.6875124.

Shirokov, M. E. (2011), Entropy reduction of quantum measurements, *Journal of mathematical physics*, *52*(5), 052,202.

Shor, P. W. (2002), The quantum channel capacity and coherent information, in *lecture notes, MSRI Workshop on Quantum Computation*.

Slepian, D. S., and J. K. Wolf (July 1973), Noiseless coding of correlated information sources, *IEEE Transactions on Information Theory*, *19*(4), 471–480.

Sohail, M. A., T. A. Atif, A. Padakandla, and S. S. Pradhan (2022), Computing sum of sources over a classical-quantum mac, *IEEE Transactions on Information Theory*, *68*(12), 7913–7934, doi:10.1109/TIT.2022.3192876.

Šupić, I., J.-D. Bancal, Y. Cai, and N. Brunner (2022), Genuine network quantum nonlocality and self-testing, *Physical Review A*, *105*(2), 022,206.

Sutter, D. (2018), Approximate quantum markov chains, in *Approximate Quantum Markov Chains*, pp. 75–100, Springer.

Takagi, R., and N. Shiraishi (2021), Correlation in catalysts enables arbitrary manipulation of quantum coherence, *arXiv preprint arXiv:2106.12592*.

Tomamichel, M. (2015), *Quantum information processing with finite resources: mathematical foundations*, vol. 5, Springer.

Turgut, S. (2007), Catalytic transformations for bipartite pure states, *Journal of Physics A: Mathematical and Theoretical*, *40*(40), 12,185.

Uhlmann, A. (1976), The "transition probability" in the state space of a*-algebra, *Reports on Mathematical Physics*, *9*(2), 273–279.

van Dam, W., and P. Hayden (2002), Embezzling entangled quantum states, *arXiv preprint quant-ph/0201041*.

Van Meter, R., and S. J. Devitt (2016), The path to scalable distributed quantum computing, *Computer*, *49*(9), 31–42.

Wagner, A. B., and V. Anantharam (2008), An improved outer bound for multiterminal source coding, *IEEE Transactions on Information Theory*, *54*(5), 1919–1937.

Weissman, T., and E. Ordentlich (2005), The empirical distribution of rate-constrained source codes, *IEEE transactions on information theory*, *51*(11), 3718–3733.

Wilde, M. M. (2011), From classical to quantum shannon theory, *arXiv preprint arXiv:1106.1445*.

Wilde, M. M. (2013a), *Quantum information theory*, Cambridge University Press (Also see arXiv:1106.1445).

Wilde, M. M. (2013b), Sequential decoding of a general classical-quantum channel, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *469*(2157), 20130,259.

Wilde, M. M. (2017), Preface to the second edition, *Quantum Information Theory*, p. xi–xii, doi: 10.1017/9781316809976.001.

Wilde, M. M. (Aug 2019), Private Communication.

Wilde, M. M., P. Hayden, F. Buscemi, and M.-H. Hsieh (2012), The information-theoretic costs of simulating quantum measurements, *Journal of Physics A: Mathematical and Theoretical*, *45*(45), 453,001.

Wilde, M. M., N. Datta, M.-H. Hsieh, and A. Winter (2013), Quantum rate-distortion coding with auxiliary resources, *IEEE Transactions on Information Theory*, *59*(10), 6755–6773.

Winter, A. (1999), Coding theorems of quantum information theory, *arXiv preprint quant-ph/9907077*.

Winter, A. (2001), The capacity of the quantum multiple-access channel, *IEEE Transactions on Information Theory*, *47*(7), 3059–3065.

Winter, A. (2002), Compression of sources of probability distributions and density operators, *arXiv preprint quant-ph/0208131*.

Winter, A. (2004), "Extrinsic" and "intrinsic" data in quantum measurements: asymptotic convex decomposition of positive operator valued measures, *Communication in Mathematical Physics*, *244*(1), 157–185.

Yimsiriwattana, A., and S. J. Lomonaco Jr (2004), Distributed quantum computing: A distributed shor algorithm, in *Quantum Information and Computation II*, vol. 5436, pp. 360–372, International Society for Optics and Photonics.

Ziegler, G. M. (2012), *Lectures on polytopes*, vol. 152, Springer Science & Business Media.