April 2020

Limitations on Protecting Information Against Quantum Adversaries

Eneet Kaur

LIMITATIONS ON PROTECTING INFORMATION AGAINST QUANTUM ADVERSARIES

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Physics and Astronomy

by
Eneet Kaur
B.Sc., University of Delhi, 2012
M.Sc., Indian Institute of Technology, Roorkee, 2014

May 2020

Acknowledgments

First and foremost, I would like to thank my advisor, Mark M. Wilde, for offering me the opportunity to work with him. My discussions with him have been instrumental in shaping my approach towards research and in writing research papers. I am also thankful to Mark for providing me with the much-needed confidence boost, and for being an incredible mentor. I would also like to thank Jonathan P. Dowling for his encouragement, the fun pizza parties, and for all his stories. I would also like to thank Andreas Winter for hosting me in Barcelona and for his collaborations. My discussions with him were valuable in shaping my approach to research. I would also like to thank Carl Miller, Stephanie Wehner, and David Elkouss for hosting me in Summer 2017 and for their valuable discussions. I want to thank Masahiro Takeoka for hosting me at NICT, for his collaboration, and his discussions. I would also like to thank Saikat Guha for our collaborations and discussions. I would also like to thank Peter Labonne Bierhorst, Daniel Sheehy, and Gestur Olafsson for taking out time to be on my Ph.D. committee. I appreciate their valuable feedback. I also thank A. R. P. Rau for being available for discussions.

I acknowledge funding support from the US Office of Naval Research and the National Science Foundation under Grant No. 1350397. I would like to thank Yao Zeng, Claire R Bullock, Paige Whittington, and other administrative staff members of the Department of Physics and Astronomy for their help with various administrative tasks.

I would also like to thank my friends and colleagues: Kevin Valson Jacob, Sumeet Khatri, Siddhartha Das, Vishal Katariya, Michelle Lollie, Sahil Saini, David Kekejian, Qingle Wang, and, Chenglong You for their fun and at times serious conversations. Anthony Brady and Kevin Valson Jacob, thanks for all the coffee breaks and discussions. I would especially like to thank Siddharth Soni for his incredible support during all these years. For keeping me grounded, for believing in me and for encouraging me throughout this journey.

I would also like to thank my mother, my father, my sister Mahima, and my brother Bisman for their support. I would also like to thank Sucheta Majumdar and Anu Kundu for all the fun times we have had, and for sticking with me for all these years.

Table of Contents

List of Figures

Abstract

The aim of this thesis is to understand the fundamental limitations on secret key distillation in various settings of quantum key distribution. We first consider quantum steering, which is a resource for one-sided device-independent quantum key distribution. We introduce a conditional mutual information based quantifier for quantum steering, which we call *intrinsic steerability*. Next, we consider quantum non-locality, which is a resource for device-independent quantum key distribution. In this context, we introduce a quantifier, *intrinsic non-locality*, which is a monotone in the resource theory of Bell non-locality. Both these quantities are inspired by intrinsic information and squashed entanglement and are based on conditional mutual information. The idea behind these quantifiers is to suppress the correlations that can be explained by a local hidden variable or by an inaccessible quantum system, thus quantifying the remaining intrinsic correlations. We then prove various properties of these two monotones, which includes the following: monotonicity under free operations, additivity under tensor product of objects, convexity, and faithfulness, among others.

Next, we prove that intrinsic steerability is an upper bound on the secret-key-agreement capacity of an assemblage, and intrinsic non-locality is an upper bound on the secret-key-agreement capacity of a quantum probability distribution. Thus we prove that these quantities are upper bounds on the achievable key rates in one-sided device-independent and device-independent quantum key distribution protocols. We also calculate these bounds for certain honest devices. The study of these upper bounds is instrumental in understanding the limitations of protocols that can be designed for various settings. These upper bounds inform us that, even if one considers the best possible protocol, there is no possibility of exceeding the upper bounds on key rates without a quantum repeater. The upper bounds introduced in this thesis are

an important step for initiating this line of research in one-sided device-independent and in device-independent quantum key distribution.

Chapter 1
Introduction and Preliminaries

1.1. Introduction

Information theory is a beautiful mathematical theory that was initiated by Claude Shannon in 1948 in his seminal paper "A Mathematical Theory of Communication" [1]. This paper has been instrumental in understanding the fundamental limits of communication over noisy channels. The main contribution of this paper was to show that, given a channel, the rate of classical information that can be reliably transmitted between two parties over this channel in the asymptotic limit is given by the channel capacity. This channel capacity is a function of the noise present in the channel.

Shannon's theory dealt primarily with classical channels and with bits as the information units. However, it was realized that one can consider quantum objects as information carriers as well. This change in the way we represent information led to the emergence of quantum information theory. Researchers began to investigate the fundamental limitations of quantum information.

A number of surprising results, such as superadditivity of communication capacities of channels and negativity of conditional entropy, among others, were discovered. An important paper, relevant to this thesis, was that of Devatak and Winter [2], in which they bounded the distillable secret key of a bipartite quantum state from below.

Besides the communication tasks considered in quantum Shannon theory, consideration of quantum objects as carriers of information also initiated a different way of thinking about security. For the first time, it became possible to think of *information-theoretic security* in cryptographic protocols.

Now, the behavior of these quantum objects is fundamentally different from that of classical carriers of information. Quantum bits, qubits, are inherently continuously valued. However, the information that we effectively use is classical; therefore, the information content or the accessible information of a qubit is equal to one bit. Also, qubits cannot be copied. Another important aspect of qubits is that of purification of quantum states. This statement is a reflection of the correlations that can be shared between two or more quantum systems. It states that, given a mixed state, one can embed the mixed state in a larger Hilbert space to obtain a pure state. These pure states do not share correlations with any other system in the universe. There are other statements regarding the uncertainty principle, teleportation, and super-dense coding that make qubits fundamentally different from bits.

The statement of purification is of considerable importance in quantum key distribution. Suppose that Alice and Bob are connected by a channel and Alice sends some information to Bob using this channel. Then, if the information sent is classical, an eavesdropper can tap the channel and copy the classical information that is being transmitted. In this scenario, *Bob and Eve have the same information*, where Eve is the eavesdropper. Now, suppose that instead of classical information, Alice sends qubits to Bob. In this scenario, the Eavesdropper cannot copy the information. Therefore, Eve cannot have the same information as Bob. However, it is possible that Eve copies some information. This copying of information is introduced in the form of information loss in the qubit that reaches Bob. To get an estimate of the loss of the information, one can consider a purification of the state shared by Alice and Bob. This gives an estimate of the total information held by the eavesdropper. This principle here, where the information is conserved and cannot be copied, is instrumental for security proofs in quantum key distribution. Quantum key distribution was first introduced by [3], and the first security proof given in [4]. It was soon realized that the Devetak-Winter [2] result can be leveraged in security proofs for quantum key distribution scenarios, thereby connecting quantum Shannon theory with quantum key distribution protocols. The Devetak-Winter formula is interesting because it applies to a large set of protocols, making it possible to obtain information-theoretic security for these protocols.

Since its introduction, one major goal of researchers in quantum key distribution has been to introduce better quantum key distribution protocols. Essentially, one would want protocols that give higher key rates for expected noise models. One can ask the following: is there a fundamental limit on key rates that can be extracted from any possible protocol? The answer is affirmative and has been explored in a number of papers [5, 6, 7, 8]. The results basically tell us about the fundamental limitations on secret key rates that one can obtain from any possible protocol, and that it is impossible to go higher than the aforementioned limits without a quantum repeater.

There are two fundamental entropic quantities that have been useful as upper bounds on key rates in quantum key distribution. The first one is conditional mutual information, and the second one is relative entropy. In this thesis, we concentrate on conditional mutual information. The quantities built from conditional mutual information quantify the correlations shared exclusively between two parties who want to share the key (i.e., those that cannot be shared with an external party) and suppress the correlations shared by an external party, thus characterizing the intrinsic information. This quantity was first used in [9] for characterizing secret correlations in a joint probability distribution P_{XYZ}, and it was also used to upper bound rates in secret key distillation from this joint probability distribution. Next, conditional mutual information was considered in the context of quantifying entanglement and was used to define the squashed entanglement measure [10]. This was then proved to be an upper bound on the distillable key of a bipartite state [11].

Table 1.1. Conditional mutual information based quantities and settings in quantum key distribution.

Resource	**Tasks**	**Measure**
Joint probability distributions	Private key distillation	Intrinsic information [9]
Bipartite states	Trusted QKD	Squashed entanglement [10]
Steerable assemblages	One-sided device-independent QKD	Intrinsic steerability [12]
Conditional probability distributions	Device-independent QKD	Intrinsic non-locality [13]

In this work, we use conditional mutual information for characterizing steerability and non-locality. These quantities are then used as upper bounds in different settings of quantum key distribution. We can summarize the contributions of this thesis in Table 1.1.

We start this thesis by outlining the basic concepts in quantum information theory, which are relevant for understanding this work. For in depth knowledge on this subject, please consult [14, 15]. In Chapter 2, we introduce different types of correlations present in quantum information, which can be resources for different settings of quantum key distribution. In Chapter 3, we introduce different settings of quantum key distribution. In Chapter 4, we outline the basics of intrinsic information and squashed entanglement. We introduce intrinsic steerability in Chapter 5 and intrinsic non-locality in Chapter 6. We then prove upper bounds on one-sided device-independent quantum key distribution and device-independent quantum key distribution in Chapter 7. We finally outline future directions and a few open questions in Chapter 8. Chapter 5 is based on [12], and Chapters 6 and 7 are based on [13].

1.2. Quantum states

In this section, we outline the basic definitions and concepts of quantum information that we use in this thesis. A Hilbert space is denoted by $\mathcal{H}$, and vectors in Hilbert space are denoted by a ket $|\psi\rangle$. The formal definition of Hilbert space is given as follows:

Definition 1 (Hilbert Space) *A Hilbert space is an inner product vector space* [1] *over complex numbers* $\mathbb{C}$*. The inner product maps a pair of vectors* $|\psi\rangle$ *and* $|\phi\rangle$ *to an element of* $\mathbb{C}$*, and has the following properties:*

- *Positivity:* $\langle\psi|\psi\rangle \geq 0$*. The equality is satisfied if and only if* $|\psi\rangle = 0$*.*
- *Linearity:* $\langle\phi|\lambda_1\psi_1 + \lambda_2\psi_2\rangle = \lambda_1 \langle\phi|\psi_1\rangle + \lambda_2 \langle\phi|\psi_2\rangle$*, where* $\lambda_1, \lambda_2 \in \mathbb{C}$*, and* $|\psi_{1,2}\rangle$*,* $|\phi\rangle$ *are vectors in* $\mathcal{H}$*.*
- *Skew symmetry:* $\langle\phi|\psi\rangle = \overline{\langle\psi|\phi\rangle}$*, where* $\bar{c}$ *denotes complex conjugation of a complex number* c*.*

Pure quantum states $|\phi\rangle$ are vectors belonging to the Hilbert space $\mathcal{H}$ with norm one, corresponding to the normalization condition $\|\psi\|_2 = 1$, where $\|\psi\|_2 = \sqrt{\langle\psi|\psi\rangle}$. Another term commonly encountered in quantum information is *mixed states*. As the name suggests, mixed states are a mixture of pure quantum states. This decomposition of a mixed state into pure states need not be unique.

We represent a quantum state, pure or mixed, as a density operator, defined as follows:

Definition 2 (Density operators) *A density operator* ρ *acting on* $\mathcal{H}$ *is a positive semidefinite, Hermitian operator with trace equal to one. This means* $\rho = \rho^\dagger$*,* $\mathrm{Tr}\left[\rho\right] = 1$*, and* $\rho \geq 0$*.*

[1] Since we are dealing with finite-dimensional systems, we stick to the simpler definition.

The inequality, $\rho \geq 0$, is an operator inequality, which implies that $\langle\psi| \rho |\psi\rangle \geq 0$, for all $|\psi\rangle \in \mathcal{H}$. We define the set of density operators over $\mathcal{H}$ as $\mathcal{S}(\mathcal{H})$. We define the set of positive semidefinite operators over $\mathcal{H}$ as $\mathcal{P}(\mathcal{H})$. In this thesis, we deal with finite-dimensional systems. Therefore, we can represent density operators as finite-dimensional Hermitian matrices.

One can also consider bipartite states shared between Alice and Bob. For this, consider two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. We can consider a tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ of these two vector spaces to define a larger Hilbert space $\mathcal{H}_{AB}$. The density operators on $\mathcal{H}_{AB}$ define the possible bipartite states shared between Alice and Bob.

Let ρ_{AB} be a bipartite state shared between Alice and Bob. One natural question to ask is the following: how can one define local states of Alice and Bob? This can be answered by taking the marginal of the state ρ_{AB}. That is, Alice's local state ρ_A is given in terms of the partial trace as $\mathrm{Tr}_B\left[\rho_{AB}\right]$. In a similar way, Bob's local state ρ_B is given by $\mathrm{Tr}_A\left[\rho_{AB}\right]$. With this, we can define *reduced density operators* as follows:

Definition 3 (Reduced density operators) *Given a bipartite state $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$, we can define its reduced density operator $\rho_A \in \mathcal{S}(\mathcal{H}_A)$ as*

$$\rho_A = \mathrm{Tr}_B\left[\rho_{AB}\right] = \sum_i (\mathbb{I}_A \otimes \langle i|_B)\rho_{AB}(\mathbb{I}_A \otimes |i\rangle_B), \tag{1.1}$$

where $\{|i\rangle_B\}$ is an orthonormal basis for $\mathcal{H}_B$ and $\mathbb{I}_A$ is the identity operator on $\mathcal{H}_A$.

Let us now suppose that Alice has a mixed state ρ_A. Then, it is always possible to enlarge the Hilbert space $\mathcal{H}_A$ to $\mathcal{H}_{AB}$, and embed the state ρ_A in a pure state ψ^ρ_{AB}, such that $\rho_A = \mathrm{Tr}_B\left[\psi^\rho_{AB}\right]$. The pure state ψ^ρ_{AB} is called a *purification* of the state ρ_A.

Example 4 *Let $\rho_A = \sum_i p_i |i\rangle\langle i|_A$. Consider a Hilbert space $\mathcal{H}_B$ isomorphic to $\mathcal{H}_A$,* *and let $\{|e_i\rangle_B\}_i$ be an orthonormal basis on $\mathcal{H}_B$. Then a particular purification of ρ_A is $|\psi^\rho\rangle_{AB} = \sum_i \sqrt{p_i}\,|i\rangle_A |e_i\rangle_B$. It is easy to see that if we trace out the B system, we obtain the marginal state ρ_A on the A system.*

Now, notice that we defined the purification by considering an arbitrary orthonormal basis $\{|e_i\rangle\}_i$ on $\mathcal{H}_B$. We know that a unitary transformation on $\{|e_i\rangle\}_i$ will define a different orthonormal basis on $\mathcal{H}_B$. This hints towards the possibility that the purification of a state is not unique and that we can access all purifications of ρ_A by an isometry acting on the purifying system. This statement is formalized in terms of non-uniqueness of purification of states as follows:

Definition 5 (Non-uniqueness of purification) *Consider two purifications $|\psi^\rho\rangle_{AB}$ and $|\phi^\rho\rangle_{AB}$ of the state ρ_A. Then, there exists an isometry V_B on $\mathcal{H}_B$ such that $|\phi^\rho\rangle_{AB} = (\mathbb{I}_A \otimes V_B)\,|\psi^\rho\rangle_{AB}$.*

We can also prove that, to construct a purification of ρ_A, it suffices to consider a purifying system B with $\dim\mathcal{H}_B = \dim\mathcal{H}_A$. [2]

Two states commonly referred to in this thesis are the maximally entangled states and maximally mixed states. The maximally entangled state Φ^d_{AB} is defined on the Hilbert space $\mathcal{H}_{AB}$, with $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = d$ as follows:

Definition 6 (Maximally entangled state) *The maximally entangled state on $\mathcal{H}_{AB}$, with dim$(\mathcal{H}_A)$ = dim$(\mathcal{H}_B) = d$, is defined as*

$$\Phi^d_{AB} = \frac{1}{d}\sum_{i,j=1}^{d} |ii\rangle\langle jj|_{AB}\,. \tag{1.2}$$

[2] This is an important concept in quantum mechanics and captures the essence of why it is possible to obtain information-theoretic security in quantum key distribution. We can always purify the system that is held by the local parties and this gives us a tool to bound the information held by an arbitrary Eavesdropper.

Definition 7 (Maximally mixed state) *The maximally mixed state on $\mathcal{H}_A$ is denoted by π_A and is defined as*

$$\pi_A^d = \frac{1}{d}\sum_i |i\rangle\langle i|_A \,. \tag{1.3}$$

The maximally entangled state Φ_{AB}^d purifies the maximally mixed state π_A^d.

1.3. Quantum operations

1.3.1. Evolution

The evolution of a quantum state $\rho_R \in \mathcal{S}(\mathcal{H}_R)$ to $\rho_A \in \mathcal{S}(\mathcal{H}_A)$ is described by a *quantum channel.* Let us denote this evolution by a map $\mathcal{N}_{R\to A}$ acting on the space of density operators. The mathematical constraints imposed on an evolution due to physical considerations are as follows:

- First, we expect $\mathcal{N}_{R\to A}$ to be a linear map from $\mathcal{S}(\mathcal{H}_R)$ to $\mathcal{S}(\mathcal{H}_A)$. That is,

  $$\mathcal{N}(\alpha\rho + \beta\sigma) = \alpha\mathcal{N}(\rho) + \beta\mathcal{N}(\sigma), \tag{1.4}$$

  where α, $\beta \in \mathbb{C}$, and $\rho, \sigma \in \mathcal{S}(\mathcal{H}_R)$.

- Second, we expect a quantum channel to transform a density operator to a density operator. Therefore, we restrict $\mathcal{N}_{R\to A}$ to be a positive map. However, we need more than the positivity condition. Suppose that we have a bipartite state ρ_{RB}, and Alice's local system is acted upon by $\mathcal{N}_{R\to A}$. Then the overall transformation of the state ρ_{RB} is given by $\sigma_{AB} = (\mathcal{N}_{R\to A} \otimes \mathrm{id})(\rho_{RB})$. Now, we expect σ_{AB} to be a density operator as well. Therefore, we need the linear map $\mathcal{N}_{R\to A}$ to be a completely positive map, which is a strictly stronger condition than a positive map. [3]

[3] A simple example of a map that is positive but not completely positive is the transpose map.

- We also require $\mathcal{N}_{R\to A}$ to be trace preserving. This again is based on the reasoning that a quantum channel should transform a density operator to a density operator.

Succinctly, quantum channels can be defined as follows:

Definition 8 (Quantum channels) *A linear map $\mathcal{N}_{R\to A}$ is a quantum channel if it is*

- *Completely positive: for any $\rho_{RB} \in \mathcal{S}\left(\mathcal{H}_{RB}\right)$, $(\mathcal{N}_{R\to A} \otimes \mathrm{id})(\rho_{RB}) \geq 0$, where* id *is the identity map on $\mathcal{S}(\mathcal{H}_B)$.*
- *Trace preserving: for any $\rho_A \in \mathcal{S}(\mathcal{H}_A)$, $\mathrm{Tr}\left[\mathcal{N}_{R\to A}(\rho_R)\right] = \mathrm{Tr}\left[\rho_R\right]$.*

Quantum channels are also referred to as CPTP maps, which stands for completely positive trace-preserving maps.

1.3.2. Quantum measurements

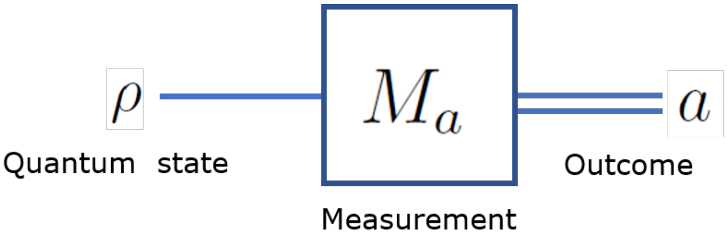

Figure 1.1. In this figure, we depict a quantum measurement. The measurement apparatus performs a measurement on quantum state ρ, with the measurement described by a set of positive semidefinite operators, and outputs a classical outcome a.

To obtain classical information about a quantum state, one performs a *quantum measurement*. This measurement may correspond to information about properties such as position, momentum, or spin of a quantum state. These properties are formally known as observables. When we perform a quantum measurement on a quan-

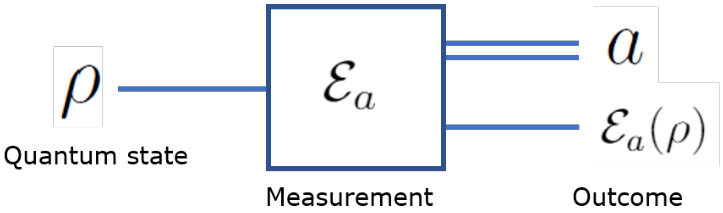

Figure 1.2. In this figure, we depict a quantum instrument. A quantum instrument takes in a quantum state as an input and gives out a classical outcome a with probability $p(a)$ and a quantum state $\mathcal{E}_a(\rho_A)/p(a)$.

tum state, the outcome a corresponds to a specific value of the observable, obtained with probability $p(a)$.

The most general way of formulating a measurement device is with a positive-operator-valued measure (POVM). A POVM is defined as a set of positive semidefinite operators $\{M_a\}_{a\in\mathcal{A}}$ such that $\sum_{a\in\mathcal{A}} M_a = \mathbb{I}_A$, and $a \in \mathcal{A}$ corresponds to the measurement outcome. The positive semi-definite operators M_a need not be orthogonal.

Now, consider a state ρ_A, and let an observer perform a POVM $\{M_a\}_{a\in\mathcal{A}}$. Then, the outcome of the measurement takes a value in $a \in \mathcal{A}$, and each of these outcomes is associated with the positive semidefinite M_a. The probability of observing the outcome a when ρ_A is measured with POVM $\{M_a\}_a$ is given by $\Pr(a) = \operatorname{Tr}[M_a\rho]$. This is known as the Born rule. The POVM does not provide any information about the state after the measurement.

We can impose a further restriction of orthogonality on the POVM elements $\{M_a\}$. That is, we can demand that the following orthogonality constraint hold $M_a M_a' = \delta_{aa'} M_a$. Such a measurement is called as Projective-Valued measure.

1.3.3. Quantum instruments

A quantum instrument is a quantum channel that takes as input a quantum state ρ and as an output gives a classical variable corresponding to a measurement outcome $a \in \mathcal{A}$, and a post-measurement state ρ^a. It can be formally defined as

Definition 9 *A quantum instrument $\mathcal{N}$ is described as*

$$\mathcal{N}(\rho) = \sum_{a\in\mathcal{A}} \mathcal{E}_a(\rho) \otimes |a\rangle\langle a|_A \,, \tag{1.5}$$

where A is a random variable defined on the alphabet $\mathcal{A}$, and the sum map $\sum_a \mathcal{E}_a$ is trace preserving.

1.4. Entropy

Let us consider a random variable X, which takes value $x \in \{0, \ldots, d\}$, where d is some positive integer. Let the probability distribution over the random variable X be given by $p_X(x)$. Suppose that we sample from this probability distribution. *The expected surprisal available with this stochastic process is called entropy.* As an example, consider the following probability distribution: $\{p(0) = 1,\ p(1) = 0\}$. The surprisal that one has by sampling from this probability distribution is equal to zero.

One can quantify the amount of information associated with a random variable by *Shannon entropy*, which is defined as follows:

Definition 10 (Shannon entropy) *Given a random variable X with the probability distribution p, the Shannon entropy of X is defined as*

$$H(p) = -\sum_x p_X(x) \log p_X(x). \tag{1.6}$$

Throughout this thesis, we use log for logarithm with base two. The Shannon entropy is maximum for a uniformly distributed random variable, as the information gain from sampling from this distribution is maximum.

Analogously, one can define the information content in a quantum state ρ_A by von Neumann entropy.

Definition 11 (von Neumann entropy) *The von Neumann entropy of a quantum state ρ_A is defined as*

$$H(A)_\rho = -\operatorname{Tr}\left[\rho \log \rho\right]. \tag{1.7}$$

The notation $H(A)_\rho$ indicates the entropy of system A in the state ρ.

Consider a density operator ρ_A with the spectral decomposition: $\rho_A = \sum_i \lambda_i |i\rangle\langle i|_A$, where λ_i's are eigenvalues of ρ, and $\{|i\rangle_A\}_i$ is an orthonormal basis. Then the von Neumann entropy of ρ is given as

$$H(A)_\rho = -\sum_i \lambda_i \log \lambda_i. \tag{1.8}$$

Since ρ is a density matrix, we know that λ_i's are real and positive. Also, $\sum_i \lambda_i = 1$ (this comes from the normalization condition). Therefore, we can associate λ_i with probabilities and interpret (1.8) in terms of the information-theoretic Shannon entropy.

The von Neumann entropy is a function of the eigenvalues of a quantum state. On applying a unitary operation to the state, the eigenvalues are unchanged. Hence, the von Neumann entropy of a quantum state is invariant under a unitary transformation. The von Neumann entropy is maximum for a maximally mixed state and is equal to zero for pure states. This is analogous to Shannon entropy, which is maximum for the uniform distribution and zero for deterministic distributions.

Definition 12 (Conditional entropy) *Given a quantum state ρ_{AB}, the conditional entropy is defined as*

$$H(A|B)_\rho = H(AB)_\rho - H(B)_\rho, \tag{1.9}$$

where $H(B)_\rho$ is the von Neumann entropy associated with the marginal state ρ_B.

If ρ_{AB} is a classical-quantum state $\rho_{AB} = \sum p_B(b) |b\rangle\langle b|_B \otimes \rho_A^b$, where $\{|b\rangle_B\}_b$ is an orthonormal basis of $\mathcal{H}_B$ and $\rho_A^b \in \mathcal{S}(\mathcal{H}_A)$, then

$$H(A|B)_\rho = \sum_b p_B(b) H(A)_{\rho^b}. \tag{1.10}$$

The above rewriting of conditional entropy is possible only if the state ρ_{AB} is a classical-quantum state.

An important inequality for entropies is

$$H(A)_\rho \geq H(A|B)_\rho. \tag{1.11}$$

This is referred to as "conditioning does not increase entropy."

The conditional entropy $H(Y|X)$ over random variables Y and X is always non-negative. Surprisingly, for a quantum state ρ_{AB}, the quantum conditional entropy $H(A|B)$ is not necessarily positive [16]. For example, consider the conditional entropy of the maximally entangled state Φ_{AB}^2: $H(A|B)_\Phi = -1$.

1.4.1. Mutual information

Another important entropic quantity is *mutual information*, which is defined as follows:

Definition 13 (Mutual information) *Given two random variables X and Y with joint probability distribution $p_{XY}(x, y)$, the mutual information is defined as*

$$I(X;Y)_p = H(X)_p + H(Y)_p - H(XY)_p \tag{1.12}$$

$$= H(X)_p - H(X|Y)_p. \tag{1.13}$$

Intuitively, mutual information is a correlation measure. It quantifies the amount of information that one has about X if one knows Y or vice versa. It is non-negative and symmetric under the exchange of X and Y.

We can also define the mutual information of a quantum state as follows:

Definition 14 *Given a bipartite quantum state ρ_{AB}, its quantum mutual information is defined as*

$$I(A;B)_\rho = H(A)_\rho - H(A|B)_\rho. \tag{1.14}$$

The mutual information information is equal to zero for a product state $\rho_{AB} = \rho_A \otimes \rho_B$. For a maximally entangled state, $H(A|B)_{\Phi^d} = -\log d$ and $I(A;B)_{\Phi^d} = 2\log d$.

Mutual information follows a data processing inequality, which implies that a local channel cannot increase the correlations between spatially separated parties:

Definition 15 (Data processing of mutual information) *The data processing inequality states that*

$$I(A;B)_\rho \geq I(A;B)_\sigma, \tag{1.15}$$

where $\sigma_{AB} = (\mathrm{id} \otimes \mathcal{N})(\rho_{AB})$.

A similar statement can be made if $\mathcal{N}$ acts on the system A. Intuitively, the data processing inequality states that a local evolution of a quantum system cannot increase correlations across the bipartite system.

1.4.2. Conditional mutual information

Another entropic quantity of importance and central to this thesis is conditional mutual information (CMI). For probability distributions, conditional mutual information is defined as follows:

Definition 16 (Conditional mutual information) *Consider three random vari-*

able X, Y, and Z. Then conditional mutual information is defined as

$$\begin{aligned} I(X;Y|Z)_p &= H(XZ)_p + H(YZ)_p - H(XYZ)_p - H(Z)_p && (1.16)\\ &= H(X|Z)_p - H(X|YZ)_p. && (1.17)\end{aligned}$$

Operationally, we can think of conditional mutual information as the amount of uncertainty between X and Y given Z. For example, suppose that X and Y are independent uniform random variables, and $Z = X$. Then, the mutual information $I(X;Y) = \log d$, while the conditional mutual information is $I(X;Y|Z) = 0$.

Unlike conditional entropy, conditioning can either increase or decrease the mutual information between two variables. That means

$$I(X;Y|Z) \ngeq I(X;Y), \qquad \text{and} \qquad I(X;Y|Z) \nleq I(X;Y). \tag{1.18}$$

First, let us illustrate that conditioning can decrease the mutual information with with the following example: X, Y, and Z form a Markov chain $X \to Y \to Z$. Then, we obtain the following equalities by invoking the chain rule of mutual information:

$$\begin{aligned} I(X;YZ) &= I(X;Z) + I(X;Y|Z) && (1.19)\\ &= I(X;Y) + I(X;Z|Y). && (1.20)\end{aligned}$$

For a short Markov chain, the condition $I(X;Z|Y) = 0$ holds, which implies that $I(X;Y|Z) \leq I(X;Y)$. Another example to showcase the above inequality is as follows: Let X be uniformly random and suppose that $X = Y = Z$. Then $I(X;Y) = \log d$, and $I(X;Y|Z) = 0$, implying $I(X;Y|Z) \leq I(X;Y)$.

Now consider the following example: Let X and Y be two Bernoulli random variables, and let $Z = X \oplus Y$. $I(X;Y) = 0$; however $I(X;Y|Z) = 1$. This shows that conditioning can increase the mutual information. That is $I(X;Y|Z) \geq I(X;Y)$. Now, suppose that we process Z according to the stochastic map $p_{\overline{Z}|Z}(\overline{z}|z) = p_{\overline{Z}|Z}(\overline{z})$. Then $I(X;Y|\overline{Z}) \leq I(X;Y|Z)$. That is, a local map on the conditioning variable can decrease the conditional mutual information.

Analogous to conditional mutual information, we can define quantum conditional mutual information as follows:

Definition 17 (Quantum conditional mutual information) *Let $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_{ABC})$ Then the conditional mutual information is defined as*

$$I(A;B|C) = H(AC)_\rho + H(BC)_\rho - H(ABC)_\rho - H(C)_\rho. \tag{1.21}$$

Some important properties of conditional mutual information that we use in this thesis are as follows:

- **Chain rule**: Mutual information can be expressed as

$$I(A;BC)_\rho = I(A;B)_\rho + I(A;C|B)_\rho. \tag{1.22}$$

- **Non-Negativity of QCMI**: Quantum conditional mutual information is non-negative for all quantum states. For a state ρ_{ABC}, the following inequality holds

$$I(A;B|C)_\rho \geq 0. \tag{1.23}$$

  This inequality is equivalent to strong subadditivity of quantum entropy [17].

- Let ρ_{ABC} be a classical-quantum state such that $\rho_{ABC} = \sum_c p(c)\, |c\rangle\!\langle c| \otimes \rho_{AB}^c$, then

$$I(A;B|C)_\rho = \sum_c p_C(c) I(A;B)_{\rho^c}. \tag{1.24}$$

1.4.3. Relative entropy

Relative entropy or Kullback-Leibler divergence is an entropic quantity that is a quantifier of "distance" between probability distributions. It is a measure of how far a given probability distribution is to another probability distribution. We define relative entropy as follows:

Definition 18 (Relative entropy) *Let p and q be probability distributions defined on the alphabet $\mathcal{X}$. Then the relative entropy of p and q is defined as follows:*

$$\begin{cases} D\left(p\|q\right) \equiv \sum_{x\in\mathcal{X}} p_X(x) \log\left[\frac{p_X(x)}{q_X(x)}\right] & \textit{if } \operatorname{supp}(p) \subseteq \operatorname{supp}(q) \\ +\infty & \textit{else} \end{cases} \tag{1.25}$$

As is evident from the definition, relative entropy is not symmetric under the exchange of p and q. Relative entropy also does not satisfy the triangle inequality. Hence, relative entropy is not a metric.

We now define quantum relative entropy, which is a natural extension of classical relative entropy as [18]. [4]

Definition 19 *Let $\rho \in \mathcal{S}(\mathcal{H})$ and $\sigma \in \mathcal{P}(\mathcal{H})$. Then the relative entropy of ρ and σ*

[4] This is one particular way of defining a quantum counterpart of relative entropy. It is, in fact, possible to define relative entropy in infinitely different ways such that for probability distributions the formula collapses to the classical one. However, the standard definition we use is justified by its operational meaning in quantum hypothesis testing [19].

is defined as

$$\begin{cases} D\left(\rho\|\sigma\right) \equiv \mathrm{Tr}\left[\rho\left[\log\rho - \log\sigma\right]\right] & \textit{if} \quad \mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma) \\ \quad + \infty & else \end{cases} \tag{1.26}$$

As with the classical relative entropy, quantum relative entropy is not symmetric under the exchange of ρ and σ and does not satisfy the triangle inequality. Therefore, relative entropy is not a distance measure. Some important properties of relative entropy that we invoke in this thesis are as follows:

- **Monotonicity**: Let $\rho \in \mathcal{S}(\mathcal{H})$ and $\sigma \in \mathcal{P}(\mathcal{H})$, and $\mathcal{N}$ be a quantum channel. Then, the relative entropy of ρ and σ can only decrease or remain the same. That is [20],

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)). \tag{1.27}$$

- **Non-negativity**: Let $\rho \in \mathcal{S}(\mathcal{H})$ and $\sigma \in \mathcal{P}(\mathcal{H})$ and $\mathrm{Tr}\left[\sigma\right] \leq 1$. Then the relative entropy is non-negative. That is,

$$D\left(\rho\|\sigma\right) \geq 0, \tag{1.28}$$

  with equality if and only if $\rho = \sigma$. This is also called as Klein's inequality.

- **Isometric invariance**: Relative entropy is invariant under the action of unitaries. That is,

$$D\left(\rho\|\sigma\right) = D\left(U\rho U^{\dagger}\|U\sigma U^{\dagger}\right). \tag{1.29}$$

1.5. Distance measures

Often in quantum information, we need to define a metric between two quantum states. Two important metrics used in quantum information are the *trace norm* and *fidelity.*

To define trace distance, we need the definition of the Schatten 1-norm $\|M\|_1$ of a linear bounded operator M. Schatten 1-norm, also known as trace norm, is defined as follows:

Definition 20 (Trace Norm) *The trace norm of a linear bounded operator M is given as follows:*

$$\|M\|_1 = \mathrm{Tr}\left[|M|\right], \tag{1.30}$$

where $|M| = \sqrt{M^\dagger M}$. Alternatively,

$$\|M\|_1 = \sum_i \lambda_i, \tag{1.31}$$

where λ_i are the singular values of M or the eigenvalues of $\sqrt{M^\dagger M}$.

The trace norm has some desirable properties such as non-negativity, triangle inequality, and isometric invariance. That is, for a linear bounded operator M acting on $\mathcal{H}$

- Non-negativity: $\|M\|_1 \geq 0$.
- Triangle inequality: $\|M + N\|_1 \leq \|M\|_1 + \|N\|_1$, where N is a linear bounded operator on $\mathcal{H}$.
- Isometric invariance: $\|UMU^\dagger\|_1 = \|M\|_1$, where U is a unitary operator on $\mathcal{H}$.

From trace norm, one can induce trace distance, which we define next.

1.5.1. Trace distance

The trace distance is an important metric used in quantum information to define distance between two quantum states. Trace distance is defined as follows:

Definition 21 *Given any two positive semi-definite operators M and N, the trace distance between them is defined as follows:*

$$\|M - N\|_1 = \mathrm{Tr}\left[|M - N|\right]. \tag{1.32}$$

Trace distance has some nice properties such as triangle inequality and monotonicity. These properties are given as follows:

- **Triangle inequality**: Let ρ, σ, and τ be quantum states. Then, the trace distance follows the inequality

$$\|\rho - \sigma\|_1 \leq \|\rho - \tau\|_1 + \|\sigma - \tau\|_1. \tag{1.33}$$

- **Monotonicity of trace distance**: This property states that the action of quantum channel on quantum states decreases the trace distance between quantum states. That is,

$$\|\mathcal{N}(\rho) - \mathcal{N}(\sigma)\|_1 \leq \|\rho - \sigma\|_1. \tag{1.34}$$

1.5.2. Fidelity

Let $|\psi\rangle$ and $|\phi\rangle$ be two pure states. Then the fidelity between the pure states is defined as the square of the overlap of two vectors. That is,

$$F(\psi, \phi) = |\langle\psi|\phi\rangle|^2. \tag{1.35}$$

We see that the fidelity is equal to one if $|\psi\rangle = |\phi\rangle$, due to the normalization condition. The fidelity is equal to zero if the states are orthogonal to each other.

We extend the definition of fidelity to mixed states according to the approach of Uhlmann [21], which is slightly complicated. Consider two mixed state ρ and σ. How do we define the overlap between two mixed states? One possibility is to consider the overlap of respective purifications of the mixed states. Let $|\phi^\rho\rangle_{RA}$ and $|\psi^\sigma\rangle_{RA}$ be respective purifications of ρ and σ with R being the purifying system isomorphic to A. Then one can define Uhlmann's fidelity as the maximum overlap between their respective purifications. That is

$$F(\rho_A, \sigma_A) = \max_{|\phi^\rho\rangle_{RA}, |\phi^\sigma\rangle_{RA}} |\langle\psi^\rho_{RA}|\psi^\sigma_{RA}\rangle|^2. \tag{1.36}$$

We can use the unitary equivalence of purifications specified in Definition 5 to rewrite the above definition of fidelity as

$$F(\rho_A, \sigma_A) = \max_{U_R} |\langle\psi^\rho_{RA}|(U_R \otimes I_A)\psi^\sigma_{RA}\rangle|^2. \tag{1.37}$$

Some important properties of fidelity are as follows:

- Isometric invariance: Fidelity is invariant under unitaries, which can be stated as
$$F(\rho, \sigma) = F(U\rho U^\dagger, U\sigma U^\dagger). \tag{1.38}$$

- Monotonicity under quantum channels: This property can be expressed as the following inequality:
$$F(\rho, \sigma) \leq F(\mathcal{N}(\rho), \mathcal{N}(\sigma)). \tag{1.39}$$

Definition 22 *Fidelity between two states ρ and σ can be used to obtain an upper* *and a lower bound on the trace distance as follows [22]:*

$$1-\sqrt{F(\rho,\sigma)} \leq \frac{1}{2}\|\rho-\sigma\|_1 \leq \sqrt{1-F(\rho,\sigma)}. \tag{1.40}$$

Chapter 2
Correlations

In this chapter, we introduce various types of correlations relevant to quantum technologies. In quantum information, correlations can exist in objects such as conditional probability distributions, assemblages (we define this term below), and in bipartite quantum states. Within each object, we can classify the correlations observed as classical correlations, quantum correlations, or no-signaling correlations. In the following sections, we define each of the aforementioned objects and later analyze the various correlations that can be present within each object.

Correlation is a term, often used in statistics, to quantify how much information is common between two or more random variables. Given two random variables X and Y such that $X = Y$, we say that these random variables are perfectly correlated. Once we know the value of X, we also know the value of Y. Given two correlated random variables, knowing one of them gives some information about the other random variable. We can lift the concept of correlations of random variables to objects such as bipartite states and assemblages.

2.1. Bipartite states, assemblages, and distributions

In this section, we define various objects relevant in quantum information theory and then explore the correlations in these objects. We first start with correlations in bipartite states.

2.1.1. Correlations in bipartite quantum systems

The objects under consideration in this section are bipartite states $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ shared between Alice and Bob.

Consider $\rho_{AB} = \rho_A \otimes \rho_B$, where $\rho_A \in \mathcal{S}(\mathcal{H}_A)$ and $\rho_B \in \mathcal{S}(\mathcal{H}_B)$ are some states of systems A and B respectively. These states are called as *product states.* These can be considered a generalization of product probability distributions corresponding to

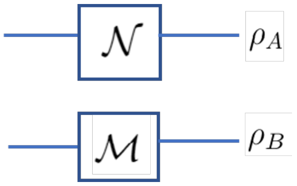

Figure 2.1. Product state.

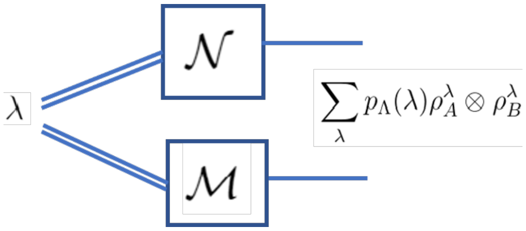
λ
$\mathcal{N}$
$\mathcal{M}$
$\sum_\lambda p_\Lambda(\lambda)\rho_A^\lambda \otimes \rho_B^\lambda$

Figure 2.2. Separable state

independent random variables. To prepare these states, Alice and Bob can perform local classical-quantum channels $\mathcal{N}_{A\to A} \otimes \mathcal{M}_{B\to B}$ on independent random variables. Next, consider separable states defined as follows:

Definition 23 (Separable states) *A state $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ that can be expressed as*

$$\rho_{AB} = \sum_\lambda p_\Lambda(\lambda)\rho_A^\lambda \otimes \rho_B^\lambda, \tag{2.1}$$

with $p_\Lambda(\lambda) \geq 0$, $\sum_\lambda p_\Lambda(\lambda) = 1$, and $\rho_A^\lambda \in \mathcal{S}(\mathcal{H}_A)$, and $\rho_B^\lambda \in \mathcal{S}(\mathcal{H}_B)$ is known as a separable state.

Suppose that Alice and Bob share a local hidden variable Λ or are allowed to communicate classically. Alice performs a quantum instrument, communicates the classical result to Bob, who applies a quantum instrument to his part of the state. This process continues for a finite number of rounds to generate a *separable state.*

The set of separable states is denoted by $\mathrm{SEP}(A\!:\!B)$. By definition, the set of separable states contains the set of product states.

A little reflection hints towards the existence of states that are not separable states. Such states are called as entangled states. That is, *the states that are not separable are called as entangled states.* In a nutshell, *entangled states have correlations that cannot be explained by underlying classical shared randomness.* This definition can also be extended to multipartite states.

Non-classical correlations in entangled states have been exploited for various quantum technologies such as quantum key distribution [23], quantum metrology [24], teleportation, among others. For this thesis, we are interested in the application of quantum entanglement in quantum key distribution. For details on quantum entanglement, please consult [25].

2.1.2. Correlation in assemblages

Suppose that Alice and Bob share a bipartite state ρ_{AB}. These states could have been distributed by an external party not trusted by Alice and Bob. Suppose that they have no information about the underlying state ρ_{AB} distributed between them. In principle, Alice and Bob can perform quantum tomography on their unknown systems to gain information about ρ_{AB}. Tomography involves performing measurements on the underlying system to obtain a characterization of the state ρ_{AB}.

What happens if Alice does not trust the measurement being performed? In this case, tomography is not possible. Instead, we consider the setting of one-sided device-independence. This untrusted measurement is modelled as a *black-box device.* The device takes a classical symbol $x \in \mathcal{X}$ as input, where $\mathcal{X}$ denotes a finite set of quantum measurements, performs some operation on the unknown state ρ_A, and gives a classical output $a \in \mathcal{A}$, with $\mathcal{A}$ denoting a finite set of measurement outcomes. Then the only way that Alice can interact with her unknown marginal state is through classical inputs and classical outputs. Let us characterize the unknown measurement performed by the device as $\{\Lambda_x^a\}_a$. During each interaction with the device, Alice obtains a, x and Bob obtains a correlated quantum state $\rho_B^{a,x}$. The state $\rho_B^{a,x}$ is defined as

$$\rho_B^{a,x} = \frac{\mathrm{Tr}_A\left[\left(\Lambda_x^a \otimes \mathbb{I}_B\right)\left(\rho_{AB}\right)\right]}{\mathrm{Tr}\left[\Lambda_x^a\left(\rho_A\right)\right]}, \tag{2.2}$$

where $\{\Lambda_x^a\}$ is an unknown POVM performed by the device, with $\Lambda_x^a \geq 0$ for all $a \in \mathcal{A}$ and $\sum_a \Lambda_x^a = I$. Alice and Bob have an assemblage $\left\{p_{\bar{A}|X}(a|x)\rho_B^{a,x}\right\}_{a,x}$, where

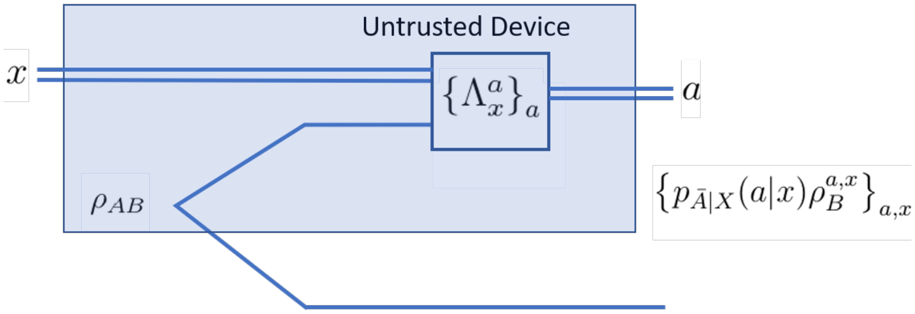

Figure 2.3. An assemblage $\left\{p_{\bar{A}|X}(a|x)\rho_B^{a,x}\right\}_{a,x}$ shared between Alice and Bob. Alice performs an untrusted measurement $\{\Lambda_x^a\}_a$ on her part of the state ρ_{AB}.

$p_{\bar{A}|X}(a|x) = \mathrm{Tr}\left[\Lambda_x^a\left(\rho_A\right)\right]$, and $\bar{A}$ is the random variable associated with the measurement outputs. The objects under consideration in this formalism are *assemblages*, defined as follows:

Definition 24 (Assemblages) *An assemblage consists of the state of Bob's subsystem and the conditional probability of Alice's outcome a (correlated with Bob's state) given the measurement choice x. This is specified as* $\{p_{\bar{A}|X}(a|x), \rho_B^{a,x}\}_{a\in\mathcal{A},x\in\mathcal{X}}$*. The sub-normalized state possessed by Bob is* $\hat{\rho}_B^{a,x} := p_{\bar{A}|X}(a|x)\rho_B^{a,x}$*.*

Consider the following example of assemblages: Suppose that Alice and Bob, unknown to them, share a maximally entangled state Φ_{AB}. Alice has a device that takes in input value $x = 0$ or $x = 1$, performs some measurement on the underlying unknown state, and returns back the output values $a = 0$ or $a = 1$. Let us suppose that if the device receives $x = 0$, then it performs a measurement in the σ_z basis, and if the device receives $x = 1$, then it performs a corresponding measurement in the σ_x basis. With this device, Alice and Bob have the following assemblage:

$$\Big\{\hat{\rho}_B^{a=0,x=0} = \frac{1}{2}|0\rangle\langle 0|_B,\ \hat{\rho}_B^{a=1,x=0} = \frac{1}{2}|1\rangle\langle 1|_B, \\ \hat{\rho}_B^{a=0,x=1} = \frac{1}{2}|+\rangle\langle +|_B,\ \hat{\rho}_B^{a=1,x=1} = \frac{1}{2}|-\rangle\langle -|_B\Big\}. \quad (2.3)$$

We can understand these assemblages from two different perspectives. We can first think of an external party distributing an object to Alice and Bob. Alice's measurement device is prepared by an external party and is not trusted or characterized. Bob's device is trusted or characterized. This perspective is often helpful when we want to think of these objects in the context of quantum key distribution. The second perspective does not involve an external party. We solely consider the operations performed by Alice and Bob and the objects created by these operations. This perspective is relevant when we consider these objects from a resource-theory perspective. Both perspectives are crucial for understanding various properties of assemblages.

We now discuss various types of assemblages: *product assemblages*, *local-hidden-state assemblages*, *steerable assemblages*, *quantum assemblages*, and *no-signaling assemblages*.

Let the underlying state ρ_{AB} be a product state. The device performs an arbitrary measurement $\{\Lambda_x^a\}_a$. The assemblage obtained can be written as $\{\hat{\rho}_B^{a,x}\}_{a,x}$. Alice and Bob can also prepare this assemblage by performing local operations in their lab. In this assemblage, there is no correlation between Alice's probability distribution and the quantum state on Bob's side. Such an assemblage is called as a *product assemblage.*

Suppose that the underlying state ρ_{AB} is a separable state. The device performs an arbitrary measurement $\{\Lambda_x^a\}_a$. The assemblage obtained can be written as follows:

$$\left\{\sum_{\lambda} p_{A|X\Lambda}(a|x,\lambda)\rho_B^{\lambda}\right\}_{a,x} \tag{2.4}$$

To prepare this assemblage, Alice and Bob share some classical randomness λ and then prepare the following assemblage $\rho_B^{a,x} = \sum_{\lambda} p_{A|X,\Lambda}(a|x,\lambda)\rho_B^{\lambda}$. All assemblages having this form are called as *local-hidden-state assemblages* or *unsteerable assemblages*.

Of course there are assemblages that are not a product or do not have a local-hidden state model. They are of interest in quantum technologies, because they share correlations that go beyond classical correlations. Such assemblages exhibit the phenomenon of *quantum steering* and are known as *steerable assemblages*, defined as follows:

Definition 25 (Steerable assemblages) *Assemblages that cannot be written as*

$$\rho_B^{a,x} = \sum_{\lambda} p_{A|X,\Lambda}(a|x,\lambda)\rho_B^{\lambda} \tag{2.5}$$

are called as steerable assemblages.

To obtain steerable assemblages, the underlying state ρ_{AB} shared by Alice and Bob must be an entangled state. Sharing an entangled state is a sufficient condition but not a necessary one to obtain a steerable assemblage. In fact, there exists an entangled state, which for any arbitrary measurement, yields a local-hidden-state assemblage [26]. Quantum states which, upon measurement, yield a steerable assemblage are known as *steerable states*.

Another question of interest here is as follows: given an assemblage $\{\hat{\rho}_B^{a,x}\}_{a,x}$, can we uniquely characterize the quantum state ρ_{AB} and the quantum measurement $\{\Lambda_x^a\}_a$ from which we obtained the assemblage? The quantum state and quantum measurement corresponding to a particular assemblage is called a *quantum strategy*. This has led to another interesting area of literature that deals with the characterization of quantum states in a one-sided device-independent scenario [27, 28]. Certain assemblages, such as the one considered in the example above, have a unique quantum strategy. However, for most assemblages, we can construct a variety of quantum

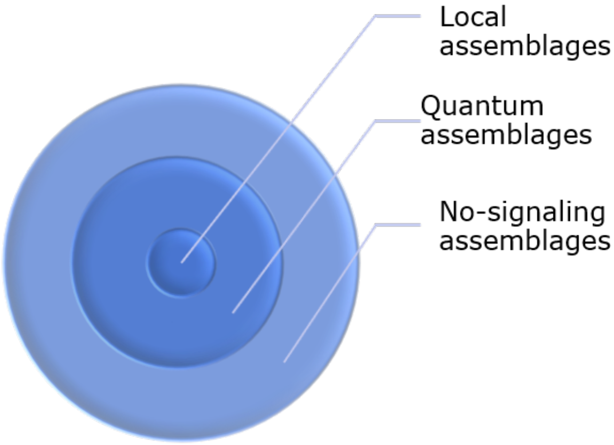

Figure 2.4. Hierarchy of assemblages.

strategies.

Consider a hypothetical scenario in which Alice wants to transmit information to Bob. She encodes the information that she wants to transmit in the measurement performed. Let us suppose that she wants to transmit $x = 0$ to Bob, and therefore she keeps on performing the measurement $x = 0$. She thinks that the local state $\rho_B^{a,x}$ on Bob's side will transmit the desired information to Bob. The problem with the above hypothetical protocol is that Bob will always have the averaged state $\sum_a p(a|x)\rho_B^{a,x}$, since he does not know the measurement result a. Bob's averaged state remains invariant with respect to Alice's measurement choices and is equal to Bob's marginal state. This can be mathematically stated as

$$\sum_a p_{\bar{A}|X}(a|x)\rho_B^{a,x} = \sum_a p_{\bar{A}|X}(a|x')\rho_B^{a,x'} = \rho_B \qquad \forall x, x'. \tag{2.6}$$

This can be easily proved by observing that $p_{\bar{A}|X}(a|x)\rho_B^{a,x} = \mathrm{Tr}_A\left[\left(\Lambda_a^x \otimes I_B\right)(\rho_{AB})\right]$. Then,

$$\sum_a p_{\bar{A}|X}(a|x)\rho_B^{a,x} = \sum_a \mathrm{Tr}_A\left[\left(\Lambda_a^x \otimes I_B\right)(\rho_{AB})\right] = \rho_B. \tag{2.7}$$

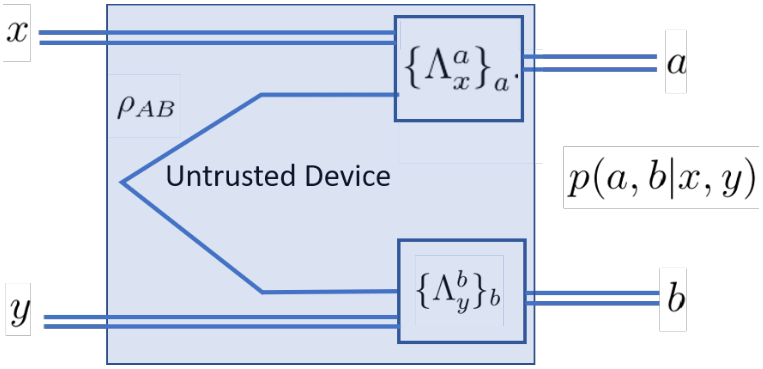

Figure 2.5. A quantum distribution

Equation (2.6) is known as the *no-signaling principle* and is a physical constraint on all assemblages that arise from an underlying quantum state and a quantum measurement. The no-signaling constraint can also be expressed equivalently in terms of conditional mutual information as $I(\bar{A};B|X)_{\hat{\rho}} = 0$ for all input probability distributions $p(x)$.

In the above discussion, we assumed that the assemblage has an underlying quantum strategy. Such assemblages are called as *quantum assemblages*. We also proved that all quantum assemblages fulfill the no-signaling constraints. Now, let us start with an assemblage that fulfills the no-signaling constraints. Such assemblages are called as no-signaling assemblages. Is it always possible to find an underlying quantum state and a POVM for a no-signaling assemblage? It is indeed possible to find a quantum strategy for every bipartite no-signaling assemblage, as proven in [29]. However, for tripartite assemblages, it has been proven in [29] that there exist assemblages that are no-signaling, yet have no underlying quantum strategy. The set of unsteerable assemblages is contained in the set of quantum assemblages, which in turn is contained in the set of no-signaling assemblages. This is depicted in Figure 2.4.

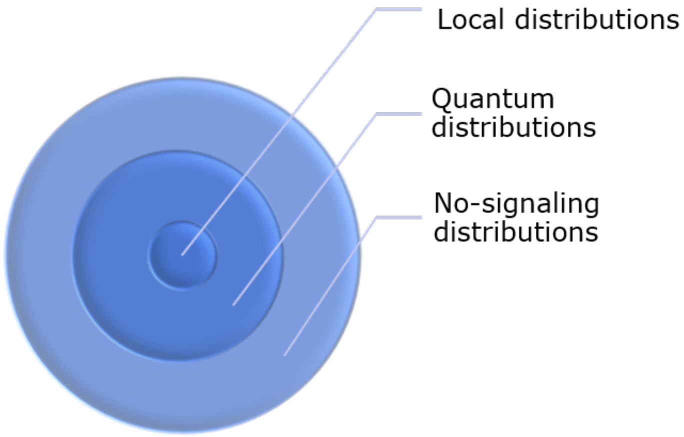

Figure 2.6. Hierarchy of distributions

2.1.3. Correlation in bipartite probability distributions

Suppose that Alice and Bob share a bipartite state ρ_{AB}. In this setting, both Alice and Bob do not trust their measurement devices. Therefore, we can think of Alice and Bob sharing a two-component black box, which takes in two inputs and gives out two outputs. Alice's component takes in an input letter $x \in \mathcal{X}$ and outputs $a \in \mathcal{A}$. Similarly, Bob's component accepts an input letter $y \in \mathcal{Y}$ and outputs $b \in \mathcal{B}$. Suppose that X and Y are finite sets of quantum measurement choices and $\mathcal{A}$ and $\mathcal{B}$ are finite sets of measurement outcomes. For simplicity, we consider $\mathcal{X} = \mathcal{Y} = [s]$ and $\mathcal{A} = \mathcal{B} = [r]$, , where s and r are natural numbers and $[n] = \{0, \ldots, n-1\}$. The box is characterized by the conditional probability distribution $\{p(a,b|x,y)\}_{a,b\in[r],x,y\in[s]}$. Then the correlations in a conditional probability distribution $\{p(a,b|x,y)\}_{a,b\in[r],x,y\in[s]}$ can be divided as follows according to the constraints that they satisfy:

- **Local distributions**: A local distribution has a local hidden variable (LHV) description written as

$$p(a,b|x,y) = \sum_{\lambda} p_{\Lambda}(\lambda) p(a|x,\lambda) p(b|y,\lambda), \tag{2.8}$$

where Λ is a local hidden variable, $p_\Lambda(\lambda)$ is the probability that the realization λ of the local hidden variable Λ occurs, $p(a|x,\lambda)$ is the probability of obtaining the outcome a given x and λ, and $p(b|y,\lambda)$ is the probability of obtaining the outcome b given y and λ. Let $\mathbf{L}$ denote the set of distributions that can be written as in (2.8). A device characterized by local distributions is also known as a *local box*.

- **Quantum distributions**: The set $\mathbf{Q}$ of quantum distributions corresponds to the set of distributions that can be written as

$$p(a,b|x,y) = \mathrm{Tr}([\Lambda_x^a \otimes \Lambda_y^b]\rho_{AB}), \tag{2.9}$$

where ρ_{AB} is a bipartite quantum state and $\{\Lambda_x^a\}_a$ and $\{\Lambda_y^b\}_b$ are POVMs characterizing Alice's and Bob's measurements with $\Lambda_x^a, \Lambda_y^b \geq 0$ for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$ and $\sum_a \Lambda_x^a = I$ and $\sum_b \Lambda_y^b = I$.

- **No-signaling distributions**: The set $\mathbf{NS}$ corresponds to the set of distributions that fulfill the following no-signaling principle:

$$\sum_a p(a,b|x,y) = \sum_a p(a,b|x',y) = p(b|y), \quad \forall x,x' \in [s] \text{ and } \quad b \in [r],\, y \in [s]. \tag{2.10}$$

$$\sum_b p(a,b|x,y) = \sum_b p(a,b|x,y') = p(a|x), \quad \forall y,y' \in [s] \text{ and } \quad a \in [r],\, x \in [s]. \tag{2.11}$$

The no-signaling constraints (2.10) and (2.11) can be expressed equivalently in terms of conditional mutual informations, namely

$$\forall p(x,y) \quad I(X;\bar{B}|Y)_p = 0 = I(Y;\bar{A}|X)_p, \tag{2.12}$$

with respect to the joint distribution $p(a,b,x,y) = p(x,y)p(a,b|x,y)$, and where $p(x,y)$ ranges over probability distributions on X and Y.

It is well known that local correlations are strictly contained in the set of quantum correlations, that is, $\mathbf{L} \subset \mathbf{Q}$. Since the correlations in $\mathbf{Q}$ fulfill the constraints in (2.10) and (2.11), we have that $\mathbf{Q} \subset \mathbf{NS}$ (see Definition 26). This is depicted in Figure 2.6. For more details on correlations, please refer to [30].

Distributions $p(a,b|x,y)$ that are not in $\mathbf{L}$ are known as *non-local distributions.* Quantum states such that there exists at least one arbitrary measurement, which results in a non-local distribution are called as *non-local states.*

An example of a correlation that belongs to the no-signaling correlations, but not the quantum correlations, is a Popescu-Rohrlich (PR) box [31], which is defined as follows:

Definition 26 (PR box) *A PR box is a device corresponding to the following correlation $p(a,b|x,y)$:*

$$\begin{aligned} p(0,0|x,y) = p(1,1|x,y) = \frac{1}{2} \quad &\text{for} \quad (x,y) \neq (1,1), \\ p(0,1|x,y) = p(1,0|x,y) = \frac{1}{2} \quad &\text{for} \quad (x,y) = (1,1), \end{aligned} \tag{2.13}$$

while $p(a,b|x,y) = 0$ for all other quadruples. This correlation is no-signaling between Alice and Bob, as defined in (2.10) *and* (2.11)*.*

2.2. Resource-theoretic framework for correlations

Over the past few years, quantum resource theories have been developed to study various phenomena in quantum information; see [32] for a review. A number of quantum resources, including coherence [33, 34], entanglement [25], asymmetric distinguishibility [35], non-locality [36], steering [37], among others, have been explored from this perspective.

In a resource theory perspective, we define three main ingredients that are intimately related. In the resource theory approach, one defines the set of free objects. These are generally objects that are not useful for a particular task or are easy to create. Here, we consider three different resource theories: resource theory of entanglement, resource theory of steering, and resource theory of non-locality. For these resource theories, the free objects are classical correlations shared between Alice and Bob. This includes separable states, local hidden assemblages, and local hidden variable distributions.

The second ingredient of resource theories is the set of restricted free operations [1]. These operations leave the set of free states invariant. Free operations can be thought of as operations that do not create the resource, and hence can be implemented freely by Alice and Bob. [2]

The free objects and free operations are intimately related. One could first fix the free resource, hence fixing the set of free operations. One can then choose the restricted set of free operations as a set of operations that leave the set of free objects invariant. Alternatively, one could first consider physical constraints that define the free operations, and then define the free objects based on these constraints. In the

[1] We use "restricted" since this need not be the full set of operations under which the set of free states is invariant.

[2] Fixing the set of free objects fixes the set of free operations but not the set of restricted free operations. Often, the mathematical structure of the free operations is not known, and considering the set of free restricted operations is particularly useful.

Table 2.1. Free and resource objects in resource theories.

Objects	**Free objects**	**Resource objects**
Bipartite states ρ_{AB}	Separable states	Entangled states
Assemblages $\hat{\rho}_B^{a,x}$	Local assemblages	Steerable assemblages
Distributions $p(a,b\|x,y)$	Local distributions	Non-local distributions

resource theories considered here, we take the former approach. [3]

The third ingredient is the resource object. One can think of these objects as resources that allow us to perform interesting tasks that might not be possible with the free objects. For the purpose of this thesis, resourceful objects include entangled states, steerable assemblages, and quantum non-local distributions; see Table 2.1. These resourceful objects are vital for quantum key distribution [3, 23], one-sided device-independent quantum key distribution [38], and device-independent quantum key distribution [39], respectively.

The next important question to consider in resource theories is the quantification of the amount of resource held in different objects. How do we quantify a resource? One way to quantify a resource is to employ a pseudo-distance measure between the resource object and the set of free objects. A higher distance to the set of free objects corresponds to higher resource content of the object. One obvious way to construct a resource quantifier $r(\rho)$ of the resource object ρ is as follows:

$$r(\rho) = \min_{\sigma \in \mathrm{F}} d(\rho, \sigma), \tag{2.14}$$

where ρ is a resource, σ is a free resource, d is a distance measure such as trace distance, and F is the set of free objects. For a pseudo-distance measure, relative

[3] This in contrast to the viewpoint taken in [32], where the free operations are the starting point.

entropy has been used in the literature.

Another way to quantify a resource is with robustness. This quantifies the amount of noise that the resource object can tolerate before turning it into a free object.

One can also take a more general approach to defining a quantifier in a resource theory. Consider a function $f : \rho \rightarrow \mathbb{R}$. This function takes in the resource object and outputs a positive real number associated to it. We demand that, to be a quantifier, the function needs to have the following properties:

- *Faithfulness*: A resource quantifier $f(\rho) = 0$, if and only if $\rho \in \mathrm{F}$. That is, the measure should be equal to zero if the object is a free resource. We also demand that if the function evaluates to zero, then the object should be a free object. This constraint is known as faithfulness of the quantifier. [4]

- *Monotonicity*: The quantifier should be monotonically decreasing under the action of free operations. This means that if one starts with an object ρ, and applies a free operation to this object to obtain $\sigma = \mathcal{N}(\rho)$, where $\mathcal{N}$ is a free operation, then $f(\rho) \geq f(\sigma)$. This is in line with the intuition of free operations, which is that the free operations should not increase the amount of resource.

If a function follows the above two properties, then the function is a quantifier for resourcefulness in the object. We can also make the following demands for a quantifier:

- Convexity: Given $\rho = \sum_i p_i \rho^i$, then $f(\rho) \leq \sum_i p_i f(\rho^i)$, for $p_i \geq 0$, and $\sum_i p_i = 1$.

- Continuity: If ρ, γ are objects satisfying $\|\rho - \gamma\|_1 \leq \varepsilon$, then $|f(\rho) - f(\gamma)| \leq g_1(\varepsilon) + g_2(\varepsilon \log d)$,

[4] This condition is sometimes relaxed just to demand that the resource quantifier should evaluate to zero for a free object. The converse statement need not be true. For example, log negativity, an entanglement quantifier, evaluates to zero for some entangled states.

Table 2.2. Resource theory framework for correlations

Resource theory	**Free objects**	**Free operations**	**Resource**
Entanglement	Separable states	Separable operations LOCC	Entangled states
Steering	Unsteerable assemblages	1W-LOCC	Steerable assemblages
Non-locality	Local distributions	WPICC	Non-local distributions

where $g_1(\varepsilon), g_2(\varepsilon \log d) \to 0$, as $\varepsilon \to 0$, and d is the dimension of the resource object. For continuous variable systems, we expect d to be replaced by finite mean energy of the object. The above properties are nice to have for a quantifier but not mandatory.

In the following sections, we will discuss the resource-theoretic framework for entanglement, steering, and non-locality. The discussions in the next three sections can be summarized by Table 2.2.

2.2.1. Resource theory of entanglement

Quantum entanglement is an important phenomenon in quantum information theory and has been vital for the development of quantum technologies. It is a uniquely quantum phenomenon, and therefore it makes sense to understand it from a resource theory perspective.

Let us begin with the following question: what are the operations that Alice and Bob can implement that leave the set of separable states invariant? An example of these operations is local operations and classical communication (LOCC) [40]. As the name suggests, LOCC consists of Alice and Bob performing local operations and communicating classically. LOCC operations can be mathematically hard to describe since the number of rounds of communication between Alice and Bob can be large. A particular way to define finite-round r-LOCC between two parties is as r-rounds of recursive one-way LOCC (LOCC$_1$) defined as

- Alice performs a local operation consisting of a quantum instrument on her side which can be given as $\mathcal{E}(\cdot) = \sum_i \mathcal{E}_i(\cdot) \otimes |i\rangle\langle i|$, such that $\sum_i \mathcal{E}_i$ is trace preserving.
- Alice communicates the classical register to Bob, and Bob applies a CPTP map $\mathcal{F}_i$. This whole operation can be mathematically described as $\mathrm{LOCC}_1(\cdot) = \sum_i \mathcal{E}_i \otimes \mathcal{F}_i(\cdot)$.

We also define LOCC$_1$ with classical communication from Bob to Alice in a similar way. Then, an LOCC operation is composed of rounds of LOCC$_1$ from Alice to Bob and then LOCC$_1$ from Bob to Alice.

Now one can ask the following question: are there correlations beyond LOCC that leave the set of separable states invariant? This larger set of operations, which contains the set of LOCC operations, is known as the set of separable operations, defined as follows:

Definition 27 (Separable operations) *Separable operations on $\mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ are CPTP maps with product Kraus operators mathematically formulated as*

$$\Lambda(\rho) = \sum_i (E_i \otimes F_i)\rho(E_i \otimes F_i)^\dagger, \tag{2.15}$$

with $\sum_i E_i^\dagger E_i \otimes F_i^\dagger F_i = \mathbb{I}$.

It was first shown in [41] that there exist maps in the set of separable operators that cannot be written as a finite-round LOCC.

It is now instructive to think of entanglement from a resource-theoretic perspective. We start by fixing the free states as the separable states. The set of operations, which leave set of free states invariant are the separability-preserving operations, which contain the LOCC operations. Since LOCC is physically easier to implement than the set of separable operations, we consider the restricted set of operations as LOCC operations. The set of entangled states are the resource states.

Some important entangled states include Werner states and isotropic states, defined as follows:

Definition 28 (Werner state [42]) *Let A and B be quantum systems, each of dimension d. A Werner state is defined for $p \in [0,1]$ as*

$$W_{AB}^{(p,d)} \equiv (1-p)\,\frac{2}{d\,(d+1)}\Pi_{AB}^{+} + p\frac{2}{d\,(d-1)}\Pi_{AB}^{-}, \tag{2.16}$$

where $\Pi_{AB}^{\pm} \equiv (I_{AB} \pm F_{AB})/2$ are the respective projections onto the symmetric and antisymmetric subspaces of A and B.

Definition 29 (Isotropic state [43]) *An isotropic state $\rho_{AB}^{(t,d)}$ is $U \otimes \bar{U}$-invariant for an arbitrary unitary U, where $\dim(\mathcal{H}_A) = d = \dim(\mathcal{H}_B)$. Such a state can be written in the following form for $t \in [0,1]$:*

$$\rho_{AB}^{(t,d)} = t\Phi_{AB}^{d} + (1-t)\frac{I_{AB} - \Phi_{AB}^{d}}{d^2 - 1}, \tag{2.17}$$

where Φ_{AB}^{d} denotes a maximally entangled state of Schmidt rank d.

There are a number of quantifiers for entangled states such as relative entropy of entanglement [44], squashed entanglement [10], entanglement robustness [45], entanglement of formation [40], log negativity [46, 47], among others. In this thesis, we only concentrate on relative entropy of entanglement and squashed entanglement.

Relative entropy of entanglement is defined as follows:

Definition 30 (Relative entropy of entanglement) *Let ρ_{AB} be a bipartite state. Then the relative entropy of entanglement of ρ_{AB} is given as*

$$E_R(\rho_{AB}) = \inf_{\sigma_{AB} \in \mathrm{SEP}(A:B)} D(\rho_{AB} \| \sigma_{AB}), \tag{2.18}$$

where $\mathrm{SEP}(A:B)$ *is the set of separable states across the partition* $A:B$*.*

Squashed entanglement was defined in [10]. We will recall the definition and intuition behind the measure in Chapter 4.

Can we calculate the amount of entanglement in a state? It turns out that entanglement cost, entanglement of formation, relative entropy of entanglement, squashed entanglement are NP-hard/NP-complete, and hence cannot be computed in polynomial time [48]. This problem is related to the separability problem [49, 50].

Of the entanglement measures stated above, log-negativity can be calculated easily. But this in no way contradicts the hardness of the separability problem due to the fact that it is not a faithful measure. It can be equal to zero even for some entangled states. It is a useful upper bound on distillable entanglement. [5]

2.2.2. Resource theory of steering

Quantum steering was first introduced by Schrödinger in 1935 [51] in order to formalize an argument made by Einstein, Podolsky, and Rosen in [52]. It refers to the following scenario: Alice and Bob share a bipartite quantum state. Alice measures her system, which can have the effect of steering the reduced state on Bob's system, depending on the measurement that she performs. She thus can influence Bob's subsystem without having access to it. However, Bob does not have any knowledge about the influence, nor can he detect it unless Alice communicates the mea-

[5] It is not known if log negativity is an upper bound on distillable key, which is different from distillable entanglement.

surement that she performed and the outcome of the measurement. For example, consider a maximally entangled state shared by Alice and Bob. Alice can measure her system in either the Pauli σ_Z basis or the Pauli σ_X basis. If she measures in the Pauli σ_Z basis, the resulting state of Bob's subsystem is represented as the ensemble $\big\{(\frac{1}{2}, |1\rangle\!\langle 1|), (\frac{1}{2}, |0\rangle\!\langle 0|)\big\}$. Alternatively, if she measures in the Pauli σ_X basis, the state of Bob's subsystem is represented as the ensemble $\big\{(\frac{1}{2}, |+\rangle\!\langle +|), (\frac{1}{2}, |-\rangle\!\langle -|)\big\}$. The notion of steering was formalized in [26], which defines it in the context of an entanglement certification task, with Alice having access to an untrusted device and Bob to a trusted quantum system.

Let us consider an entangled state ρ_{AB}. Suppose that Alice performs a POVM $\{\Lambda_x^a\}_a$ on the state ρ_{AB}. Then, it is not necessary that the resulting assemblage is a steerable assemblage. To put it concisely, entanglement is a necessary but not sufficient criterion for the observation of steering. An example of this phenomenon is a two-qubit Werner states $\psi^p = (1-p)\pi_{AB} + p\phi^-_{AB}$, where ϕ^- is the singlet state. A Werner state is entangled if and only if $p \geq \frac{1}{3}$, and steerable if $p \geq \frac{1}{2}$. That is, for $\frac{1}{3} \leq p \leq \frac{1}{2}$, the state is entangled but not steerable. This example demonstrates that exhibiting steering is more difficult than having entanglement.

The resource theory of steering was formalized in [37], and we give a basic overview below. The resource objects are steerable assemblages, and the free objects are local-hidden-state assemblages. The set of free operations consists of one-way classical communication from Bob to Alice and local operations (1W-LOCC) that leave the set of free assemblages invariant. The set of 1W-LOCC also contains operations in which Bob is also allowed to perform a quantum instrument on his system and communicate the classical outcome prior to the measurement choice by Alice [37, Definition 1] (thus, he can influence the input to her black box).

Before defining 1W-LOCC operations, let us introduce some notation. The subnormalized state possessed by Bob is $\hat{\rho}_B^{a,x} := p_{\bar{A}|X}(a|x)\rho_B^{a,x}$. Taking $p_X(x)$ as a probability distribution over measurement choices of Alice, we can then embed the assemblage $\{\hat{\rho}_B^{a,x}\}_{a,x}$ in a classical-quantum state as follows: $\rho_{X\bar{A}B} := \sum_{a,x} p_X(x) |x\rangle\langle x|_X \otimes |a\rangle\langle a|_{\bar{A}} \otimes \hat{\rho}_B^{a,x}$,where $\{|x\rangle_X\}_x$ and $\{|a\rangle_{\bar{A}}\}_a$ are orthonormal bases.

We now define 1W-LOCC operations. Starting with a given assemblage $\{\hat{\rho}_B^{a,x}\}_{a,x}$, it is possible for Bob to perform a quantum instrument on his system, specified as the following measurement channel acting on an input state σ_B:

$$\mathcal{M}_{B\to B'Y}(\sigma_B) := \sum_y \mathcal{K}_y(\sigma_B) \otimes |y\rangle\langle y|_Y \,, \tag{2.19}$$

where

$$\mathcal{K}_y(\sigma_B) := \sum_t K_{y,t}\sigma_B K_{y,t}^{\dagger}. \tag{2.20}$$

The sum map $\sum_y \mathcal{K}_y$ is trace preserving, i.e.,

$$\sum_{y,t} K_{y,t}^{\dagger} K_{y,t} = I_B, \tag{2.21}$$

and each $K_{y,t}$ is a Kraus operator, taking a vector in $\mathcal{H}_B$ to a vector in $\mathcal{H}_{B'}$. Bob can then communicate the classical result y to Alice, who chooses the input x to her black box according to a classical channel $p_{X|Y}(x|y)$. The state after these operations is

$$\rho_{X\bar{A}B'Y} := \sum_{a,x,y} p_{X|Y}(x|y) \, |x\rangle\langle x|_X \otimes |a\rangle\langle a|_{\bar{A}} \otimes \mathcal{K}_y(\hat{\rho}_B^{a,x}) \otimes |y\rangle\langle y|_Y \,. \tag{2.22}$$

A pictorial representation of 1W-LOCC operations is given in Figure 2.7.

It is quite simple to see that if we allow Alice to communicate classically with Bob, then she can easily send (a, x) to Bob. With this information, Bob can prepare the state $\rho_B^{a,x}$. Hence, if Alice and Bob are allowed to perform an LOCC operation,

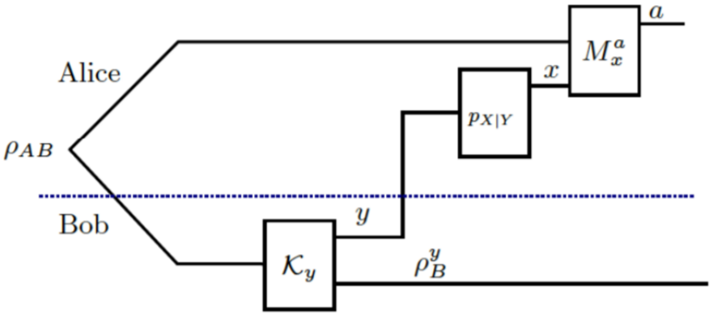

Figure 2.7. This figure represents a 1W-LOCC operation acting on an assemblage. Bob is allowed to send classical information y to Alice, who chooses the input x to her black box according to $p_{X|Y}$.

they can create a steerable assemblage. Therefore, LOCC is not a free operation in the resource theory of steering.

The 1W-LOCC that we have considered above allows for classical communication from Bob to Alice. This communication can take place prior to Alice giving the input to her device or can take place after Alice gives the input. The former scenario is a bit more complicated to handle and can be experimentally hard to implement. We thus consider a simpler, restricted class of free operations in which Bob cannot influence Alice's input to her black box.

In considering this restricted class, we are motivated by practical, relativistic constraints that can potentially limit the performance of Alice and Bob's quantum devices in any quantum steering protocol. Typically, in any such protocol, Alice, Bob, and the source of their systems are spatially separated, and furthermore, their quantum devices typically have a finite coherence time. If Alice were to wait to receive a signal from Bob before taking any action on her system, the performance of her device could potentially get much worse than it would be if she were simply instead to input to her system as soon as she receives it from the source. This perspective motivates a restricted class of 1W-LOCC operations in which any classical communication from Bob reaches Alice only after she has received the output $\bar{A}$ from her black box. We refer to these free operations as *restricted 1W-LOCC*.

Definition 31 (Restricted 1W-LOCC) *Let* $\{\hat{\rho}_B^{a,x}\}$ *be an assemblage, and let* $\left\{p_{X|X_f}, p_{\bar{A}|\bar{A}XX_fZ}, \{\mathcal{K}_z\}\right\}$ *denote a restricted 1W-LOCC operation that results in an assemblage* $\left\{\hat{\rho}_f^{a,x}\right\}$

$$\hat{\rho}_f^{a,x} = \sum_{a,x,z} p_{X|X_f}(x|x_f) p_{\bar{A}_F|\bar{A}XX_fZ}(a_f|a,x,x_f,z)\mathcal{K}_z(\rho_B^{a,x}). \tag{2.23}$$

The only difference between restricted 1W-LOCC and 1W-LOCC is that Alice's inputs are no longer dependent on classical information that Bob sends.

There is another possibility that can be considered: Alice classically communicates to Bob before giving the input to her device. This will not lead to the creation of a steerable assemblage. This kind of operation can be recast in terms of 1W-LOCC. Therefore, we need not consider it separately [37].

A number of quantifiers for quantum steering have been introduced in the literature. This includes the following: steerable weight [53], robustness [54], relative entropy of steering [37, 55] and intrinsic steerability, among others. Of relevance to this thesis are relative entropy of steering and intrinsic steerability. We will define intrinsic steerability in Chapter 5 and relative entropy of steering in Chapter 7.

2.2.3. Resource theory of Bell non-locality

Bell non-locality, formalized in [56], reflects a fundamental difference between classical and quantum correlations. Studied and introduced as a foundational concept in quantum mechanics, it has proven to be an important resource in quantum key distribution. A detailed review of Bell non-locality can be found in [30].

Entanglement is a necessary but not sufficient criteria for observation of non-locality. As an example consider the two-qubit Werner states. These states are Bell non-local for $p \geq \frac{1}{\sqrt{2}}$, steerable for $p \geq \frac{1}{2}$, and entangled for $p \geq \frac{1}{3}$. We thus observe that Bell non-locality is a stronger form of correlations than entanglement or steering.

The resource theory of Bell non-locality was formalized in [36], and we provide an overview here. The resource objects are the Bell non-local distributions $p(a, b|x, y)$.

Having fixed the resource, now let us consider the free operations. It is intuitive to see that if classical communication is allowed between Alice and Bob, then they can generate any possible probability distribution. This rules out LOCC or 1W-LOCC as possible sets of free operations. We now consider operations that do not involve communication between Alice and Bob after they have obtained a, x, y, and b. One such class of operations is local operations and shared randomness. This class includes the set of operations in which Alice and Bob are allowed to process information locally, without any communication.

Physically, local operations and shared randomness [57, 58] refers to an operation in which Alice and Bob share unlimited free randomness and can perform local operations on

- the inputs given by Alice and Bob to their respective components,
- the outputs of the two components to give the final outputs to Alice and Bob.

The local operations and shared randomness act on the initial correlation $p_i(a, b|x, y)$ corresponding to the device, in order to yield a final, modified correlation $p_f(a, b|x, y)$. These operations can be parametrized as follows [59]:

$$p_f(a_f, b_f|x_f, y_f) := \sum_{a,b,x,y} O^{(L)}(a_f, b_f|a, b, x, y, x_f, y_f)p_i(a, b|x, y)I^{(L)}(x, y|x_f, y_f). \tag{2.24}$$

Here, $I^{(L)}$ corresponds to a local correlation for a local device that takes in the inputs x_f and y_f from Alice and Bob, uses shared randomness, and performs local operations to yield new inputs x and y for the main device characterized by p_i. This can be written as

$$I^{(L)}(x, y|x_f, y_f) = \sum_{\lambda_2} p_{\Lambda_2}(\lambda_2) I_A(x|x_f, \lambda_2) I_B(y|y_f, \lambda_2), \tag{2.25}$$

where $p_{\Lambda_2}(\lambda_2)$ corresponds to the probability distribution of the shared classical variable Λ_2, $I_A(x|x_f, \lambda_2)$ corresponds to the probability of obtaining x given x_f and λ_2, and $I_B(y|y_f, \lambda_2)$ corresponds to the probability of obtaining y given y_f and λ_2.

Once the initial device p_i generates the outputs a and b, it can be post-processed by a local device that is characterized by the local correlation $O^{(L)}$. This can be written as

$$O^{(L)}(a_f, b_f|a, b, x, y, x_f, y_f) = \sum_{\lambda_1} p_{\Lambda_1}(\lambda_1) O_A(a_f|a, x, x_f, \lambda_1) O_B(b_f|b, y, y_f, \lambda_1). \tag{2.26}$$

This device takes in a, b, x, y, x_f, y_f and gives the final outputs a_f, b_f by using shared randomness and performing local operations on the inputs. Here, $p_{\Lambda_1}(\lambda)$ is a probability distribution over the classical shared random variable λ_1, $O_A(a_f|a, x, x_f, \lambda_1)$ is a conditional probability distribution for obtaining a_f given x, x_f, λ_1, a, and $O_B(b_f|b, y$ $, y_f, \lambda_1)$ is a conditional probability distribution for obtaining b_f given y, y_f, λ_1, b. See Figure 2.8 for a pictorial representation of the most general transformation of local operations and shared randomness on a correlation $p_i(a, b|x, y)$.

In the resource theory of Bell non-locality [36, 59], the resources are non-local distributions $p(a, b|x, y)$. Local operations and shared randomness are one possible set of free operations in this resource theory [36]. It can be shown from the definition

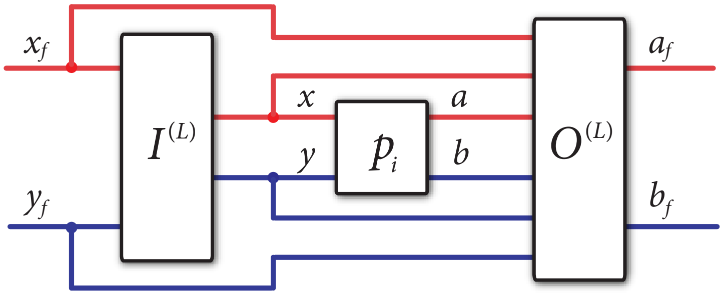

Figure 2.8. This figure depicts how local operations and shared randomness can act on an initial correlation $p_i(a, b|x, y)$ to produce a final correlation $p_f(a_f, b_f|x_f, y_f)$.

of a local distribution that the action of the local operations and shared randomness transforms a local distribution to a distribution in **L**. Furthermore, a quantum distribution remains in the set **Q** when acted upon by these free operations. To see this, replace the local boxes $O^{(L)}$ and $I^{(L)}$ in (2.24) by separable states shared between Alice and Bob with the local states encoding the probability distributions required in (2.25) and (2.26) and the measurements as projective measurements.

Another class of operations that keeps the set of local correlations invariant is called as wirings and prior-to-input classical communication. In this class, Alice and Bob are allowed to communicate before the inputs are given to their respective devices. For details, please see [59].

A number of quantifiers have been introduced for quantifying non-locality of a probability distribution. Some of these include the following: relative entropy of non-locality [60], intrinsic non-locality [13], and squashed non-locality [61]. We will give detailed definitions for these quantifiers in Chapter 6.

Chapter 3
Quantum Key Distribution

In this chapter, we review the basics of quantum key distribution. For in-depth details, please consult [62, 63, 64]. The main goal of quantum key distribution is to establish secure keys between Alice and Bob. By secure keys, we mean that Alice and Bob share a string of random variables that are not known to any Eavesdropper. We will give a mathematically precise definition later. The security of the established key relies on quantum correlations shared between two distant parties: Alice and Bob. It relies on quantum phenomena such as the monogamous nature of quantum correlations, the uncertainty principle, and the no-cloning theorem. With these properties, we bypass the need for relying on computational hardness, which is a fundamental requirement for classical cryptography. Instead, we obtain information-theoretic security, which relies on physical principles. The strength of quantum key distribution relies on the fact that the security is guaranteed irrespective of progress in computation: classical or quantum. Any public classical information collected by the eavesdropper during the protocol cannot be used to break the security of the established key with computational developments in the future. This is in contrast to computational security, where computational advances can threaten the security of keys established retroactively.

Every quantum key distribution protocol has the following three assumptions:

- Quantum mechanics or the no-signaling principle is correct.
- Alice and Bob's devices are not communicating with the eavesdropper or publicly leaking information.
- There exists a classical authenticated channel between Alice and Bob.

The existence of an authenticated channel is required because Alice and Bob need to confirm that they are communicating with each other and not with an eavesdropper impersonating as Alice or Bob. Without this assumption, quantum key distribution is susceptible to a man-in-the-middle attack.

Now, to establish a classical authenticated channel, Alice and Bob can use public key cryptography, which relies on computational security. Here we encounter a conundrum: If the security of quantum key distribution is contingent on the security of the public key authentication, which relies on computational hardness, are the keys generated from a quantum key distribution protocol secure? If the secret keys used for authentication are secure *during the run of a quantum key distribution protocol*, then the keys generated from a QKD protocol are also secure. Once the protocol for generating the key ends, the keys used for authentication can be released publicly and will not affect the security of the generated keys. That is, the key used for authentication, which relies on computational hardness, need only be secure for a short period. For more discussion on the aforementioned assumptions, please see [65].

A generic quantum key distribution protocol includes the following steps:

- Alice and Bob share an unknown bipartite state ρ_{AB}.

- Alice and Bob perform measurements on their systems to obtain measurement outcomes. These outcomes will have some correlations depending on the state shared between Alice and Bob.

- Using their data of measurement outcomes and choices, Alice and Bob quantify the amount of information that the Eavesdropper has about the outcomes. One particular way to proceed with the above analysis is to consider the set $\mathcal{S}$ of states compatible with the observed correlations. Define a set $\bar{\mathcal{S}}$ that consists of a purification ψ^{ρ}_{ABE} of ρ_{AB} jointly held by Alice, Bob, and the Eavesdropper. An optimization over this set is used to quantify the amount of information

Table 3.1. Different types of quantum key distribution based on trust assumptions on measurement devices.

Setting	**Measurement on Alice's side**	**Measurement on Bob's side**	**Resource**
Trusted QKD	Trusted	Trusted	Entanglement
1S-DI-QKD	Untrusted	Trusted	Steering
DI-QKD	Untrusted	Untrusted	Bell non-locality

that an Eavesdropper has about the measurement outcomes.

- Alice and Bob perform local operations and authenticated classical communication, which involves error correction and privacy amplification, to obtain the final key. The eavesdropper can passively copy any classical communication exchanges between Alice and Bob.

Quantifying the amount of information that Eve has about the measurement outcomes is the bottleneck for most of the security proofs. The reason is that the set of states compatible with the observed measurement outcomes can be large. To quantify Eve's information, techniques need to be developed that can optimize over such large sets of compatible states. For progress in this area, please consult [66, 39, 67].

In QKD protocols, measurement devices play a crucial part. They are instrumental in obtaining information about unknown quantum states. Therefore, one needs to understand how much we can trust the measurement devices. This trust assumption has led to different settings in quantum key distribution. In the first one, Alice and Bob trust their measurement settings. This means that they know the POVM implemented by their device. This corresponds to trusted quantum key distribution. In the second scenario, Bob knows the POVMs implemented by his device; however,

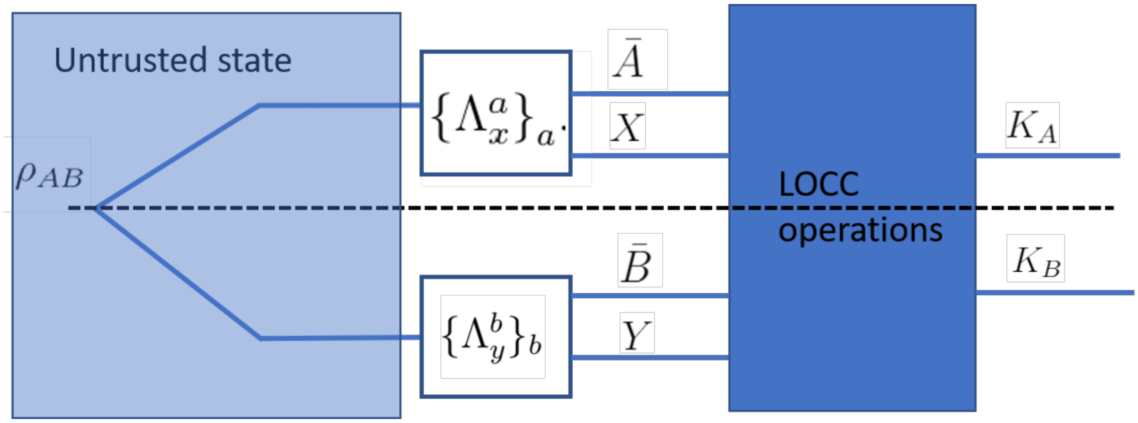

Figure 3.1. Trusted QKD

Alice does not know the POVMs implemented by her device. This is the setting of one-sided device-independent quantum key distribution (1S-DI-QKD). In the third scenario, both Alice and Bob do not trust the POVMs implemented by their device. We refer to this as device-independent quantum key distribution (DI-QKD).

3.1. Trusted quantum key distribution

Trusted quantum key distribution is the well-studied setting of quantum key distribution. Primarily, in this setting, we consider two types of protocols: prepare-and-measure (PM) and entanglement-based (EB) protocols. As the name suggests, in a prepare-and-measure protocol, Alice prepares a state and sends it to Bob, who measures the received state. In an entanglement-based protocol, an untrusted source prepares a bipartite state and sends one share each to Alice and Bob. Alice and Bob measure their shares of the state to obtain classical variables, from which they extract a key. Although PM protocols and EB protocols are seemingly different on the surface, one can show that for every PM protocol there exists an EB protocol and vice versa [68]. This equivalence is extremely useful because PM protocols are easier to implement and EB protocols are slightly easier to analyze.

Most prepare-and-measure protocols consist of the following steps:

- The following transmission round is performed n times:

  - Alice needs to encode the bit values 0 and 1. She encodes the bits in the state $\psi_i^{0,1}$, where $i \in \{1, \ldots, m\}$ reflects the encoding basis, and ψ_i^0 and ψ_i^1 are the states used to encode 0 and 1, respectively. For the s^{th} transmission round, Alice stores the values of the encoding basis in random variable X_1^s and the encoded bits in X_2^s.
  - Alice sends the states through an insecure quantum channel, which can be controlled by an eavesdropper. This implies that Alice and Bob do not know the state that is received by Bob.
  - Bob then measures the received state in some basis ψ_j and stores the values of the measuring basis for the s^{th} round in Y_1^s and the bit values in Y_2^s.

- At the end of n transmission rounds, Alice has $X_1 = X_1^1 X_1^2 \ldots X_1^n$. Similarly, we can define X_2, Y_1, and Y_2.

- Alice and Bob then perform a sifting process. Alice announces the basis X_1 in which she prepared the state and Bob announces the basis Y_1 in which he measured the state. They only keep X_2^k and Y_2^k for which the measurement basis is the same, that is, $X_1^k = Y_1^k$. This gives the raw key, $\tilde{X}_2$ and $\tilde{Y}_2$, on which they perform error correction and privacy amplification. These protocols involve local operations and public communication. The classical communication takes place over a classical authenticated channel.

The security proof for most PM protocols relies on using correlations observed in X_2 and Y_2, along with the knowledge of X_1 and Y_1, to determine the channel implemented by the eavesdropper. This step can be hard because the number of channels compatible with Alice and Bob's observations can be large.

As discussed above, a slightly easier approach to security proofs is to define a corresponding entanglement-based protocol. In entanglement-based protocols, Alice prepares a bipartite state $\rho_{AA'}$ and sends the system A' over an insecure quantum channel $\mathcal{N}_{A'\to B}$ that can be controlled by an eavesdropper. As a consequence, Alice and Bob share the state $\rho_{AB} = \mathcal{N}_{A'\to B}(\rho_{AA'})$. Alice and Bob measure this state in a random basis to obtain the measurement outcomes. To establish the security of this protocol, Alice and Bob determine the states ρ_{AB} that are compatible with their measurement outcomes. Because we assume that all information leaked from the channel is collected by Eve, the state held by Eve is a purification ψ^{ρ}_{ABE} of a compatible state ρ_{AB}. Even the set of joint states held by Alice, Bob, and the eavesdropper can be extremely large and therefore optimizing over this set is hard. Several techniques have been developed in [69, 39, 4] to solve this problem for various protocols.

The first quantum key distribution protocol was introduced in the seminal work of [3] with a security proof provided in [4]. Following that, there have been several other prepare-and-measure protocols introduced, such as the six-state protocol [70] and the B92 protocol [71]. Ekert [23] also introduced a protocol that uses bipartite entanglement to obtain secure keys and relies on the violation of a Bell inequality to detect the presence of an Eavesdropper.

3.1.1. Resource in QKD

To obtain a non-zero secret key rate in a QKD protocol, *the underlying channel implemented by an eavesdropper should not be entanglement breaking.* An entanglement-breaking channel is defined as follows [72]:

Definition 32 (Entanglement-breaking channels) *A channel $\mathcal{N}$ is an entanglement breaking channel if its Choi matrix $\Phi^{\mathcal{N}}_{AB}$ is separable.*

One can easily understand this requirement in the context of EB protocols. To extract a secret key in these protocols, the shared state must be an entangled state.

If the shared state is separable, then Alice and Bob cannot generate a secure key from this state via a quantum key distribution protocol. As seen here, entanglement in the bipartite state ρ_{AB} is a necessary condition, but it is not known to be a sufficient condition for obtaining a secure key. The sufficient condition has been extensively studied in [73, 74, 75].

In quantum key distribution, we assume that an eavesdropper controls the channel connecting Alice and Bob. This means that *a priori* Alice and Bob do not know the channel. They might assume a certain model for the channel that is later verified during the protocol using their measurement statistics. Also, any noise, that in part, can be due to physical constraints such as fiber losses, is attributed to the eavesdropper. If the noise added by the physical process is greater than a certain threshold, then the bipartite state is no longer entangled. Hence, no secure key can be extracted from it.

As an example, suppose that an optical fiber with transmissivity η connects Alice and Bob. We attribute the fiber losses to an eavesdropper. We assume that the eavesdropper is collecting all information being leaked from the fiber. This fundamentally limits the distance at which a secret key can be distilled. This intuition is formalized for various channels in [6, 7, 8]. The attribution of all losses in the channel to an eavesdropper is a powerful assumption. However, to obtain information-theoretic security, this assumption is necessary.

Now, if we inspect the workings of the security proof of BB84, B91, and the six-state protocol, we see that there is an underlying assumption that Alice's encoding and Bob's measurement devices are fully characterized. If we consider the entanglement-based counterpart, Alice and Bob's measurements are fully characterized. One can also consider an extreme scenario in which Alice and Bob do not characterize their measurement devices. We do not know the POVMs corresponding

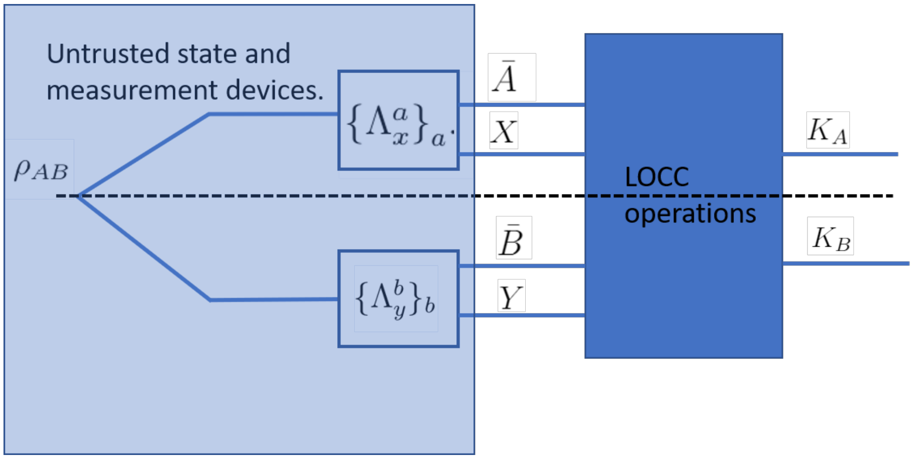

Figure 3.2. Device-independent quantum key distribution

to the measurement devices. Can we extract secret keys from an unknown bipartite state without trusting the measurement performed on these states? Fortunately, we can still exploit the correlations present in quantum mechanics without trusting measurement devices and quantum states. In this context, we study device-independent quantum key distribution, which we explain in the next section.

3.2. Device-independent quantum key distribution

3.2.1. Model

Consider a scenario in which Alice and Bob receive a device manufactured by some malicious untrusted party. This box can contain an entangled state, a separable state, or be described by a no-signaling correlation; however, Alice and Bob do not know the contents of the box. They can access this device with classical inputs $x, y \ \in \{0, 1, \ldots, m\}$. Once Alice and Bob choose the classical inputs, the device performs some arbitrary action and gives Alice and Bob some classical output $a, b \in \{0, 1, \ldots, m\}$. So the only way that Alice and Bob can interact with their devices is through classical inputs and classical outputs. We can thus characterize these devices with a conditional probability distribution $P_{\bar{A}\bar{B}|XY}$, where $\bar{A}, \bar{B}$ are random variables associated with Alice and Bob's respective outcomes, and X, Y are random variables associated with Alice and Bob's respective inputs. This device is called a black box because we do not know the inner workings of the box. The question now becomes the following: can Alice and Bob still do something useful with the generated correlations without knowing about the inner working of the device?

Interestingly, with the following assumptions and minimal trust on their devices, Alice and Bob can still construct protocols to extract secret keys from the conditional probability distribution $P_{AB|XY}$:

- There is no extraneous/unwanted communication from Alice and Bob's device.
- There is a trusted random number generator to produce classical variables.
- They have trusted classical devices (e.g., memories and computing devices) to store and process the classical data generated by their quantum devices.
- They are connected by authenticated classical public channels.

The first assumption is required to make sure that the device is not leaking out classical data. Without this, the device can communicate the classical data to Eve, and there can be no secure key. The second assumption is required because Alice and Bob's measurement choices should not be known in advance to the eavesdropper or the device manufacturer. If the choices are known in advance, then the manufacturer can rig the devices. The fourth assumption is required because Alice and Bob need to be sure that they are communicating with each other.

A basic device-independent protocol consists of the following steps:

- The following round is performed n times.
  - Alice and Bob give inputs X_i and Y_i to the device.
  - The device gives outputs $\bar{A}_i$ and $\bar{B}_i$.

At the end of these rounds, Alice has $X = X_1X_2 \dots X_n$ and $\bar{A} = \bar{A}_1\bar{A}_2 \dots \bar{A}_n$. Bob has $Y = Y_1Y_2 \dots Y_n$ and $\bar{B} = \bar{B}_1\bar{B}_2 \dots \bar{B}_n$.

- Alice and Bob announce their measurement settings. For m random rounds, Alice and Bob share their outputs $\bar{A}_j$ and $\bar{B}_j$, $j \in \{1, \dots, m\}$. They use this information to bound the value by which the outputs $\bar{A}_j$, X_j, Y_j, and $\bar{B}_j$ violate a Bell inequality. They discard the data collected in these m rounds. The rounds in which Alice and Bob's measurement settings do not match are also subsequently discarded. The remaining outcomes form a raw key.
- If the Bell inequality violation is sufficiently strong, they proceed with local operations and public communication, which includes error correction and privacy amplification.

3.2.2. Resource in DI-QKD

Suppose that that an untrusted device is defined by a local correlation $P_{\bar{A}\bar{B}|XY} = \sum_{\lambda} p_{\Lambda}(\lambda) p_{\bar{A}|X,\Lambda} p_{\bar{B}|Y,\Lambda}$. The eavesdropper can copy the classical information λ, and when Alice and Bob reveal their measurement choices, induce a classical channel from X, Λ, Y to obtain the outcomes $\bar{A}$ and $\bar{B}$. Thus, we see that no secret key can be extracted from a local correlation in a device-independent QKD protocol.

To check for non-local distributions, we check for violations of Bell inequalities. The first Bell inequality was formulated by John Bell in his seminal paper "On the Einstein Podolsky Rosen paradox" [56]. This inequality was generalized in [76] to make it realizable with experiments. Ever since its formulation, several other Bell inequalities have been explored. These include Mermin inequalities [77] and tilted CHSH inequalities [78], among others.

A crucial part of the aforementioned protocol is a Bell inequality test, which is defined as follows:

Definition 33 (Bell inequality) *Let $p(a,b|x,y) \in \mathbf{L}$, where $\mathbf{L}$ is described in* (2.8). *Then there exists an inequality*

$$\sum_{a,b,x,y} s(a,b,x,y)p(a,b|x,y) \leq S, \tag{3.1}$$

satisfied for all $p \in \mathbf{L}$ and not satisfied for some $p \in \mathbf{Q}$, where $\mathbf{Q}$ is described in (2.9). *Here $S > 0$, and $s(a,b,x,y) \in \mathbb{R}$.*

The objective of a Bell inequality test is to quantify Eve's information in a device-independent way. This proves to be a challenging task, especially because here, we need to optimize over *a set of states as well as measurements* that are compatible with the observed measurement statistics. Most DI-QKD protocols rely on the CHSH Bell inequality. This is because the two-input two-output setting of the CHSH inequality allows us to invoke Jordan's lemma, which enables us to assume a dimension bound on Alice and Bob's underlying state. Establishing security proofs for DI-QKD protocols with other Bell inequalities is still an open line of research.

In a device-independent scenario, one can consider different models of an eavesdropper. Besides the assumptions that we made above, we can also assume that Alice, Bob, and Eve's systems are governed by quantum mechanics. This implies that the conditional probability distribution describing the black box has an underlying quantum strategy. The two-component box is assumed to have an underlying quantum state ρ_{AB}, and Eve has the purification of this state ψ^{ρ}_{ABE}. Such an eavesdropper is known as a *quantum eavesdropper* and has a quantum extension of the distribution, defined as follows:

Definition 34 (Quantum extension) *Let $p(a,b|x,y) \in \mathbf{Q}$. A distribution in the set $\mathbf{Q}$ arises from an underlying state ρ_{AB} and POVMs characterized by $\{\Lambda^a_x\}_a$ and*

$\left\{\Lambda_y^b\right\}_b$. *Now, consider a quantum state* ρ_{ABE} *such that* $\mathrm{Tr}_E\left(\rho_{ABE}\right) = \rho_{AB}$. *Then, a quantum extension of* $p(ab|xy)$ *is defined as*

$$p(a,b|x,y)\rho_E^{a,b,x,y} = \mathrm{Tr}_{AB}\left[(\Lambda_x^a \otimes \Lambda_y^b \otimes I_E)\rho_{ABE}\right]. \tag{3.2}$$

Lower bounds on secret key rates for protocols with a quantum eavesdropper have been considered in [39, 66].

If the distribution describing their device is no-signaling, then we can also assume that the eavesdropper is no-signaling. We can further model a no-signaling eavesdropper in two ways. In the first one, we can model the eavesdropper as a *no-signaling quantum eavesdropper*. A no-signaling quantum eavesdropper has a no-signaling quantum extension of the probability distribution $p(a,b|x,y)$, defined as follows:

Definition 35 (No-signaling quantum extension) *The assemblage* $\hat{\rho}_E^{a,b,x,y} = p(a,b|x,y)$ $\rho_E^{a,b,x,y}$ *is a no-signaling quantum extension of* $p(a,b|x,y)$ *if the following no-signaling conditions are satisfied:*

$$\sum_a p(a,b|x,y)\rho_E^{a,b,x,y} = \sum_a p(a,b|x',y)\rho_E^{a,b,x',y} \quad \forall x,x' \in \mathcal{X}. \tag{3.3}$$

$$\sum_b p(a,b|x,y)\rho_E^{a,b,x,y} = \sum_b p(a,b|x,y')\rho_E^{a,b,x,y'} \quad \forall y,y' \in \mathcal{Y}. \tag{3.4}$$

In the second model, one assumes that the eavesdropper has a *no-signaling extension*, defined as follows:

Definition 36 (No-signaling extension) *The conditional probability distribution* $p(a,b,c|x,y,z)$ *is a no-signaling extension of* $p(a,b|x,y)$ *if the following no-signaling* *condition holds*

$$\sum_{c} p(a,b,c|x,y,z) = \sum_{c'} p(a,b,c|x,y,z) \quad \forall z, z'. \tag{3.5}$$

The security of various DI-QKD protocols has been proven for different models of the eavesdropper [79, 80, 81].

Although DI-QKD is the holy grail of quantum key distribution, its practical implementation is extremely challenging. To implement DI-QKD, it is imperative to perform a loophole-free Bell test, which was only recently demonstrated in [82, 83, 84]. Recall that in trusted quantum key distribution, we attributed only the noise in the transmission channel to the eavesdropper. With DI-QKD protocols, we attribute even the noise in the measurement device to the eavesdropper. This makes DI-QKD protocols extremely sensitive to noise.

Above we have mentioned two scenarios: one in which Alice and Bob trust their measurement settings, and in the other, they do not trust their measurement settings. One can also consider several in-between scenarios. There are three different models of QKD that have been considered: semi-device-independent quantum key distribution [85], one-sided device-independent quantum key distribution [67, 38], and measurement device-independent quantum key distribution [86, 87]. With semi-device-independent QKD, the dimensions of the quantum systems and measurements are trusted and known. In one-sided device-independent QKD, we do not trust the preparation phase and measurement devices with either Alice or Bob. In measurement-device-independent QKD, we trust the preparation phase but do not trust the measurement devices.

In the next section, we investigate one-sided device-independent quantum key distribution.

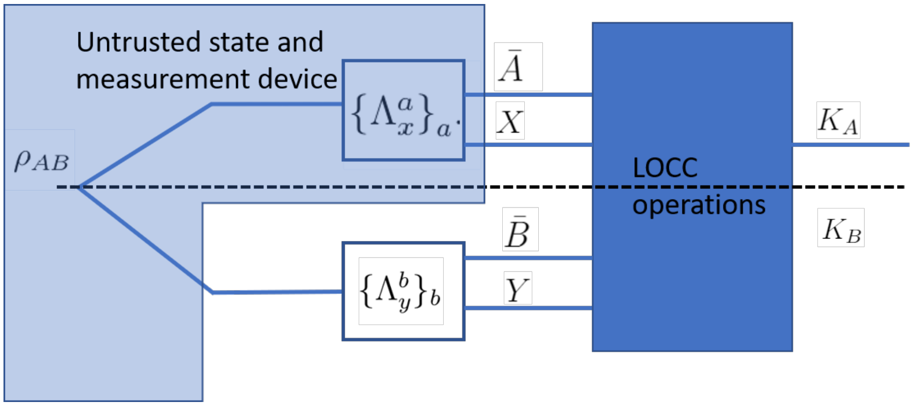

Figure 3.3. One-sided device-independent quantum key distribution

3.3. One-sided device-independent QKD

3.3.1. Model in 1S-DI-QKD

One attractive alternative to DI-QKD is 1S-DI-QKD introduced in [67, 38]. Consider a scenario in which, say, a bank and a user want to share a secret key. The bank, since they have a lot of money, invest in equipment that is resistant to noise, and has highly calibrated preparation and measurement devices. However, the user cannot afford to keep the highly calibrated instruments due to high maintenance costs and, therefore, cannot trust the equipment. The user's equipment effectively behaves like a black box and can only be accessed via a classical input and a classical output. Quantum key distribution protocols in which one of the parties does not trust its measurement equipment are technically known as one-sided device-independent protocols. The aforementioned scenario is reminiscent of the one we discussed in quantum steering, in which one of the parties is characterized by a conditional probability distribution, and the other party is characterized by a correlated quantum state.

The only difference between DI-QKD and one-sided device-independent QKD comes from the treatment of the variable Y. In DI-QKD, we assume that we do not have the characterization of Y; however, in 1S-DI-QKD, we know the measurement/action being carried out by the device when the user inputs Y.

The first 1S-DI-QKD was introduced in [67], where they proved that the prepare-and-measure BB84 protocol could be made one-sided device-independent on Bobs side, albeit with lower key rates. The proof requires a memoryless assumption, as discussed in [88]. In [38], it was shown that to obtain secure key rates in this setting, Alice and Bob's assemblage needs to violate a steering inequality, establishing a connection with quantum steering. It is, in fact, simple to prove that if an assemblage has a local-hidden-state model, then a 1S-DI-QKD protocol cannot extract a secret key from the assemblage.

3.3.2. Resource required

The resource required in 1S-DI-QKD is steerable assemblages, which can be witnessed by a steering inequality. A steering inequality is the counterpart of a Bell inequality in quantum steering. It is a witness for steerable assemblages and has the following form:

Definition 37 (Steering inequality) *A steering inequality is given as follows:*

$$\beta = \mathrm{Tr}\left[\sum_{b,y} \Lambda_y^b \hat{\rho}^{b,y}\right] \leq \beta_{\mathrm{LHS}}, \tag{3.6}$$

where $\left\{\Lambda_y^b\right\}_{b,y}$ are the POVM elements and $\hat{\rho}^{b,y}$ is the assemblage, β_{LHS} is the maximal value of β that can be obtained by any local-hidden-state assemblage.

3.4. Attacks by eavesdropper

Often the first attempt to obtain secure key rates for a quantum key distribution protocol is in the asymptotic regime, in which the number n of transmission rounds tends to infinity. Let l denote the length of the secure key that can be extracted in n rounds. Then, the rate of the protocol is given by $\lim_{n\to\infty}\frac{l}{n}$. Of course, this is not a realistic assumption; however, it can be insightful for determining the key rates that one can expect from a particular protocol.

Security is also first generally proved under an independent and identically distribution (i.i.d.) assumption or collective attack assumption. Under this assumption, Eve's actions, which appear as a noisy quantum channel $\mathcal{N}$ to Alice and Bob in PM protocols, remain the same for each transmission round. She is allowed to make arbitrary collective measurements on her collected quantum systems at the end of the protocol. In the entanglement-based scenario, we assume that the quantum state, probability distribution, or the assemblage considered remains the same for each transmission round.

Collective attacks are not the most general attacks that can be carried out by an eavesdropper. For general attacks, the channel implemented by an eavesdropper in a particular transmission round can depend on the information collected in the previous rounds. In this context, techniques such as de Finetti reductions [89] and the post-selection technique [90] can be employed for trusted quantum key distribution. To obtain security under these general attacks, we only need to prove security for collective attacks and then invoke the above techniques. The techniques, with a reduction in the key rates, give a secure key for arbitrary attacks. In DI-QKD, de Finetti reductions [91] and entropy accumulation [66] can be employed to deal with general attacks. However, the use of these techniques is not as straightforward as in the trusted setting. For general attacks, the secure key rate tends to rates obtained in collective attacks, for a large number of transmission rounds.

3.5. Lower bound on asymptotic key rates and collective attacks

In this section, we discuss techniques for obtaining lower bounds on key rates for a QKD protocol. For this, we first discuss lower bounds on the rates of secret key distillation protocols.

Let Alice and Bob share n copies of an arbitrary state ρ_{AB}. The collective joint state on Alice, Bob, and Eve's systems is given by the purification ρ_{ABE} of the shared state ρ_{AB}. Alice and Bob can perform local operations and public communication protocols on $\rho_{ABE}^{\otimes n}$ to obtain an ε-secure key $\omega_{K_A K_B E}$ where K_A and K_B are classical registers, such that

$$\frac{1}{2}\|\bar{\Phi}_{K_A K_B} \otimes \sigma_E - \omega_{K_A K_B E}\|_1 \leq \varepsilon, \tag{3.7}$$

where $\bar{\Phi}_{K_A K_B} = \frac{1}{|d|}\sum_{i=1}^{d} |ii\rangle\langle ii|_{K_A K_B}$ is the ideal key and $\omega_{K_A K_B E}$ is the joint state at the end of the protocol. Such a protocol is called an (n, R, ε) protocol, where $R = \frac{\log_2 d}{n}$. A rate R is achievable if for all $\varepsilon \in (0, 1], \delta > 0$, and sufficiently large n, there exists an $(n, R - \delta, \varepsilon)$ protocol. Then the maximum achievable rate is the distillable key of the state $\rho_{ABE}^{\otimes n}$.

Since distillable key involves an optimization over all possible LOCC protocols, it can be hard to calculate in general. Therefore, we obtain lower bounds and upper bounds on the rate and, in this way, narrow down the region of possible values of the distillable key. One such well known lower bound is the Devatak-Winter bound introduced in [2]. This is a lower bound on the one-way distillable key of the state and hence is also a lower bound on the distillable key. Here, one-way distillable key means that classical communication is allowed only in one direction, either from Alice to Bob or Bob to Alice. The Devatak-Winter formula states that the key rate R is bounded from below as follows:

$$R \geq I(X;Y)_\rho - \chi(Y;E)_\rho, \tag{3.8}$$

where $I(X;Y)$ is the mutual information between X and Y, and $\chi(Y;E)$ is the Holevo information between Y and the eavesdropper's system. The state $\rho_{XYE} =$

$\sum_{x,y} |x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y \otimes \mathrm{Tr}_{AB}\left[\left(\Lambda_a^x \otimes \Lambda_b^y\right)\rho_{ABE}\right]$, where $\{\Lambda_x^a\}_a$ and $\left\{\Lambda_y^b\right\}_b$ are POVMs characterizing Alice and Bob's measurement operators. For protocols with communication from Alice to Bob, we can replace Y in the Holevo information term with X.

The distillable key of a state is closely related to key rates in quantum key distribution. The missing link is that in a quantum key distribution protocol, Alice and Bob cannot make any assumption on the form of the state $\omega_{K_A K_B E}$ because they do not know the underlying bipartite state ρ_{AB}. However, Alice and Bob do have certain observations from their measurement outcomes, which they can use to characterize the set of states compatible with the outcomes. To take this into account, we need to perform an optimization over the set of states in the Holevo-information term in the Devatak-Winter formula as follows:

$$R \geq I(X;Y) - \max_{\rho_{YE}} \chi(Y;E). \tag{3.9}$$

This optimization makes it challenging to calculate lower bounds. In several protocols, we cannot perform this optimization and have to come up with novel ways of obtaining upper bounds on the Holevo information.

3.5.1. Honest devices

A QKD protocol is designed such that, for any noisy channel, it yields a secret key with high probability or it aborts. The key rates that one would obtain from these protocols are calculated over expected noise models. It is important to point out here that the protocol will be secure for *any* noise model. But to calculate key rates, we assume a specific noise model. Ideally, we would want a protocol to give a high key rate for the expected noise models. Some common forms of noise models are the depolarizing channel, thermal channel, pure-loss channel, erasure channel, among others.

In the DI-QKD literature, one often encounters the term *honest device.* An honest device means that the noise in the device behaves as expected. A DI-QKD protocol will be secure for all noise models, but we calculate the key rates for the honest devices.

Chapter 4
Intrinsic Information and Squashed Entanglement

In this chapter, we introduce information-theoretic quantities based on conditional mutual information, which are relevant in the context of secret key distillation. We first formalize the definition of distillable key for both classical and quantum correlations. Then we motivate the use of conditional mutual information-based measures for quantifying secret correlations in a joint probability distribution or a bipartite quantum state. We revisit the definitions of intrinsic information, introduced by Ueli Maurer in [9] and its improvements as introduced in [92] and [93]. We recall the definition of squashed entanglement [10], an entanglement measure inspired by intrinsic information. With these definitions established, we explore the connection of squashed entanglement with key rates in quantum key distribution.

4.1. Secret key distillation protocols

In secret key distillation protocols, the goal is to establish a secret key bit string shared between Alice and Bob. An eavesdropper should have little or no knowledge about this bit string. The security of secret bit strings can be established under various assumptions. One common assumption is that of computational complexity, which is currently employed in most cryptographic protocols. Under the computational complexity setting, security is based on the hardness of certain mathematical problems. The security proofs in quantum key distribution do not rely on *computational assumptions* but rather on certain *physical assumptions*. Keys established with security based on physical assumptions have information-theoretic security.

In secret key distillation protocols, we often use the term LOPC, which stands for local operations and public communication. This kind of operation is a concatenation of the following one-way LOPC operations, described as follows:

- Alice performs the following quantum instrument: $\left(\Lambda_{A\to A\bar{A}}\otimes \mathrm{id}_{BE}\right)\left(\rho_{ABE}\right) =$

$\sum_a \left(\mathcal{E}_a \otimes \mathrm{id}_{BE}\right)(\rho_{ABE}) \otimes |a\rangle\langle a|_{\bar{A}}$, where the sum map $\sum_a \mathcal{E}_a$ is trace preserving.

- Alice sends a over an authenticated public channel. Eve can copy all the classical information exchanged between Alice and Bob over the authenticated channel. Then the state at the end of protocol is $\rho_{A\bar{A}B\bar{B}E\bar{E}} = \sum_a(\mathcal{E}_a \otimes \mathrm{id}_{BE})(\rho_{ABE}) \otimes |aaa\rangle\langle aaa|_{\bar{A}\bar{B}\bar{E}}$.

We can replace Alice with Bob in the above one-way LOPC operations.

We can now define an LOPC operation Λ as a concatenation of n rounds of one-way LOPC operations $\{\Lambda\}_{n\in\mathbb{N}}$. An example of an LOPC operation is as follows: $\Lambda : \rho_{ABE}^{\otimes n} \to \rho_{K_AK_BEA^nB^nE^n}$ where $\rho_{K_AK_BEA^nB^nE^n}$ is final state of this LOPC protocol.

An ideal secret key can be defined as follows

$$\bar{\Phi}_{K_AK_BE} = \bar{\Phi}_{K_AK_B} \otimes \rho_E, \tag{4.1}$$

with ρ_E an arbitrary state of the eavesdropper's system E and $\bar{\Phi}_{K_AK_B} = \sum_i \frac{1}{d} |ii\rangle\langle ii|_{K_AK_B}$.

4.1.1. Classical secret key distillation protocol

Consider three parties Alice, Bob, and an eavesdropper, with access to i.i.d. random variables. Let $\bar{A}$ be a random variable with Alice, $\bar{B}$ a random variable with Bob, and Z a random variable with Eve. All three parties know the joint probability distribution $P_{\bar{A}\bar{B}Z}$ associated with these three random variables (this is a physical assumption in this setting). Suppose that Alice and Bob want to extract a secret key, unknown to Eve, from $P_{\bar{A}\bar{B}Z}$ by using an LOPC operation. The secret key rate, for this setting, is the maximum rate at which Alice and Bob can extract a secret key from n independent instances of the probability distribution. This rate is denoted as $K_D(\bar{A};\bar{B}|Z)$. For a precise definition, please see [9].

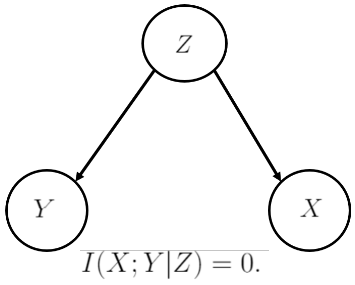

Figure 4.1. Short Markov chain

4.1.2. Quantum secret key distillation protocol

A similar definition of secret key distillation protocols can be formulated for quantum states. Instead of access to i.i.d. random variables, Alice, Bob, and Eve have access to n copies of a tripartite quantum state ρ_{ABE}. This state is also known to the three parties (physical assumption in this setting). Alice and Bob can perform an LOPC operation $\Lambda_{A^nB^n\to K_AK_B}$ to extract a state $\rho_{K_AK_BE} = \Lambda_{A^nB^nE^n\to K_AK_BE}(\rho_{A^nB^nE^n})$, such that

$$\frac{1}{2}\|\rho_{K_AK_BE} - \bar{\Phi}_{K_AK_B} \otimes \rho_E\|_1 \leq \varepsilon, \tag{4.2}$$

where $\varepsilon > 0$, and $\bar{\Phi}_{K_AK_B} = \frac{1}{d}\sum_{k=1}^{d} |kk\rangle\langle kk|_{K_AK_B}$. The aforementioned protocol is an (n, R, ε) secret key distillation protocol with the rate $R = \frac{\log_2 d}{n}$. A rate R of secret key distillation is achievable for ρ_{ABE} if there exists an $(n, R-\delta, \varepsilon)$ secret key distillation protocol for all $\varepsilon \in (0, 1)$, $\delta > 0$, and sufficiently large n. The distillable key $K_D(A; B|E)_\rho$ of a state ρ_{ABE} is equal to the supremum of all achievable rates.

4.2. Entanglement measures, secrecy measures, and conditional mutual information

Suppose that we are given a joint probability distribution $P_{\bar{A}\bar{B}Z}$ or a tripartite state ρ_{ABE}. Can we find the distillable key of these objects? Finding this quantity is not easy, and so we resort to finding tight upper bounds and lower bounds. Mutual information-based quantities are useful in the context of obtaining upper bounds on $K_D(\bar{A}; \bar{B}|Z)_p$ and $K_D(A; B|E)_\rho$.

The intuition behind using conditional mutual information-based quantities for quantification of secrecy can be traced back to the following property of conditional mutual information:

> *The conditional mutual information $I(X;Y|Z)$ of a short Markov chain, $X \to Z \to Y$, is equal to zero.*

The correlations existing between X and Y can be quantified by $I(X;Y)$. However, X and Y might have some common correlations with random variable Z. Therefore, one needs to 'squash' these common correlations to measure the intrinsic correlations, shared just by X and Y (which can form the basis of a secret key).

Apart from the quantification of secret key rate in triples, conditional mutual information is also useful in quantifying the entanglement of a state. For this, first recall that a separable state can be written as $\rho_{AB} = \sum_\lambda p_\Lambda(\lambda)\rho_A^\lambda \otimes \rho_B^\lambda$. Now for this state, the conditional mutual information is equal to zero, i.e., $I(A;B|\Lambda)_\rho = 0$. Now, consider any bipartite state σ_{AB}, and for this state, let us define the following function: $\inf_{\sigma_{AB\Lambda}} I(A;B|\Lambda)_{\sigma_{AB\Lambda}}$, where $\sigma_{AB\Lambda} = \sum_\lambda p_\Lambda(\lambda) |\lambda\rangle\langle\lambda|_\Lambda \otimes \sigma_{AB}^\lambda$, such that $\mathrm{Tr}_\Lambda [\sigma_{AB\Lambda}] = \sigma_{AB}$. For every separable state, we can find a decomposition such that $\inf_\Lambda I(A;B|\Lambda)_{\rho_{AB\Lambda}} = 0$. This gives us an intuition that the CMI of quantum states can be helpful in quantifying entanglement of a quantum state [94, 95].

4.2.1. Intrinsic information

In this section, we define intrinsic information, which was first introduced in [9].

In the last section, we discussed that given a triple X, Y, and Z, the CMI $I(X;Y|Z)$ can be used as a measure for secret correlations. However, recall that the conditional mutual information is not monotone with respect to local operations on the conditioning variable. This needs to be taken into account when we define a measure for secret correlations. That is, we should allow for all classical channels

$P_{Z'|Z}$ to be performed on Z. With this, we define *intrinsic information* as follows [9]:

Definition 38 (Intrinsic information) *Given three discrete random variables X, Y, and Z, the intrinsic information between X and Y given Z is defined as*

$$I(X;Y\downarrow Z)=\inf_{P_{\bar{Z}|Z}} I(X;Y|\bar{Z}), \tag{4.3}$$

where the infimum is over all possible classical channels from Z to $\bar{Z}$.

For the secret key distillation protocol, the eavesdropper is allowed to perform any operation on his random variable Z. Now, we know that the conditional mutual information is not monotone with respect to local operations on the conditioning system, as discussed in Section 1.4.2.. To take this into account, we define the intrinsic information with an infimum over all the local operations, $P_{\bar{Z}|Z}$, that the eavesdropper can perform. The range of $\bar{Z}$, in principle, can be extremely large. It was, however, shown in [96] that it suffices to take the range of $\bar{Z}$ to be equal to that of Z. Therefore, we can replace the infimum in the definition with a minimization. In [9], it was proved that $K_D(X;Y|Z)_P \leq I(X;Y\downarrow Z)_P$. That is, the distillable key of the probability distribution P_{XYZ} is bounded from above by the intrinsic information of the distribution.

4.2.2. Reduced intrinsic information

How tight is intrinsic information as an upper bound on the secret key rate? Are there any probability distributions with non-zero intrinsic information but with no distillable key? These questions were explored in [92]. In this work, Renner and Wolf considered a particular probability distribution and proved for the first time a gap between its intrinsic information and asymptotic secret key distillation. In this work, they also introduced a quantity known as *reduced intrinsic information*, defined as follows:

Definition 39 (Reduced intrinsic information) *Given are three discrete random variables X, Y, and Z. The intrinsic information between X and Y given Z is defined as follows:*

$$I(X;Y \downarrow\downarrow Z) = \inf_{P_{U|XYZ}} (I(X;Y\|UZ) + H(U)), \tag{4.4}$$

where the infimum is over all possible classical channels from ZXY to U.

Two important properties of reduced intrinsic information are as follows:

- $I(X;Y \downarrow\downarrow Z) \leq I(X;Y \downarrow Z)$
- $K_D(X;Y|Z) \leq I(X;Y \downarrow\downarrow Z)$.

Subsequently, in [93], an improved upper bound on secret key distillation capacity was introduced. This new, improved upper bound is defined as follows:

Definition 40 *Given are three discrete random variables X, Y, and Z. Then the improved reduced intrinsic information between X and Y given Z is defined as*

$$I(X;Y\|Z) = \inf_J I(X;Y \downarrow J) + I(XY;J|Z), \tag{4.5}$$

where the infimum is over any arbitrary correlated random variable J.

This improved bound is less than reduced intrinsic information and is also an upper bound on the distillable key.

4.2.3. Squashed entanglement

As explained above, conditional mutual information is also useful in quantifying entanglement in a quantum state. We already saw some evidence of this in Section 4.2.

in the form of CMI being equal to zero for separable states. This intuition was formalized in [10] as follows:

Definition 41 (Squashed entanglement) *The squashed entanglement of the state ρ_{AB} is defined as*

$$E_{sq}(A:B)_\rho = \inf_{\rho_{ABE}} \frac{1}{2} I(A;B|E), \tag{4.6}$$

where the infimum is over all extensions ρ_{ABE} of ρ_{AB} such that $\rho_{AB} = \mathrm{Tr}_E\left[\rho_{ABE}\right]$.

The connection between entanglement measures and conditional mutual information has also been explored in [94, 95].

Squashed entanglement can be thought of as quantifying quantum correlations in ρ_{AB} that are inaccessible to any other quantum system (hence the conditioning on the extension system). Once we know the bipartite state ρ_{AB}, one can construct its purification $\psi^{\rho}_{ABE'}$. Now, any extension ρ_{ABE} of ρ_{AB} can be reached by acting with a local operation $\Lambda_{E'\to E}$ on the purifying system. Therefore, we have the following equivalence [10]:

$$\inf_{\rho_{ABE}} I(A;B|E)_\rho = \inf_{\Lambda_{E'\to E}} I(A;B|E)_{\Lambda(\psi^\rho)} \tag{4.7}$$

This has a remarkable resemblance to the definition of intrinsic information. However, unlike intrinsic information for which we can replace the infimum with a minimum, such a result is not known for squashed entanglement. This suggests that the quantity is inherently uncomputable. As a consequence, it is not currently known whether squashed entanglement is computable. By uncomputable, we mean that we do not even know if this quantity can be computed at all. This is because there is no dimension bound on the extension system. Therefore, any algorithm that tries to calculate this quantity would not be able to determine at what point to stop. However, this does not imply that the squashed entanglement is useless. We can always make a particular choice of the extension system and thus obtain an upper bound on squashed entanglement.

Squashed entanglement has several nice properties, including faithfulness [97, 98], continuity [97], additivity on product states and super-additivity in general [10], and monotonicity under LOCC [10]. These properties, as discussed in Section 2.2.1., ensure that squashed entanglement is a monotone in the resource theory of entanglement and is a quantifier of entanglement in a bipartite state.

Now, we pose a different question: Is this measure useful in the context of any protocols involving entanglement? This question was positively answered in [11], which proved that squashed entanglement is an upper bound on distillable entanglement of the state [10], and is an upper bound on the distillable key [11].

In quantum key distribution, we introduce particular protocols to generate secret keys. These protocols outline the particulars of measurement choices, error correction, and privacy amplification codes. To test the effectiveness of the protocol, we calculate key rates obtained from these protocols for expected noise models. This is important since we want protocols to give high key rates for expected noise models. However, how do we know if this protocol is the best that one can have? It might be possible that with different measurement choices, error correction codes, or privacy amplification, one might be able to construct a better protocol. To understand these limitations, we rely on upper bounds on the key rates that can be generated from expected noise models. One such upper bound is given by the squashed entanglement of the state. The state for which the squashed entanglement is calculated is given by the noise model that we have for the unknown channel/device. We then compare the asymptotic key rates that we obtain from particular protocols to the calculated upper bounds.

Apart from squashed entanglement, relative entropy has also been proven to be an upper bound on the distillable key [5]. In certain cases, it is a tighter upper bound than squashed entanglement; in other cases the converse is true [99].

4.3. CMI based measures for steering and non-locality

In the sections above, we discussed the use of intrinsic information to upper bound the secret key rate generated from a joint probability distribution. We also discussed the upper bounds on the distillable key of quantum states in terms of squashed entanglement. The latter is useful in the trusted setting of quantum key distribution.

Now consider the setting of 1S-DI-QKD. We know that for 1S-DI-QKD, we require the underlying state to be a steerable state. We also know that there are entangled states which are not steerable and that squashed entanglement is a faithful measure. Therefore, we can deduce that squashed entanglement will not be a tight upper bound on the distillable secret key in the 1S-DI-QKD setting. A similar argument can be made for device-independent quantum key distribution. Therefore, we need to introduce different functions of assemblages and probability distributions to obtain better upper bounds for these settings. These functions should, at the very least, evaluate to zero for unsteerable assemblages and Bell local distributions. Now recall that an unsteerable assemblage is defined as follows:

$$\hat{\rho}_B^{a,x} = \sum_{\lambda} p_{\bar{A}|X}(a|x,\lambda)\rho_B^{\lambda}. \tag{4.8}$$

Associated to this assemblage, let us define the following state:

$$\rho_{\bar{A}XB\Lambda} = \sum_{a,x,\lambda} p_{\Lambda}(\lambda) p_{\bar{A}|X,\Lambda}(a|x,\lambda) \, |a,x,\lambda\rangle\!\langle a,x,\lambda|_{\bar{A}X\Lambda} \otimes \rho_B^{\lambda}. \tag{4.9}$$

For this state $I(A;B|X\Lambda)_{\rho} = 0$. This statement again relies on the property of CMI evaluating to zero for Markov chains.

Now, recall that a local distribution is defined as follows:

$$p(a,b|x,y) = \sum_{\lambda} p(\lambda)p(a|x,\lambda)p(b|y,\lambda). \tag{4.10}$$

Let us embed this distribution in the following state:

$$\sigma_{\bar{A}\bar{B}XY\Lambda} = \sum_{a,x,y,b,\lambda} p(\lambda)p(x)p(a|x,\lambda)p(b|y,\lambda)\,|a,b,x,y,\lambda\rangle\!\langle a,b,x,y,\lambda|_{\bar{A}\bar{B}XY\Lambda}\,. \tag{4.11}$$

For a local distribution, $I(\bar{A};\bar{B}|XY\Lambda)_{\sigma} = 0$.

This indicates the possibility of constructing CMI based measures for quantum steering and non-locality. In Chapter 5, we propose an information-theoretic quantifier for steering called intrinsic steerability, which uses conditional mutual information to measure the deviation of a given assemblage from one having a local hidden-state model. In Chapter 6, we introduce intrinsic non-locality as a quantifier for Bell non-locality, and we prove that it satisfies certain desirable properties such as faithfulness, convexity, and monotonicity under local operations and shared randomness. In Chapter 7, we prove that intrinsic steerability is an upper bound on the distillable key of an assemblage, and that intrinsic non-locality is an upper bound on the distillable key of a probability distribution.

Chapter 5
Intrinsic Steerability

In previous chapters, we discussed the role of conditional mutual information for quantifying secret correlations in noisy distributions and quantum correlations in bipartite quantum states. In this chapter, we introduce intrinsic steerability and restricted intrinsic steerability, which are measures of quantum correlations in an assemblage. We define intrinsic steerability and restricted intrinsic steerability in Section 5.1.. We then discuss the proofs of various properties of intrinsic steerability in Section 5.3. and restricted intrinsic steerability in Section 5.4.. We conclude with open questions regarding these quantities in Section 5.5.. The results in this chapter are based on [12]. As seen in Section 4.3., the definitions of steering quantifiers are inspired by the approach of [94, 95] to quantifying non-Markovianity in Bayesian networks, which in turn bears resemblance to squashed entanglement [10] and intrinsic information [9]. To see this, consider that correlations in any unsteerable assemblage can be explained by a hidden variable, which implies that such an assemblage has a Markov-chain structure. Assemblages with this structure have zero conditional mutual information when conditioning on the shared variable [100]. So our primary idea is to take a non-signaling extension of an assemblage, remove the correlations that can be explained by a shared variable (by conditioning), and then quantify the remaining intrinsic correlations.

We will eventually establish a connection between the quantifiers introduced in this chapter to 1S-DI-QKD. To this end, it is instructive to think of the conditioning system as being held by an eavesdropper. This is only necessary when we understand an assemblage as a resource in 1S-DI-QKD.

We recall here the definitions of an assemblage as discussed in Section 2.1.. An *assemblage* consists of the state of Bob's subsystem and the conditional probability of

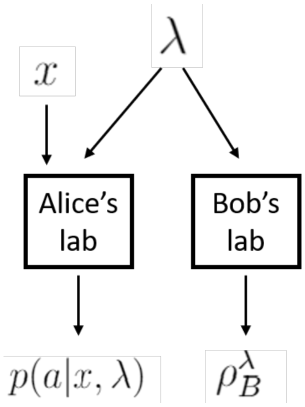

Figure 5.1. Markov chain structure in an unsteerable assemblage

Alice's outcome a (correlated with Bob's state) given the measurement choice x. This is specified as $\{p_{\bar{A}|X}(a|x), \rho_B^{a,x}\}_{a\in\mathcal{A},x\in\mathcal{X}}$. The sub-normalized state possessed by Bob is $\hat{\rho}_B^{a,x} := p_{\bar{A}|X}(a|x)\rho_B^{a,x}$. Taking $p_X(x)$ as a probability distribution over measurement choices, we can then embed the assemblage $\{\hat{\rho}_B^{a,x}\}_{a,x}$ in a classical-quantum state as follows:

$$\rho_{X\bar{A}B} := \sum_{a,x} p_X(x)\,|x\rangle\!\langle x|_X \otimes |a\rangle\!\langle a|_{\bar{A}} \otimes \hat{\rho}_B^{a,x}, \tag{5.1}$$

where $\{|x\rangle_X\}_x$ and $\{|a\rangle_{\bar{A}}\}_a$ are orthonormal bases.

5.1. Definitions of intrinsic steerability and restricted intrinsic steerability

As discussed in Section 2.2.2., the free operations allowed in the context of quantum steering are 1W-LOCC operations. We require this definition when defining intrinsic steerability. The intrinsic steerability of an assemblage is defined as follows:

Definition 42 (Intrinsic Steerability) *Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ denote an assemblage, and let $\rho_{X\bar{A}B'Y}$ be a state resulting from a 1W-LOCC operation. Consider a non-signaling* *extension $\rho_{X\bar{A}B'EY}$ of $\rho_{X\bar{A}B'Y}$ of the following form:*

$$\rho_{X\bar{A}B'EY} := \sum_{x,a,y} p_{X|Y}(x|y) \, |x\rangle\!\langle x|_X \otimes |a\rangle\!\langle a|_{\bar{A}} \otimes \hat{\rho}^{a,x,y}_{B'E} \otimes |y\rangle\!\langle y|_Y \,, \tag{5.2}$$

where $\hat{\rho}^{a,x,y}_{B'E}$ satisfies $\mathrm{Tr}_E(\hat{\rho}^{a,x,y}_{B'E}) = \mathcal{K}_y(\hat{\rho}^{a,x}_B)$ and the following no-signaling constraints:

$$\sum_a \hat{\rho}^{a,x,y}_{B'E} = \sum_a \hat{\rho}^{a,x',y}_{B'E} \ \forall x, x' \in \mathcal{X}, \ y \in \mathcal{Y}. \tag{5.3}$$

We define the intrinsic steerability of a given assemblage as follows:

$$S(\bar{A}; B)_{\hat{\rho}} := \sup_{\left\{p_{X|Y}, \{\mathcal{K}_y\}_y\right\}} \inf_{\rho_{X\bar{A}B'EY}} I(X\bar{A}; B'|EY)_\rho, \tag{5.4}$$

where the supremum is with respect to all quantum instruments, consisting of trace non-increasing maps $\{\mathcal{K}_y\}_y$ such that the sum map $\sum_y \mathcal{K}_y$ is trace preserving, and all classical channels $p_{X|Y}$ leading to Alice's input choice x. The infimum is with respect to all non-signaling extensions of $\rho_{X\bar{A}B'Y}$. Using the no-signaling constraints, which imply that $I(X; B'|EY)_\rho = 0$, we can write

$$S(\bar{A}; B)_{\hat{\rho}} := \sup_{\left\{p_{X|Y}, \{\mathcal{K}_y\}_y\right\}} \inf_{\rho_{X\bar{A}B'EY}} I(\bar{A}; B'|EXY)_\rho. \tag{5.5}$$

Definition 42 might seem rather complicated with the number of systems involved and the number of objects involved in the optimizations. While undesirable, we note that other steering quantifiers, such as the relative entropy of steering [37, 12], feature similar complications, and this seems unavoidable, having to do with the structure of assemblages and 1W-LOCC operations.

The idea behind the definition of intrinsic steerability is to measure the correlations between Alice and Bob's systems after conditioning on all systems that an eavesdropper could have. The worst possible scenario is that the eavesdropper possesses an arbitrary non-signaling extension of $\mathcal{K}_y(\hat{\rho}_B^{a,x})$. Here, we allow Alice and Bob first to pick a particular 1W-LOCC strategy to maximize their correlations, and Eve reacts to this strategy by choosing the extensions, to minimize their correlations. There could be another definition in which we first allow Eve to choose an extension. Then, Alice and Bob carry out a 1W-LOCC operation. However, this would mean a less powerful eavesdropper.

To simplify the rather complicated definition of intrinsic steerability, we are motivated to find alternate definitions. One such quantifier is *restricted intrinsic steerability*, defined as follows:

Definition 43 (Restricted Intrinsic Steerability) *Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ denote an assemblage, and let $\rho_{X\bar{A}B}$ denote a corresponding classical–quantum state. Consider a non-signaling extension $\rho_{X\bar{A}BE}$ of $\rho_{X\bar{A}B}$ of the following form:*

$$\rho_{X\bar{A}B'E} := \sum_{a,x} p_X(x) \, |x\rangle\langle x|_X \otimes |a\rangle\langle a|_{\bar{A}} \otimes \hat{\rho}_{BE}^{a,x}, \tag{5.6}$$

where $\hat{\rho}_{BE}^{a,x}$ satisfies $\mathrm{Tr}_E(\hat{\rho}_{BE}^{a,x}) = \hat{\rho}_B^{a,x}$ and the following no-signaling constraints:

$$\sum_a \hat{\rho}_{BE}^{a,x} = \sum_a \hat{\rho}_{BE}^{a,x'} \ \forall x, x' \in \mathcal{X}. \tag{5.7}$$

We define the restricted intrinsic steerability of $\{\hat{\rho}_B^{a,x}\}_{a,x}$ as follows:

$$S^R(\bar{A}; B)_{\hat{\rho}} := \sup_{p_X} \inf_{\rho_{X\bar{A}BE}} I(X\bar{A}; B|E)_{\rho}, \tag{5.8}$$

where the supremum is with respect to all probability distributions p_X and the infimum is with respect to all non-signaling extensions of $\rho_{X\bar{A}B}$. Using the no-signaling *constraints, which imply that* $I(X;B|E)_\rho = 0$*, it follows that*

$$S^R(\bar{A};B)_{\hat{\rho}} := \sup_{p_X} \inf_{\rho_{X\bar{A}BE}} I(\bar{A};B|EX)_\rho. \tag{5.9}$$

This definition is simpler because Eve's choice of extension is no longer dependent on the local operations performed by Alice and Bob on their initial assemblage. Instead, in this definition, Eve has an extension of the initial assemblage. There is also no supremum over the Kraus operators $\{\mathcal{K}_y\}_y$ in the definition of restricted intrinsic steerability. The drawback of this quantifier is that it is not known to be a monotone under 1W-LOCC operations. It is, however, a monotone under restricted 1W-LOCC operations.

By inspecting definitions, we can conclude that intrinsic steerability is never smaller than restricted intrinsic steerability: $S(\bar{A};B)_{\hat{\rho}} \geq S^R(\bar{A};B)_{\hat{\rho}}$. This follows because the restricted intrinsic steerability involves a supremization over particular 1W-LOCC strategies that are included in the supremization in the definition of the intrinsic steerability.

By invoking strong subadditivity of quantum entropy [17] and an upper bound in terms of the dimensions, we conclude that $0 \leq S(\bar{A};B)_{\hat{\rho}} \leq \log_2|\bar{A}|$. Similarly, using known bounds on conditional mutual information, the expression in (5.9), and the fact that taking an infimum over classical extensions E does not decrease $S^R(\bar{A};B)_{\hat{\rho}}$, we find that $0 \leq S^R(\bar{A};B)_{\hat{\rho}} \leq \min\{\log_2|\bar{A}|, \log_2|B|\}$.

One can also consider different constraints on an assemblage and its extensions. In the definition of intrinsic steerability and restricted intrinsic steerability, the eavesdropper has a non-signaling extension. This makes the eavesdropper compatible with the no-signaling theory. However, it is also possible to define a quantum extension system, which makes the eavesdropper compatible with quantum theory.

Given a bipartite no-signaling assemblage $\hat{\rho}_B^{a,x}$. From [29], it is possible to construct a quantum strategy for this assemblage. That is, we can always find a set of bipartite states and POVMs that give rise to a bipartite assemblage. Let a particular strategy be a quantum state ρ_{AB} and a POVM $\{\Lambda_x^a\}_a$. A particular quantum extension of the assemblage can be defined in terms of a particular quantum strategy as follows:

$$\hat{\rho}_{BE}^{a,x} = \mathrm{Tr}_A \left[(\Lambda_x^a \otimes I_{BE}) \rho_{ABE}) \right]. \tag{5.10}$$

It is possible to define a quantifier similar to intrinsic steerability in Definition 42, and restricted intrinsic steerability in Definition 43. However, for bipartite assemblages, these new definitions collapse to the original one by invoking the result in [29]. This is because the set of bipartite no-signaling extensions is equal to the set of bipartite quantum extensions. However, this difference would be crucial if one were to define intrinsic steerability over multipartite assemblages.

5.2. Examples

We now calculate the intrinsic steerability of an assemblage considered in Section 2.1.2.. Consider the following assemblage resulting from Pauli σ_Z or σ_X measurements on one share of a maximally entangled state $|\Phi\rangle_{AB} := (|00\rangle_{AB} + |11\rangle_{AB})/\sqrt{2}$, consisting of the following four subnormalized states: $\hat{\rho}_B^{a=0,x=0} = \frac{1}{2}|0\rangle\langle 0|_B$, $\hat{\rho}_B^{a=1,x=0} = \frac{1}{2}|1\rangle\langle 1|_B$, $\hat{\rho}_B^{a=0,x=1} = \frac{1}{2}|+\rangle\langle +|_B$, $\hat{\rho}_B^{a=1,x=1} = \frac{1}{2}|-\rangle\langle -|_B$. The non-signaling constraint imposes a constraint that any non-signaling extension of the above assemblage has the form $\hat{\rho}_B^{a,x} \otimes \omega_E$ for all $a, x \in \{0, 1\}$ and for some state ω_E (see proof of Example 44 below). Thus, in this sense this assemblage is unextendible, similar to maximally entangled states, and features a certain kind of monogamy against non-signaling adversaries. As a consequence, we find that this assemblage has exactly one bit of intrinsic steerability.

Example 44 *Consider a maximally entangled state*

$$|\Phi\rangle_{AB} := \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}). \tag{5.11}$$

Let measurement $x = 0$ be Pauli σ_Z on system A, with outcomes $a = 0$ and $a = 1$. Let measurement $x = 1$ be Pauli σ_X on system A, with outcomes $a = 0$ and $a = 1$. This leads to the following assemblage:

$$\left\{ \begin{array}{cc} \hat{\rho}_B^{a=0,x=0} = \frac{1}{2}|0\rangle\langle 0|_B, & \hat{\rho}_B^{a=1,x=0} = \frac{1}{2}|1\rangle\langle 1|_B, \\ \hat{\rho}_B^{a=0,x=1} = \frac{1}{2}|+\rangle\langle +|_B, & \hat{\rho}_B^{a=1,x=1} = \frac{1}{2}|-\rangle\langle -|_B \end{array} \right\}, \tag{5.12}$$

which has one bit of intrinsic steerability and restricted intrinsic steerability:

$$S(\bar{A}; B)_{\hat{\rho}} = S^R(\bar{A}; B)_{\hat{\rho}} = 1. \tag{5.13}$$

Proof. Arbitrary extensions of each of the above subnormalized states are as follows:

$$\begin{aligned} \hat{\rho}_{BE}^{a=0,x=0} &= \frac{1}{2}|0\rangle\langle 0|_B \otimes \omega_E^{00}, \quad \hat{\rho}_{BE}^{a=1,x=0} = \frac{1}{2}|1\rangle\langle 1|_B \otimes \omega_E^{10}, \\ \hat{\rho}_{BE}^{a=0,x=1} &= \frac{1}{2}|+\rangle\langle +|_B \otimes \omega_E^{01}, \quad \hat{\rho}_{BE}^{a=1,x=1} = \frac{1}{2}|-\rangle\langle -|_B \otimes \omega_E^{11}, \end{aligned} \tag{5.14}$$

where $\omega_E^{ij} \geq 0$ and $\mathrm{Tr}(\omega_E^{ij}) = 1$ for all $i, j \in \{0, 1\}$. The no-signaling constraint is as follows:

$$\hat{\rho}_{BE}^{a=0,x=0} + \hat{\rho}_{BE}^{a=1,x=0} = \hat{\rho}_{BE}^{a=0,x=1} + \hat{\rho}_{BE}^{a=1,x=1}. \tag{5.15}$$

Writing out the left-hand side of (5.15) in matrix form, we find that

$$\frac{1}{2}|0\rangle\langle 0|_B \otimes \omega_E^{00} + \frac{1}{2}|1\rangle\langle 1|_B \otimes \omega_E^{10} = \frac{1}{2}\begin{bmatrix} \omega_E^{00} & 0 \\ 0 & \omega_E^{10} \end{bmatrix}. \tag{5.16}$$

Writing out the right-hand side of (5.15) in matrix form, we find that

$$\begin{aligned} &\frac{1}{2}|+\rangle\langle +|_B \otimes \omega_E^{01} + \frac{1}{2}|-\rangle\langle -|_B \otimes \omega_E^{11} \\ &= \frac{1}{4}\left[|0\rangle\langle 0|_B + |1\rangle\langle 0|_B + |0\rangle\langle 1|_B + |1\rangle\langle 1|_B\right] \otimes \omega_E^{01} \\ &\quad + \frac{1}{4}\left[|0\rangle\langle 0|_B - |1\rangle\langle 0|_B - |0\rangle\langle 1|_B + |1\rangle\langle 1|_B\right] \otimes \omega_E^{11} && (5.17) \\ &= \frac{1}{2}|0\rangle\langle 0|_B \otimes \left(\frac{\omega_E^{01} + \omega_E^{11}}{2}\right) + \frac{1}{2}|1\rangle\langle 0|_B \otimes \left(\frac{\omega_E^{01} - \omega_E^{11}}{2}\right) \\ &\quad + \frac{1}{2}|0\rangle\langle 1|_B \otimes \left(\frac{\omega_E^{01} - \omega_E^{11}}{2}\right) + \frac{1}{2}|1\rangle\langle 1|_B \otimes \left(\frac{\omega_E^{01} + \omega_E^{11}}{2}\right) && (5.18) \\ &= \frac{1}{2}\begin{bmatrix} \frac{\omega_E^{01}+\omega_E^{11}}{2} & \frac{\omega_E^{01}-\omega_E^{11}}{2} \\ \frac{\omega_E^{01}-\omega_E^{11}}{2} & \frac{\omega_E^{01}+\omega_E^{11}}{2} \end{bmatrix}. && (5.19) \end{aligned}$$

So equating them, we find that the following equation (no-signaling constraint) should be satisfied

$$\begin{bmatrix} \omega_E^{00} & 0 \\ 0 & \omega_E^{10} \end{bmatrix} = \begin{bmatrix} \frac{\omega_E^{01}+\omega_E^{11}}{2} & \frac{\omega_E^{01}-\omega_E^{11}}{2} \\ \frac{\omega_E^{01}-\omega_E^{11}}{2} & \frac{\omega_E^{01}+\omega_E^{11}}{2} \end{bmatrix}. \tag{5.20}$$

This implies that $\omega_E^{01} = \omega_E^{11}$, which in turn implies that $\omega_E^{10} = \omega_E^{01} = \omega_E^{11} = \omega_E^{00}$. Thus, the only possible extension allowed in order to satisfy the no-signaling constraint is a product extension independent of a and x, meaning one of the following form:

$$\begin{aligned}\hat{\rho}_{BE}^{a=0,x=0} &= \frac{1}{2}|0\rangle\langle 0|_B \otimes \omega_E, \quad \hat{\rho}_{BE}^{a=1,x=0} = \frac{1}{2}|1\rangle\langle 1|_B \otimes \omega_E, \\ \hat{\rho}_{BE}^{a=0,x=1} &= \frac{1}{2}|+\rangle\langle +|_B \otimes \omega_E, \quad \hat{\rho}_{BE}^{a=1,x=1} = \frac{1}{2}|-\rangle\langle -|_B \otimes \omega_E, \end{aligned} \tag{5.21}$$

where $\omega_E \geq 0$ and $\mathrm{Tr}(\omega_E) = 1$. We can then evaluate the restricted intrinsic steerability in terms of the following classical–quantum state:

$$\left[\begin{array}{c} \frac{1}{2}|0\rangle\langle 0|_X \otimes |0\rangle\langle 0|_{\bar{A}} \otimes \frac{1}{2}|0\rangle\langle 0|_B + \frac{1}{2}|0\rangle\langle 0|_X \otimes |1\rangle\langle 1|_{\bar{A}} \otimes \frac{1}{2}|1\rangle\langle 1|_B \\ \\ +\frac{1}{2}|1\rangle\langle 1|_X \otimes |0\rangle\langle 0|_{\bar{A}} \otimes \frac{1}{2}|+\rangle\langle +|_B + \frac{1}{2}|1\rangle\langle 1|_X \otimes |1\rangle\langle 1|_{\bar{A}} \otimes \frac{1}{2}|-\rangle\langle -|_B \end{array}\right] \otimes \omega_E. \tag{5.22}$$

The conditional mutual information of this state is as follows:

$$I(X\bar{A};B|E) = I(X\bar{A};B) = H(B) - H(B|X\bar{A}) = H(B) = 1, \tag{5.23}$$

so that this assemblage has *one bit of restricted intrinsic steerability*. The first equality follows because the system E is product regardless of the extension, due to the above analysis with the no-signaling constraint. The second equality follows by expanding the mutual information. The third equality follows because the state of the B system is pure when conditioned on systems $X\bar{A}$. The final equality follows because the reduced state on the B system is maximally mixed. Also, it is clear that this is the maximum value of the restricted intrinsic steerability, given that it is always bounded from above by $\log\dim(\mathcal{H}_B)$ or $\log\dim(\mathcal{H}_{\bar{A}})$. By considering the upper bound $\log\dim(\mathcal{H}_{\bar{A}})$ for intrinsic steerability, we see that this assemblage achieves the upper bound on intrinsic steerability and thus has one bit of intrinsic steerability. ■

We generalize this to an assemblage resulting from an arbitrary pure bipartite state being measured in the Schmidt basis and the basis Fourier conjugate to this one. We find that this assemblage has the same kind of monogamy against non-signaling adversaries and that it has restricted intrinsic steerability equal to the entropy of entanglement [101] of the state being measured.

Example 45 *Consider a pure bipartite state $|\varphi\rangle_{AB}$ in its Schmidt basis:*

$$|\varphi\rangle_{AB} := \sum_{j=0}^{d-1} \alpha_j |j\rangle_A \otimes |j\rangle_B, \tag{5.24}$$

where $|\alpha_j| \neq 0$ for all $j \in \{0, \dots, d-1\}$. Let measurement $x = 0$ be a measurement $\{|j\rangle\langle j|_A\}_j$ in the Schmidt basis on system A, with outcomes $a = j \in \{0, \dots, d-1\}$. Let measurement $x = 1$ be a measurement $\{|\widetilde{j}\rangle\langle\widetilde{j}|_A\}_j$ in the Fourier conjugate basis, where

$$|\widetilde{j}\rangle_A := \frac{1}{\sqrt{d}} \sum_k e^{2\pi ijk/d} |k\rangle_A, \tag{5.25}$$

on system A, with outcomes $a = j \in \{0, \dots, d-1\}$. This leads to the following assemblage:

$$\left\{ \left\{ \hat{\rho}_B^{a=j,x=0} = |\alpha_j|^2 \, |j\rangle\langle j|_B \right\}_j, \, \left\{ \hat{\rho}_B^{a=j,x=1} = \frac{1}{d} Z^\dagger(j) |\psi\rangle\langle\psi|_B Z(j) \right\}_j \right\}, \tag{5.26}$$

where $|\psi\rangle_B := \sum_j \alpha_j |j\rangle_B$. This assemblage has $H(\{|\alpha_j|^2\}_j) = H(A)_\varphi$ bits of restricted intrinsic steerability. Note that this is equal to the entropy of entanglement of the state $|\varphi\rangle_{AB}$. If the state $|\varphi\rangle_{AB}$ is maximally entangled so that $\alpha_j = 1/\sqrt{d}$, then the resulting assemblage has $\log_2(d)$ bits of intrinsic steerability.

Proof. It is clear that the post-measurement state for Bob $\hat{\rho}_B^{a=j,x=0}$ is as above. For the other case, consider that

$$\langle\widetilde{j}|_A \otimes I_B|\varphi\rangle_{AB} = \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1} e^{-2\pi ijk/d}\langle k|_A \sum_{l=0}^{d-1}\alpha_l|l\rangle_A \otimes |l\rangle_B \tag{5.27}$$

$$= \frac{1}{\sqrt{d}}\sum_{k,l=0}^{d-1} \alpha_k e^{-2\pi ijk/d}\langle k|l\rangle_A \otimes |l\rangle_B \tag{5.28}$$

$$= \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1} \alpha_k e^{-2\pi ijk/d}|k\rangle_B. \tag{5.29}$$

Now defining the unitary operator $Z(j)$ by $Z(j)|k\rangle = e^{2\pi ijk/d}|k\rangle$ for $j,k \in \{0,\dots,d-1\}$, we can write

$$\langle\widetilde{j}|_A \otimes I_B|\varphi\rangle_{AB} = \frac{1}{\sqrt{d}}Z^\dagger(j)|\psi\rangle_B, \tag{5.30}$$

confirming the post-measurement subnormalized states $\hat{\rho}_B^{a=j,x=1}$. Arbitrary extensions of each of the above subnormalized states are as follows:

$$\hat{\rho}_{BE}^{a=j,x=0} = |\alpha_j|^2\,|j\rangle\langle j|_B \otimes \omega_E^j, \tag{5.31}$$

$$\hat{\rho}_{BE}^{a=j,x=1} = \frac{1}{d}Z^\dagger(j)|\psi\rangle\langle\psi|_B Z(j) \otimes \tau_E^j, \tag{5.32}$$

where $\omega_E^j, \tau_E^j \geq 0$ and $\mathrm{Tr}(\omega_E^j) = \mathrm{Tr}(\tau_E^j) = 1$ for all $j \in \{0,\dots,d-1\}$. The no-signaling constraint is as follows:

$$\sum_{j=0}^{d-1}\hat{\rho}_{BE}^{a=j,x=0} = \sum_{j=0}^{d-1}\hat{\rho}_{BE}^{a=j,x=1}, \tag{5.33}$$

which is the same as

$$\sum_{k=0}^{d-1}|k\rangle\langle k|_B \otimes |\alpha_k|^2\,\omega_E^k$$
$$= \sum_{j=0}^{d-1}\frac{1}{d}Z^\dagger(j)|\psi\rangle\langle\psi|_B Z(j) \otimes \tau_E^j \tag{5.34}$$

$$= \sum_{j,k,k'=0}^{d-1} \frac{1}{d} \alpha_k \alpha_{k'}^* e^{-2\pi i j(k-k')/d} |k\rangle\langle k'|_B \otimes \tau_E^j \tag{5.35}$$

$$= \sum_{k,k'=0}^{d-1} |k\rangle\langle k'|_B \otimes \frac{1}{d} \alpha_k \alpha_{k'}^* \sum_{j=0}^{d-1} e^{-2\pi i j(k-k')/d} \tau_E^j. \tag{5.36}$$

Set $k' = 0$. For $k \in \{0, 1, \ldots, d-1\}$, we get the following constraints from the no-signaling condition:

$$\omega_E^0 = \frac{1}{d} \sum_{j=0}^{d-1} \tau_E^j, \tag{5.37}$$

$$\sum_{j=0}^{d-1} e^{-2\pi i jk/d} \tau_E^j = 0. \tag{5.38}$$

We can conclude that τ_E^j is independent of j, so that $\tau_E^j = \omega_E^0$ for all $j \in \{0, \ldots, d-1\}$. To see this, let us solve the above equations, thinking of ω_E^0 as fixed and τ_E^j as free for all $j \in \{0, \ldots, d-1\}$. Consider that

$$\sum_{j=0}^{d-1} e^{-2\pi i jk/d} = 0 \quad \forall k \in \{1, \ldots, d-1\}. \tag{5.39}$$

Then we can see that $\tau_E^0 = \tau_E^1 = \ldots = \tau_E^{d-1} = \omega_E^0$ is one of the solutions of the equations in (5.37)–(5.38). Since the equations are linearly independent, it is a unique solution. Now considering the other blocks in (5.34) (i.e., for $k = k' = 1, \ldots, d-1$), we find that $\omega_E^1 = \ldots = \omega_E^{d-1} = \omega_E^0$. Thus, the only possible extension allowed in order to satisfy the no-signaling constraint is a product extension independent of a and x, meaning one of the following form:

$$\hat{\rho}_{BE}^{a=j,x=0} = |\alpha_j|^2 \, |j\rangle\langle j|_B \otimes \omega_E, \tag{5.40}$$

$$\hat{\rho}_{BE}^{a=j,x=1} = \frac{1}{d} Z^\dagger(j) |\psi\rangle\langle\psi|_B Z(j) \otimes \omega_E, \tag{5.41}$$

where $\omega_E \geq 0$ and $\text{Tr}(\omega_E) = 1$. We can then evaluate the restricted intrinsic steerability in terms of the following classical–quantum state:

$$\Bigg[p|0\rangle\langle 0|_X \otimes \sum_j |j\rangle\langle j|_{\bar{A}} \otimes |\alpha_j|^2 \, |j\rangle\langle j|_B + (1-p) \, |1\rangle\langle 1|_X \otimes \\ \sum_j |j\rangle\langle j|_{\bar{A}} \otimes \frac{1}{d} Z^\dagger(j)|\psi\rangle\langle\psi|_B Z(j) \Bigg] \otimes \omega_E, \quad (5.42)$$

where $(p, 1-p)$ is a probability distribution for the input x. The conditional mutual information of this state is as follows:

$$I(X\bar{A}; B|E) = I(X\bar{A}; B) = H(B) - H(B|X\bar{A}) \quad (5.43)$$

$$= H(B) = H(\{|\alpha_j|^2\}), \quad (5.44)$$

so that this assemblage has $H(\{|\alpha_j|^2\})$ *bits of restricted intrinsic steerability*. The first step follows because the system E is product regardless of the extension, due to the above analysis with the no-signaling constraint. The second step follows by expanding the mutual information. The third step follows because the state of the B system is pure when conditioned on systems $X\bar{A}$. The final step follows because the reduced state on the B system is $\sum_j |\alpha_j|^2 \, |j\rangle\langle j|_B$, which can be seen from

$$\text{Tr}_{X\bar{A}} \Bigg(p|0\rangle\langle 0|_X \otimes \sum_j |j\rangle\langle j|_{\bar{A}} \otimes |\alpha_j|^2 \, |j\rangle\langle j|_B + (1-p) \, |1\rangle\langle 1|_X \otimes \\ \sum_j |j\rangle\langle j|_{\bar{A}} \otimes \frac{1}{d} Z^\dagger(j)|\psi\rangle\langle\psi|_B Z(j) \Bigg) \\ = p \sum_j |\alpha_j|^2 \, |j\rangle\langle j|_B + (1-p) \sum_j \frac{1}{d} Z^\dagger(j)|\psi\rangle\langle\psi|_B Z(j) \quad (5.45)$$

$$= p\sum_j |\alpha_j|^2 |j\rangle\langle j|_B + (1-p)\sum_j |\alpha_j|^2 |j\rangle\langle j|_B \tag{5.46}$$

$$= \sum_j |\alpha_j|^2 |j\rangle\langle j|_B. \tag{5.47}$$

This state is independent of the input probability distribution, so that the maximum is achieved for any choice of $p \in (0,1)$.

If the state $|\varphi\rangle_{AB}$ is maximally entangled, then $H(\{|\alpha_j|^2\}) = \log_2(d)$. Given the upper bound $\log(\dim(\mathcal{H}_{\bar{A}})) = \log_2(d)$ on intrinsic steerability, we see that the upper bound is achieved in this case. ■

5.3. Properties of intrinsic steerability

In this section, we prove that intrinsic steerability is a steering monotone.

Theorem 46 *The intrinsic steerability $S(\bar{A};B)_{\hat{\rho}}$ is a convex steering monotone. That is, it does not increase on average under deterministic 1W-LOCC, it vanishes for an assemblage having a local-hidden-state model, and it is convex.*

Proof. The proof of the theorem follows from Proposition 47, Proposition 48, and Proposition 50. ■

Proposition 47 *Intrinsic steerability vanishes for assemblages having a LHS model.*

Proof. To prove this, consider a particular non-signaling extension for an assemblage with a local-hidden-state model as follows:

$$\sum_{x,a,\lambda,y} p_{X|Y}(x|y)\, |x\rangle\!\langle x|_X \otimes p_{\bar{A}|X\Lambda}(a|x,\lambda)\, |a\rangle\!\langle a|_{\bar{A}} \otimes \sum_t K_{y,t}\hat{\rho}^{\lambda}_B K^{\dagger}_{y,t} \otimes p_\Lambda(\lambda)\, |\lambda\rangle\!\langle\lambda|_E \otimes |y\rangle\!\langle y|_Y\,. \tag{5.48}$$

For this non-signaling extension, conditioned on the values λ and y, systems $X\bar{A}$ and B' are in a product state, so that the conditional mutual information $I(X\bar{A};B'|EY)$

vanishes. Since the argument holds for all quantum instruments $\{\mathcal{K}_y\}_y$ and channels $p_{X|Y}$, then $S(\bar{A}; B)_\rho = 0$. ■

We now prove that intrinsic steerability is a monotone under 1W-LOCC operations. This condition is essential from a resource-theoretic perspective: a quantifier of the resource should not increase under free operations.

Proposition 48 (1W-LOCC monotone) *Let* $\{\hat{\rho}_B^{a,x}\}_{a,x}$ *be an assemblage, and suppose that*

$$\left\{\hat{\rho}_{B_f,z}^{a_f,x_f} := \sum_{a,x} p(a_f|x_f, x, a, z)p(x|x_f, z)\mathcal{K}_z(\hat{\rho}_B^{a,x})/p(z)\right\}_{a_f,x_f}, \qquad (5.49)$$

is an assemblage that arises from it by the action of a general 1W-LOCC operation, where

$$p(z) := \operatorname{Tr}\left(\mathcal{K}_z\left(\sum_a \hat{\rho}_B^{a,x}\right)\right) = \operatorname{Tr}(\mathcal{K}_z(\rho_B)). \qquad (5.50)$$

Then,

$$\sum_z p(z)S(\bar{A}_f; B_f)_{\hat{\rho}_z} \leq S(\bar{A}; B)_{\hat{\rho}}. \qquad (5.51)$$

Proof. First, we give a proof sketch for the monotonicity of intrinsic steerability on average under deterministic 1W-LOCC: $S(\bar{A}; B)_{\hat{\rho}} \geq \sum_z p_Z(z)S(\bar{A}_f; B_f)_{\hat{\rho}_z}$, where $\hat{\rho}_z := \{\hat{\rho}_{B_f,z}^{a_f,x_f}\}_{a_f,x_f}$ is the assemblage resulting from a 1W-LOCC operation on the initial assemblage $\{\hat{\rho}_B^{a,x}\}_{a,x}$ and is given as follows [37]:

$$\hat{\rho}_{B_f,z}^{a_f,x_f} := \sum_{a,x} p(a_f|a, x, x_f, z)p(x|x_f, z)\mathcal{K}_z(\hat{\rho}_B^{a,x}). \qquad (5.52)$$

In the above, $p(a_f|a, x, x_f, z)$ and $p(x|x_f, z)$ are local classical channels that Alice uses, respectively, to pick the output a_f of the final assemblage and the input x to her initial assemblage. The set $\{\mathcal{K}_z\}_z$ of completely positive maps is such that the sum map $\sum_z \mathcal{K}_z$ is trace preserving and thus corresponds to a quantum instrument acting on Bob's system. The definition of the intrinsic steerability involves a supremum over measurements of the system B_f of the final assemblage and classical channels for the input X_f to the final assemblage. Using data processing and when given Z, we can say that system $\bar{A}_f$ was obtained by processing systems $XX_f\bar{A}$. Then, the two successive measurements on Bob's system can be thought of as a single measurement. Since the intrinsic steerability involves a supremum over all possible measurements, the result follows.

We now give a detailed proof. To see this, consider that, in accordance with the definition of $S(\bar{A}_f; B_f)_{\hat{\rho}_z}$, the assemblages $\{\hat{\rho}^{a_f,x_f}_{B_f,z}\}_{a_f,x_f}$ can be further preprocessed by a z-dependent 1W-LOCC $\{p_{X_f|YZ=z}, \{\mathcal{L}^{(z)}_y\}_y\}$, resulting in the following state:

$$\sigma^z_{X_f\bar{A}_fB'_fY} := \sum_{a_f,x_f,y} p(x_f|yz)[x_f] \otimes [a_f] \otimes \mathcal{L}^{(z)}_y(\hat{\rho}^{a_f,x_f}_{B_f,z}) \otimes [y]. \tag{5.53}$$

Notation 49 *In the above and in what follows, we employ a shorthand* $[x] \equiv |x\rangle\langle x|_X$ *or* $[a] \equiv |a\rangle\langle a|_{\bar{A}}$*, etc.*

The state in (5.53) is extended by the following one:

$$\begin{aligned}\sigma^z_{X_fX\bar{A}_f\bar{A}B'_fY} := \sum_{a_f,a,x,x_f,y} & p(x_f|yz)[x_f] \otimes p(x|x_f,z)[x] \otimes \\ & p(a_f|x_f,x,a,z)[a_f] \otimes [a] \otimes \frac{\mathcal{L}^{(z)}_y(\hat{\rho}^{a,x}_B)}{p(z)} \otimes [y],\end{aligned} \tag{5.54}$$

which in turn are elements of the following classical–quantum state:

$$\sigma_{X_fX\bar{A}_f\bar{A}B'_fYZ} := \sum_z \sigma^z_{X_fX\bar{A}_f\bar{A}B'_fY} \otimes p(z)[z]. \tag{5.55}$$

An *arbitrary* non-signaling extension of the state in (5.53), according to that needed in the definition of $S(\bar{A}_f; B_f)_{\hat{\rho}_z}$, is as follows:

$$\sigma^z_{X_f \bar{A}_f B'_f EY} := \sum_{a_f, x_f, y} p(x_f|yz)[x_f] \otimes [a_f] \otimes \hat{\tau}^{a_f, x_f, y, z}_{B'_f E} \otimes [y], \tag{5.56}$$

where $\hat{\tau}^{a_f, x_f, y, z}_{B'_f E}$ satisfies

$$\mathrm{Tr}_E(\hat{\tau}^{a_f, x_f, y, z}_{B'_f E}) = \mathcal{L}^{(z)}_y(\hat{\rho}^{a_f, x_f}_{B_f, z}), \tag{5.57}$$

$$\sum_{a_f} \hat{\tau}^{a_f, x_f, y, z}_{B'_f E} = \sum_{a_f} \hat{\tau}^{a_f, x'_f, y, z}_{B'_f E}$$

$$\forall x_f, x'_f \in \mathcal{X}_f,\ y \in \mathcal{Y},\ z \in \mathcal{Z}. \tag{5.58}$$

A *particular* non-signaling extension of the state in (5.53), according to that needed in the definition of $S(\bar{A}_f; B_f)_{\hat{\rho}_z}$, is as follows:

$$\begin{aligned}\zeta^z_{X_f \bar{A}_f B'_f EY} := \sum_{a_f, x_f, y} p(x_f|yz)[x_f] \otimes [a_f] \\ \otimes \sum_{a,x} p(a_f|x_f, x, a, z) p(x|x_f, z) \hat{\omega}^{a,x,y,z}_{B'_f E} \otimes [y],\end{aligned} \tag{5.59}$$

where $\hat{\omega}^{a,x,y,z}_{B'_f E}$ satisfies

$$\mathrm{Tr}_E(\hat{\omega}^{a,x,y,z}_{B'_f E}) = \frac{\mathcal{L}^{(z)}_y(\mathcal{K}_z(\hat{\rho}^{a,x}_B))}{p(z)}, \tag{5.60}$$

$$\sum_a \hat{\omega}^{a,x,y,z}_{B'_f E} = \sum_a \hat{\omega}^{a,x,y,z}_{B'_f E} \quad \forall x, x' \in \mathcal{X},\ y \in \mathcal{Y},\ z \in \mathcal{Z}. \tag{5.61}$$

The operator $\hat{\omega}^{a,x,y,z}_{B'_f E}$ will serve as an arbitrary non-signaling extension needed in the definition of $S(\bar{A};B)_{\hat{\rho}}$. Let $\zeta_{X_f \bar{A}_f B'_f EYZ}$ denote the following state:

$$\zeta_{X_f \bar{A}_f B'_f EYZ} := \sum_z \zeta^z_{X_f \bar{A}_f B'_f EY} \otimes p(z)[z]. \tag{5.62}$$

This in turn is a marginal of the following state:

$$\zeta_{X_f X \bar{A}_f \bar{A} B'_f EYZ} := \sum_{a_f,a,x_f,x,y} p(x_f|yz)[x_f] \otimes p(x|x_f,z)[x] \otimes \\ p(a_f|x_f,x,a,z)[a_f] \otimes [a] \otimes \hat{\omega}^{a,x,y,z}_{B'_f E} \otimes [y] \otimes p(z)[z]. \tag{5.63}$$

Consider that

$$\sum_z p(z) \inf_{\text{ext. in (5.56)}} I(X_f \bar{A}_f; B'_f|EY)_{\sigma^z}$$

$$\leq \sum_z p(z) I(X_f \bar{A}_f; B'_f|EY)_{\zeta^z} \tag{5.64}$$

$$= I(X_f \bar{A}_f; B'_f|EYZ)_{\zeta} \tag{5.65}$$

$$\leq I(X_f X \bar{A}; B'_f|EYZ)_{\zeta} \tag{5.66}$$

$$= I(X\bar{A}; B'_f|EYZ)_{\zeta} + I(X_f; B'_f|EYZX\bar{A})_{\zeta} \tag{5.67}$$

$$= I(X\bar{A}; B'_f|EYZ)_{\zeta}. \tag{5.68}$$

The first inequality follows because the extension state $\zeta^z_{X_f \bar{A}_f B'_f EY}$ is a particular kind of non-signaling extension required in the definition of $S(\bar{A}_f;B_f)_{\hat{\rho}_z}$. The first equality follows because system Z is classical and thus can be incorporated as a conditioning system in the conditional mutual information. The second inequality follows from local data processing for the conditional mutual information: given Z, the system $\bar{A}_f$ arises from local processing of systems $X_f X \bar{A}$. The second equality follows from the chain rule for conditional mutual information. The final equality follows from the fact that systems $B'_f E$ are independent of X_f when given the classical systems $YZX\bar{A}$ (one can inspect the state in (5.63) to see this explicitly). Since the above chain of inequalities holds for any non-signaling extension of the form in (5.59), we can conclude that

$$\sum_z p(z) \inf_{\text{ext. in (5.56)}} I(X_f\bar{A}_f; B'_f|EY)_{\sigma^z} \leq \inf_{\text{ext. in (5.59)}} I(X\bar{A}; B'_f|EYZ)_\zeta. \tag{5.69}$$

Now we can take the supremum of both sides with respect to 1W-LOCC operations $\{p_{X_f|YZ=z}, \{\mathcal{L}_y^{(z)}\}_y\}_z$ and we find that

$$\sup_{\{p_{X_f|YZ=z},\{\mathcal{L}_y^{(z)}\}_y\}_z} \sum_z p(z) \inf_{\text{ext. in (5.56)}} I(X_f\bar{A}_f; B'_f|EY)_{\sigma^z} \\ \leq \sup_{\{p_{X_f|YZ=z},\{\mathcal{L}_y^{(z)}\}_y\}_z} \inf_{\text{ext. in (5.59)}} I(X\bar{A}; B'_f|EYZ)_\zeta. \tag{5.70}$$

Since the 1W-LOCC operation $\{p_{X_f|YZ=z}, \{\mathcal{L}_y^{(z)}\}_y\}_z$ is a particular 1W-LOCC operation that can be performed on the original assemblage $\{\hat{\rho}_B^{a,x}\}_{a,x}$, we find that

$$\sup_{\{p_{X_f|YZ=z},\{\mathcal{L}_y^{(z)}\}_y\}_z} \inf_{\text{ext. in (5.59)}} I(X\bar{A}; B'_f|EYZ)_\zeta \leq S(\bar{A}; B)_{\hat{\rho}}. \tag{5.71}$$

Since each z-dependent 1W-LOCC operation $\{p_{X_f|YZ=z}, \{\mathcal{L}_y^{(z)}\}_y\}$ depends only on a particular value of z, we can then exchange the supremum and the sum over z in (5.70) to conclude that

$$\sup_{\{p_{X_f|YZ=z},\{\mathcal{L}_y^{(z)}\}_y\}_z} \sum_z p(z) \inf_{\text{ext. in (5.56)}} I(X_f\bar{A}_f; B'_f|EY)_{\sigma^z} \\ = \sum_z p(z) \sup_{\{p_{X_f|YZ=z},\{\mathcal{L}_y^{(z)}\}_y\}} \inf_{\text{ext. in (5.56)}} I(X_f\bar{A}_f; B'_f|EY)_{\sigma^z} \tag{5.72}$$

$$= \sum_z p(z) S(\bar{A}_f; B_f)_{\hat{\rho}_z}. \tag{5.73}$$

This concludes the proof. ■

We now prove that the intrinsic steerability is convex. The physical interpretation of this statement is that steering cannot increase when mixing two assemblages.

Proposition 50 (Convexity) *Let* $\{\hat{\rho}_B^{a,x}\}_{a,x}$ *and* $\{\hat{\sigma}_B^{a,x}\}_{a,x}$ *be assemblages, and let* $\lambda \in [0,1]$*. Let* $\{\hat{\tau}_B^{a,x}\}_{a,x}$ *be a mixture of the two assemblages, defined as*

$$\hat{\tau}_B^{a,x} := \lambda \hat{\rho}_B^{a,x} + (1-\lambda)\hat{\sigma}_B^{a,x}. \tag{5.74}$$

Then

$$S(\bar{A}; B)_{\hat{\tau}} \leq \lambda S(\bar{A}; B)_{\hat{\rho}} + (1-\lambda) S(\bar{A}; B)_{\hat{\sigma}}. \tag{5.75}$$

Proof. We first give a proof sketch for the convexity of intrinsic steerability. Let $\lambda \in [0,1]$. Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ and $\{\hat{\sigma}_B^{a,x}\}_{a,x}$ be two assemblages, and consider an assemblage $\{\hat{\tau}_B^{a,x} := \lambda \hat{\rho}_B^{a,x} + (1-\lambda)\hat{\sigma}_B^{a,x}\}_{a,x}$. Convexity of the intrinsic steerability is the following statement: $S(\bar{A}; B)_{\hat{\tau}} \leq \lambda S(\bar{A}; B)_{\hat{\rho}} + (1-\lambda) S(\bar{A}; B)_{\hat{\sigma}}$. A proof for convexity is similar to known proofs for the convexity of squashed entanglement [10]. To prove convexity, first consider arbitrary non-signaling extensions of $\{\hat{\rho}_B^{a,x}\}_{a,x}$ and $\{\hat{\sigma}_B^{a,x}\}_{a,x}$. Embedding these in a larger classical–quantum state with a label chosen according to λ gives a particular non-signaling extension of $\hat{\tau}$. Convexity then follows from a property of conditional mutual information and because the intrinsic steerability involves an infimum over all non-signaling extensions.

We now give a detailed proof. Let $\{p_{X|Y}, \{\mathcal{K}_y\}_y\}$ denote an arbitrary 1W-LOCC operation, which leads to the following classical–quantum state:

$$\tau_{X\bar{A}B'Y} := \sum_{a,x,y} p_{X|Y}(x|y)\, |x\rangle\!\langle x|_X \otimes |a\rangle\!\langle a|_{\bar{A}} \otimes \mathcal{K}_y(\hat{\tau}^{a,x}_B) \otimes |y\rangle\!\langle y|_Y \,. \tag{5.76}$$

An *arbitrary* non-signaling extension of this state, is as follows:

$$\tau_{X\bar{A}B'YE} := \sum_{a,x,y} p_{X|Y}(x|y)\, |x\rangle\!\langle x|_X \otimes |a\rangle\!\langle a|_{\bar{A}} \otimes \hat{\tau}^{a,x,y}_{B'E} \otimes |y\rangle\!\langle y|_Y \,, \tag{5.77}$$

where

$$\mathrm{Tr}_E(\hat{\tau}^{a,x,y}_{B'E}) = \mathcal{K}_y(\hat{\tau}^{a,x}_B), \tag{5.78}$$

$$\sum_a \hat{\tau}^{a,x,y}_{B'E} = \sum_a \hat{\tau}^{a,x',y}_{B'E} \qquad \forall x,x' \in \mathcal{X},\ y \in \mathcal{Y}. \tag{5.79}$$

Let $\hat{\rho}^{a,x,y}_{B'E}$ and $\hat{\sigma}^{a,x,y}_{B'E}$ be *arbitrary* non-signaling extensions of $\mathcal{K}_y(\hat{\rho}^{a,x}_B)$ and $\mathcal{K}_y(\hat{\sigma}^{a,x}_B)$, satisfying

$$\mathrm{Tr}_E(\hat{\rho}^{a,x,y}_{B'E}) = \mathcal{K}_y(\hat{\rho}^{a,x}_B), \tag{5.80}$$

$$\sum_a \hat{\rho}^{a,x,y}_{B'E} = \sum_a \hat{\rho}^{a,x',y}_{B'E} \qquad \forall x,x' \in \mathcal{X},\ y \in \mathcal{Y}, \tag{5.81}$$

$$\mathrm{Tr}_E(\hat{\sigma}^{a,x,y}_{B'E}) = \mathcal{K}_y(\hat{\sigma}^{a,x}_B), \tag{5.82}$$

$$\sum_a \hat{\sigma}^{a,x,y}_{B'E} = \sum_a \hat{\sigma}^{a,x',y}_{B'E} \qquad \forall x,x' \in \mathcal{X},\ y \in \mathcal{Y}. \tag{5.83}$$

These lead to the following states:

$$\rho_{X\bar{A}B'YE} := \sum_{a,x,y} p_{X|Y}(x|y)\, |x\rangle\!\langle x|_X \otimes |a\rangle\!\langle a|_{\bar{A}} \otimes \hat{\rho}^{a,x,y}_{B'E} \otimes |y\rangle\!\langle y|_Y \,, \tag{5.84}$$

$$\sigma_{X\bar{A}B'YE} := \sum_{a,x,y} p_{X|Y}(x|y)\, |x\rangle\!\langle x|_X \otimes |a\rangle\!\langle a|_{\bar{A}} \otimes \hat{\sigma}^{a,x,y}_{B'E} \otimes |y\rangle\!\langle y|_Y \,. \tag{5.85}$$

A *particular* non-signaling extension $\tau'_{X\bar{A}B'YEE'}$ of $\tau_{\bar{A}B'XY}$, given by

$$\tau'_{X\bar{A}B'YEE'} := \sum_{a,x,y} p_{X|Y}(x|y) \, |x\rangle\!\langle x|_X \otimes |a\rangle\!\langle a|_{\bar{A}} \otimes \\ (\lambda \hat{\rho}^{a,x,y}_{B'E} \otimes |0\rangle\!\langle 0|_{E'} + (1-\lambda)\hat{\sigma}^{a,x,y}_{B'E} \otimes |1\rangle\!\langle 1|_{E'}) \otimes |y\rangle\!\langle y|_Y . \quad (5.86)$$

Then consider that

$$\inf_{\text{ext. in (5.77)}} I(X\bar{A};B'|EY)_\tau \leq I(X\bar{A};B'|EYE')_{\tau'} \quad (5.87)$$

$$= \lambda I(X\bar{A};B'|EY)_\rho + (1-\lambda) I(X\bar{A};B'|EY)_\sigma. \quad (5.88)$$

Since the inequality above holds for all general non-signaling extensions of the form in (5.84) and (5.85), we conclude that

$$\inf_{\text{ext. in (5.77)}} I(X\bar{A};B'|EY)_\tau \\ \leq \lambda \inf_{\text{ext. in (5.84)}} I(X\bar{A};B'|EY)_\rho + (1-\lambda) \inf_{\text{ext. in (5.85)}} I(X\bar{A};B'|EY)_\sigma. \quad (5.89)$$

Now taking a supremum over all 1W-LOCC operations, we find that

$$S(\bar{A};B)_{\hat{\tau}}$$

$$= \sup_{\left\{p_{X|Y},\{\mathcal{K}_y\}_y\right\}} \inf_{\text{ext. in (5.77)}} I(X\bar{A};B'|EY)_\tau \quad (5.90)$$

$$\leq \sup_{\left\{p_{X|Y},\{\mathcal{K}_y\}_y\right\}} \left(\lambda \inf_{\text{ext. in (5.84)}} I(X\bar{A};B'|EY)_\rho + (1-\lambda) \inf_{\text{ext. in (5.85)}} I(X\bar{A};B'|EY)_\sigma\right) \quad (5.91)$$

$$\leq \lambda \sup_{\left\{p_{X|Y},\{\mathcal{K}_y\}_y\right\}} \inf_{\text{ext. in (5.84)}} I(X\bar{A};B'|EY)_\rho$$

$$+ (1-\lambda) \sup_{\left\{p_{X|Y},\{\mathcal{K}_y\}_y\right\}} \inf_{\text{ext. in (5.85)}} I(X\bar{A};B'|EY)_\sigma \tag{5.92}$$

$$= \lambda S(\bar{A};B)_{\hat{\rho}} + (1-\lambda) S(\bar{A};B)_{\hat{\sigma}}. \tag{5.93}$$

This concludes the proof. ■

We now consider a superadditivity property of assemblages, which holds for intrinsic steerability. Suppose that Alice has two quantum systems A_1 and A_2 and suppose that Bob has two quantum systems B_1 and B_2. Alice could perform a local measurement on A_1 chosen according to x_1 and with output a_1. Similarly, Alice could perform a local measurement on A_2 chosen according to x_2 and with output a_2. This process realizes a joint assemblage $\{\hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2}\}_{a_1,a_2,x_1,x_2}$ obeying certain no-signaling constraints, but it also realizes some local assemblages as well. One would expect that the steering available in the joint assemblage should never be smaller than the sum of the steering available in the local assemblages, and this is what the following proposition addresses:

Proposition 51 (Superadditivity) *Let $\{\hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2}\}_{a_1,a_2,x_1,x_2}$ be an assemblage for which the following additional no-signaling constraints hold*

$$\sum_{a_2} \hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2} = \sum_{a_2} \hat{\rho}^{a_1,a_2,x_1,x_2'}_{B_1B_2} := \hat{\theta}^{a_1,x_1}_{B_1B_2} \qquad \forall x_2, x_2',$$

$$\sum_{a_1} \hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2} = \sum_{a_1} \hat{\rho}^{a_1,a_2,x_1',x_2}_{B_1B_2} := \hat{\kappa}^{a_2,x_2}_{B_1B_2} \qquad \forall x_1, x_1',$$

Let $\{\operatorname{Tr}_{B_2}(\hat{\theta}^{a_1,x_1}_{B_1B_2})\}_{a_1,x_1}$ and $\{\operatorname{Tr}_{B_1}(\hat{\kappa}^{a_2,x_2}_{B_1B_2})\}_{a_2,x_2}$ be reduced, local assemblages arising from the joint assemblage $\{\hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2}\}_{a_1,a_2,x_1,x_2}$. Then intrinsic steerability is superadditive in the following sense:

$$S(\bar{A}_1\bar{A}_2;B_1B_2)_{\hat{\rho}} \geq S(\bar{A}_1;B_1)_{\hat{\theta}} + S(\bar{A}_2;B_2)_{\hat{\kappa}}. \tag{5.94}$$

Proof. The core idea behind our proof of Proposition 51 is to exploit the chain rule for conditional mutual information. First, pick a 1W-LOCC strategy where Alice's inputs X_1 and X_2 depend only on measurement outcomes Y_1 and Y_2 of B_1 and B_2, respectively. The chain rule and non-negativity of conditional mutual information imply that

$$I(X_1X_2\bar{A}_1\bar{A}_2;B_1B_2|EY_1Y_2)_\rho \geq I(X_1\bar{A}_1;B_1|EY_1Y_2)_\rho + I(X_2\bar{A}_2;B_2|EB_1Y_1Y_2)_\rho, \tag{5.95}$$

where system E denotes a non-signaling extension system. The idea is then to take EY_2 as a non-signaling extension for $X_1\bar{A}_1B_1Y_1$, systems EB_1Y_1 as a non-signaling extension for $X_2\bar{A}_2B_2Y_2$, and work from there.

For the proof, suppose that we apply to the assemblage $\{\hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2}\}_{a_1,a_2,x_1,x_2}$ a general 1W-LOCC operation $\{p_{X_1X_2|Y},\{\mathcal{K}_y\}_y\}$, resulting in the following classical–quantum state:

$$\rho_{\bar{A}_1X_1\bar{A}_2X_2YB_1'B_2'} := \sum_{a_1,x_1,a_2,x_2,y} p_{X_1X_2|Y}(x_1,x_2|y)[a_1]\otimes[x_1]\otimes[a_2]\otimes[x_2]\otimes[y]\otimes\mathcal{K}_y(\hat{\rho}^{a_1,x_1,a_2,x_2}_{B_1B_2}). \tag{5.96}$$

Let $\hat{\rho}^{a_1,x_1,a_2,x_2,y}_{B_1'B_2'E}$ be a non-signaling extension of $\mathcal{K}_y(\rho^{a_1,x_1,a_2,x_2}_{B_1B_2})$ and consider the following extension of the above state:

$$\rho_{\bar{A}_1X_1\bar{A}_2X_2YB_1'B_2'} := \sum_{a_1,x_1,a_2,x_2,y} p_{X_1X_2|Y}(x_1,x_2|y)[a_1]\otimes[x_1]\otimes[a_2]\otimes[x_2]\otimes[y]\otimes\hat{\rho}^{a_1,x_1,a_2,x_2,y}_{B_1'B_2'E}. \tag{5.97}$$

A particular "product" 1W-LOCC operation has the form $\{p_{X_1|Y_1}p_{X_2|Y_2},\{\mathcal{L}_{y_1}\otimes\mathcal{M}_{y_2}\}_{y_1,y_2}\}$ and results in the following state:

$$\omega_{\bar{A}_1X_1\bar{A}_2X_2Y_1Y_2B_1'B_2'} := \sum_{a_1,x_1,a_2,x_2,y} p_{X_1|Y_1}(x_1|y_1)p_{X_2|Y_2}(x_2|y_2)[a_1]\otimes[x_1]\otimes[a_2]\otimes[x_2]\otimes[y_1]$$

$$\otimes [y_2] \otimes (\mathcal{L}_{y_1} \otimes \mathcal{M}_{y_2})\, \hat{\rho}^{a_1,x_1,a_2,x_2}_{B_1B_2}. \quad (5.98)$$

Let $\hat{\omega}^{a_1,x_1,a_2,x_2,y_1,y_2}_{B_1'B_2'E}$ be a non-signaling extension of $(\mathcal{L}_{y_1} \otimes \mathcal{M}_{y_2})\, (\hat{\rho}^{a_1,x_1,a_2,x_2}_{B_1B_2})$, and define the following state:

$$\omega_{\bar{A}_1X_1\bar{A}_2X_2Y_1Y_2B_1'B_2'E} := \sum_{a_1,x_1,a_2,x_2,y} p_{X_1|Y_1}(x_1|y_1)p_{X_2|Y_2}(x_2|y_2)[a_1]$$
$$\otimes [x_1] \otimes [a_2] \otimes [x_2] \otimes [y_1] \otimes [y_2] \otimes \hat{\omega}^{a_1,x_1,a_2,x_2,y_1,y_2}_{B_1'B_2'E}. \quad (5.99)$$

Let $\hat{\theta}^{a_1,x_1,y_1}_{B_1'F}$ be a non-signaling extension of $\mathcal{L}_{y_1}(\hat{\theta}^{a_1,x_1}_{B_1})$ and let $\hat{\kappa}^{a_2,x_2,y_2}_{B_2'G}$ be a non-signaling extension of $\mathcal{M}_{y_2}(\hat{\kappa}^{a_2,x_2}_{B_2})$, leading to the following classical–quantum states:

$$\theta_{X_1\bar{A}_1B_1'FY_1} := \sum_{x_1,a_1} p_{X_1|Y_1}(x_1|y_1)[x_1] \otimes [a_1] \otimes \hat{\theta}^{a_1,x_1,y_1}_{B_1'F} \otimes [y_1], \quad (5.100)$$

$$\kappa_{X_2\bar{A}_2B_2'GY_2} := \sum_{x_2,a_2} p_{X_2|Y_2}(x_2|y_2)[x_2] \otimes [a_2] \otimes \hat{\kappa}^{a_2,x_2,y_2}_{B_2'G} \otimes [y_2]. \quad (5.101)$$

Consider that

$$I(\bar{A}_1X_1\bar{A}_2X_2; B_1'B_2'|EY_1Y_2)_\omega$$
$$= I(\bar{A}_1X_1\bar{A}_2X_2; B_1'|EY_1Y_2)_\omega + I(\bar{A}_1X_1\bar{A}_2X_2; B_2'|EB_1'Y_1Y_2)_\omega \quad (5.102)$$
$$= I(\bar{A}_1X_1; B_1'|EY_1Y_2)_\omega + I(\bar{A}_2X_2; B_1'|EY_1Y_2\bar{A}_1X_1)_\omega$$
$$+ (\bar{A}_2X_2; B_2'|EB_1'Y_1Y_2)_\omega + I(\bar{A}_1X_1; B_2'|EB_1'Y_1Y_2\bar{A}_2X_2)_\omega \quad (5.103)$$
$$\geq I(\bar{A}_1X_1; B_1'|EY_1Y_2)_\omega + I(\bar{A}_2X_2; B_2'|EB_1'Y_1Y_2)_\omega \quad (5.104)$$
$$\geq \inf_{\text{ext. in (5.100)}} I(\bar{A}_1X_1; B_1'|FY_1)_\theta + \inf_{\text{ext. in (5.101)}} I(\bar{A}_2X_2; B_2'|GY_2)_\kappa. \quad (5.105)$$

The first two equalities follow from the chain rule for conditional mutual information. The first inequality follows by dropping two of the terms and from the fact that the conditional mutual information is non-negative. To see the last inequality, consider that the state $\sum_{a_2,x_2,y_2} \hat{\omega}^{a_1,x_1,a_2,x_2,y_1,y_2}_{B_1'E} \otimes [y_2]$ is a particular non-signaling extension of $\mathcal{L}_{y_1}(\hat{\theta}^{a_1,x_1}_{B_1})$ and the state $\sum_{a_1,x_1,y_1} \hat{\omega}^{a_1,x_1,a_2,x_2,y_1,y_2}_{B_1'B_2'E} \otimes [y_1]$ is a particular non-signaling extension of $\mathcal{M}_{y_2}(\hat{\kappa}^{a_2,x_2}_{B_2})$, such that an infimization over arbitrary respective non-signaling extensions $\hat{\theta}^{a_1,x_1,y_1}_{B_1'F}$ and $\hat{\kappa}^{a_2,x_2,y_2}_{B_2'G}$ can never lead to higher values of the conditional mutual informations. Since we have shown the inequality above for an arbitrary non-signaling extension $\hat{\omega}^{a_1,x_1,a_2,x_2,y_1,y_2}_{B_1'B_2'E}$, we can conclude that

$$\begin{aligned}&\inf_{\text{ext. in (5.99)}} I(\bar{A}_1X_1\bar{A}_2X_2;B_1'B_2'|EY_1Y_2)_\omega \\ &\quad\geq \inf_{\text{ext. in (5.100)}} I(\bar{A}_1X_1;B_1'|FY_1)_\theta + \inf_{\text{ext. in (5.101)}} I(\bar{A}_2X_2;B_2'|GY_2)_\kappa, \end{aligned} \tag{5.106}$$

which in turn implies that

$$\begin{aligned}&\sup_{\{p_{X_1|Y_1}p_{X_2|Y_2},\{\mathcal{L}_{y_1}\otimes\mathcal{M}_{y_2}\}_{y_1,y_2}\}} \inf_{\text{ext. in (5.99)}} I(\bar{A}_1X_1\bar{A}_2X_2;B_1'B_2'|EY_1Y_2)_\omega \\ &\quad\geq \inf_{\text{ext. in (5.100)}} I(\bar{A}_1X_1;B_1'|FY_1)_\theta + \inf_{\text{ext. in (5.101)}} I(\bar{A}_2X_2;B_2'|GY_2)_\kappa. \end{aligned} \tag{5.107}$$

The reduced 1W-LOCC operations $\{p_{X_1|Y_1},\{\mathcal{L}_{y_1}\}_{y_1}\}$ and $\{p_{X_2|Y_2},\{\mathcal{M}_{y_2}\}_{y_2}\}$ are arbitrary, and so we can conclude that

$$\begin{aligned}&\sup_{\{p_{X_1|Y_1}p_{X_2|Y_2},\{\mathcal{L}_{y_1}\otimes\mathcal{M}_{y_2}\}_{y_1,y_2}\}} \inf_{\text{ext. in (5.99)}} I(\bar{A}_1X_1\bar{A}_2X_2;B_1'B_2'|EY_1Y_2)_\omega \\ &\geq \sup_{\{p_{X_1|Y_1},\{\mathcal{L}_{y_1}\}_{y_1}\}} \inf_{\text{ext. in (5.100)}} I(\bar{A}_1X_1;B_1'|FY_1)_\theta \\ &\qquad + \sup_{\{p_{X_2|Y_2},\{\mathcal{M}_{y_2}\}_{y_2}\}} \inf_{\text{ext. in (5.101)}} I(\bar{A}_2X_2;B_2'|GY_2)_\kappa \qquad (5.108) \\ &= S(\bar{A}_1;B_1)_{\hat{\theta}} + S(\bar{A}_2;B_2)_{\hat{\kappa}}. \qquad (5.109)\end{aligned}$$

Finally, since the 1W-LOCC operation $\{p_{X_1|Y_1}p_{X_2|Y_2}, \{\mathcal{L}_{y_1} \otimes \mathcal{M}_{y_2}\}_{y_1,y_2}\}$ has a particular product form, we could never achieve a lower value of the quantity on the LHS by allowing for an arbitrary 1W-LOCC operation, implying the desired superadditivity:

$$S(\bar{A}_1\bar{A}_2; B_1B_2)_{\hat{\rho}} \geq S(\bar{A}_1; B_1)_{\hat{\theta}} + S(\bar{A}_2; B_2)_{\hat{\kappa}}. \tag{5.110}$$

This concludes the proof. ■

5.4. Properties of restricted instrinsic steerability

We prove that the restricted intrinsic steerability is a steering monotone with respect to restricted 1W-LOCC and that it is convex.

Theorem 52 *The restricted intrinsic steerability $S^R(\bar{A}; B)_{\hat{\rho}}$ is a convex steering monotone with respect to restricted 1W-LOCC. That is, it does not increase under restricted deterministic 1W-LOCC, it vanishes for assemblages having a local-hidden-state model, and it is convex.*

Proof. The proof follows from Proposition 53, 54, and 55. ■

Proposition 53 *Restricted intrinsic steerability vanishes for assemblages having an LHS model.*

Proof. To prove this, consider the following non-signaling, classical extension of an unsteerable assemblage:

$$\rho_{X\bar{A}BE} := \sum_{a,x} p_X(x) \left|x\middle\rangle\!\middle\langle x\right|_X \otimes p_{\bar{A}|X\Lambda}(a|x,\lambda) \left|a\middle\rangle\!\middle\langle a\right|_{\bar{A}} \otimes \hat{\rho}^{\lambda}_B \otimes p_\Lambda(\lambda) \left|\lambda\middle\rangle\!\middle\langle \lambda\right|_E . \tag{5.111}$$

Then,

$$I(X\bar{A}; B|E)_\rho = \sum_\lambda p_\Lambda(\lambda) I(X\bar{A}; B)_{\rho^\lambda}, \tag{5.112}$$

where

$$\rho^{\lambda}_{X\bar{A}B} = \sum_{a,x} p_X(x)\, |x\rangle\langle x|_X \otimes p_{\bar{A}|X\Lambda}(a|x,\lambda)\, |a\rangle\langle a|_{\bar{A}} \otimes \rho^{\lambda}_{B}, \tag{5.113}$$

and we have used the fact that the conditional mutual information can be written as a convex combination of mutual informations for a classical conditioning system. By inspection, we see that systems $X\bar{A}$ and B are independent when given the shared variable $\Lambda = \lambda$. By choosing system E to contain the shared random variable Λ, the result is that the systems form a Markov chain $X\bar{A} - E - B$, so that the conditional mutual information $I(X\bar{A}; B|E)_{\rho}$ is equal to zero. Since this argument holds for any probability distribution p_X, we conclude that $S^R(\bar{A}; B)_{\hat{\rho}} = 0$. ■

We now prove that restricted intrinsic steerability is a 1W-LOCC monotone.

Proposition 54 (Restricted 1W-LOCC monotone) *Let $\{\hat{\rho}^{a,x}_B\}_{a,x}$ be an assemblage, and let $\{p_{X|X_f}, p_{\bar{A}_f|\bar{A}XX_fZ}, \{\mathcal{K}_z\}_z\}$ denote a restricted 1W-LOCC operation that results in an assemblage $\{\hat{\sigma}^{a_f,x_f}_{B'}\}_{a_f,x_f}$, defined as*

$$\hat{\sigma}^{a_f,x_f}_{B'} := \sum_{a,x,z} p_{X|X_f}(x|x_f) p_{\bar{A}_f|\bar{A}XX_fZ}(a_f|a,x,x_f,z) \mathcal{K}_z(\hat{\rho}^{a,x}_B). \tag{5.114}$$

Then

$$S^R(\bar{A}; B)_{\hat{\rho}} \geq S^R(\bar{A}_f; B')_{\hat{\sigma}}. \tag{5.115}$$

Proof. Taking a distribution p_{X_f} over the black-box inputs of the final assemblage, we can embed the state of the final assemblage into the following classical–quantum state:

$$\sigma_{X_f\bar{A}_fB'} := \sum_{x_f,a_f} p_{X_f}(x_f)[x_f] \otimes [a_f] \otimes \hat{\sigma}^{a_f,x_f}_{B'}, \tag{5.116}$$

which is a marginal of the following state:

$$\sigma_{X_f X \bar{A}_f \bar{A} Z B'} := \sum_{x_f, a_f, a, x, z} p_{X_f}(x_f)[x_f] \otimes p_{X|X_f}(x|x_f)[x] \otimes p_{\bar{A}_f|\bar{A}XX_fZ}(a_f|a, x, x_f, z)[a_f] \otimes$$

$$[a] \otimes [z] \otimes \mathcal{K}_z(\hat{\rho}_B^{a,x}). \quad (5.117)$$

An *arbitrary* non-signaling extension of the state in (5.116) is as follows:

$$\sigma_{X_f \bar{A}_f B' E} := \sum_{x_f, a_f} p_{X_f}(x_f)[x_f] \otimes [a_f] \otimes \hat{\sigma}_{B'E}^{a_f, x_f}, \quad (5.118)$$

where

$$\mathrm{Tr}_E(\hat{\sigma}_{B'E}^{a_f, x_f}) = \hat{\sigma}_{B'}^{a_f, x_f}, \quad (5.119)$$

$$\sum_{a_f} \hat{\sigma}_{B'E}^{a_f, x_f} = \sum_{a_f} \hat{\sigma}_{B'E}^{a_f, x_f} \quad \forall x_f, x_f' \in \mathcal{X}_f. \quad (5.120)$$

A *particular* non-signaling extension of the state in (5.116) is as follows:

$$\omega_{X_f \bar{A}_f B' E Z} := \sum_{x_f, a_f} p_{X_f}(x_f)[x_f] \otimes$$

$$[a_f] \otimes \sum_{x_f, a_f, a, x, z} p_{X|X_f}(x|x_f) p_{\bar{A}_f|\bar{A}XX_fZ}(a_f|a, x, x_f, z) \mathcal{K}_z(\hat{\rho}_{BE}^{a,x}) \otimes [z], \quad (5.121)$$

where

$$\mathrm{Tr}_E(\hat{\rho}_{BE}^{a,x}) = \hat{\rho}_B^{a,x}, \quad (5.122)$$

$$\sum_a \hat{\rho}_{BE}^{a,x} = \sum_a \hat{\rho}_{BE}^{a,x'} \quad \forall x, x' \in \mathcal{X}. \quad (5.123)$$

The state $\omega_{X_f \bar{A}_f B' E}$ is a marginal of the following state:

$$\omega_{X_f X \bar{A}_f \bar{A} B' E Z} := \sum_{x_f, a_f, a, x, z} p_{X_f}(x_f)[x_f] p_{X|X_f}(x|x_f)[x] p_{\bar{A}_f|\bar{A}XX_fZ}(a_f|a, x, x_f, z)[a_f]$$

$$\otimes [a] \otimes \mathcal{K}_z(\hat{\rho}^{a,x}_{BE}) \otimes [z]. \quad (5.124)$$

Let $\rho_{X\bar{A}BE}$ be the following state:

$$\rho_{X\bar{A}BE} := \sum_{x_f,a,x} p_{X_f}(x_f)[x_f] \otimes p_{X|X_f}(x|x_f)[x] \otimes [a] \otimes \hat{\rho}^{a,x}_{BE}. \quad (5.125)$$

Consider that

$$\inf_{\text{ext. in (5.118)}} I(X_f\bar{A}_f; B'|E)_\sigma \leq I(X_f\bar{A}_f; B'|EZ)_\omega \quad (5.126)$$

$$\leq I(X_f\bar{A}_f X\bar{A}; B'|EZ)_\omega \quad (5.127)$$

$$= I(X\bar{A}; B'|EZ)_\omega + I(X_f; B'|EZX\bar{A})_\omega + I(\bar{A}_f; B'|EZX_fX\bar{A})_\omega \quad (5.128)$$

$$= I(X\bar{A}; B'|EZ)_\omega \quad (5.129)$$

$$\leq I(X\bar{A}; B'Z|E)_\omega \quad (5.130)$$

$$\leq I(X\bar{A}; B|E)_\rho. \quad (5.131)$$

The first inequality follows because the non-signaling extension in (5.121) is a particular kind of non-signaling extension. The second inequality follows from data processing. The first equality follows from the chain rule for conditional mutual information. The second equality follows from various Markov-chain structures when inspecting (5.124): X_f is independent of $B'E$ when given $ZX\bar{A}$, and $\bar{A}_f$ is independent of $B'E$ when given $ZX_fX\bar{A}$, so that $I(X_f; B'|EZX\bar{A})_\omega = I(\bar{A}_f; B'|EZX_fX\bar{A})_\omega = 0$. The third inequality follows by applying the chain rule for and non-negativity of conditional mutual information. The last inequality follows again from data processing. Since the inequality holds for all non-signaling extensions of the form in (5.125), we can conclude that

$$\inf_{\text{ext. in (5.118)}} I(X_f\bar{A}_f;B'|E)_\sigma \leq \inf_{\text{ext. in (5.125)}} I(X\bar{A};B|E)_\rho \tag{5.132}$$

$$\leq \sup_{p_X} \inf_{\text{ext. in (5.125)}} I(X\bar{A};B|E)_\rho. \tag{5.133}$$

Since the inequality above holds for an arbitrary choice of p_{X_f}, we can finally conclude that

$$\sup_{p_{X_f}} \inf_{\text{ext. in (5.118)}} I(X_f\bar{A}_f;B'|E)_\sigma \leq \sup_{p_X} \inf_{\text{ext. in (5.125)}} I(X\bar{A};B|E)_\rho, \tag{5.134}$$

which is equivalent to the statement of the proposition. ■

The proof of convexity of the restricted intrinsic steerability is along the same lines as that for intrinsic steerability, given already in the proof of Proposition 50. We summarize the result as the following proposition:

Proposition 55 (Convexity) *Let $\{\hat{\rho}_B^{a,x}\}_{a,x}$ and $\{\hat{\sigma}_B^{a,x}\}_{a,x}$ be assemblages, and let $\lambda \in [0,1]$. Let $\{\hat{\tau}_B^{a,x}\}_{a,x}$ be a mixture of the two assemblages, defined as*

$$\hat{\tau}_B^{a,x} := \lambda\hat{\rho}_B^{a,x} + (1-\lambda)\hat{\sigma}_B^{a,x}. \tag{5.135}$$

Then

$$S^R(\bar{A};B)_{\hat{\tau}} \leq \lambda S^R(\bar{A};B)_{\hat{\rho}} + (1-\lambda)S^R(\bar{A};B)_{\hat{\sigma}}. \tag{5.136}$$

Proposition 56 (Superadditivity and Additivity) *Let $\{\hat{\rho}_{B_1B_2}^{a_1,a_2,x_1,x_2}\}_{a_1,a_2,x_1,x_2}$ be an assemblage for which the following additional no-signaling constraints hold*

$$\sum_{a_2} \hat{\rho}_{B_1B_2}^{a_1,a_2,x_1,x_2} = \sum_{a_2} \hat{\rho}_{B_1B_2}^{a_1,a_2,x_1,x_2'} := \hat{\theta}_{B_1B_2}^{a_1,x_1} \qquad \forall x_2, x_2', \tag{5.137}$$

$$\sum_{a_1} \hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2} = \sum_{a_1} \hat{\rho}^{a_1,a_2,x_1',x_2}_{B_1B_2} := \hat{\kappa}^{a_2,x_2}_{B_1B_2} \qquad \forall x_1, x_1', \tag{5.138}$$

Let $\{\mathrm{Tr}_{B_2}(\hat{\theta}^{a_1,x_1}_{B_1B_2})\}_{a_1,x_1}$ *and* $\{\mathrm{Tr}_{B_1}(\hat{\kappa}^{a_2,x_2}_{B_1B_2})\}_{a_2,x_2}$ *be reduced assemblages arising from the joint assemblage* $\{\hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2}\}_{a_1,a_2,x_1,x_2}$*. Then the restricted intrinsic steerability is superadditive in the following sense:*

$$S^R(\bar{A}_1\bar{A}_2; B_1B_2)_{\hat{\rho}} \geq S^R(\bar{A}_1; B_1)_{\hat{\theta}} + S^R(\bar{A}_2; B_2)_{\hat{\kappa}}. \tag{5.139}$$

If the assemblage $\{\hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2}\}_{a_1,a_2,x_1,x_2}$ *has a tensor-product form, so that* $\hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2} = \hat{\theta}^{a_1,x_1}_{B_1} \otimes \hat{\kappa}^{a_2,x_2}_{B_2}$ *for assemblages* $\{\hat{\theta}^{a_1,x_1}_{B_1}\}_{a_1,x_1}$ *and* $\{\hat{\kappa}^{a_2,x_2}_{B_2}\}_{a_2,x_2}$*, then the restricted intrinsic steerability is additive:*

$$S^R(\bar{A}_1\bar{A}_2; B_1B_2)_{\hat{\rho}} = S^R(\bar{A}_1; B_1)_{\hat{\theta}} + S^R(\bar{A}_2; B_2)_{\hat{\kappa}}. \tag{5.140}$$

Proof. The superadditivity of restricted intrinsic steerability is similar to the proof above for intrinsic steerability. Thus, to prove the additivity of intrinsic steerability with respect to product assemblages, it is sufficient to prove the following subadditivity inequality:

$$S^R(\bar{A}_1\bar{A}_2; B_1B_2)_{\hat{\rho}} \leq S^R(\bar{A}_1; B_1)_{\hat{\theta}} + S^R(\bar{A}_2; B_2)_{\hat{\kappa}}. \tag{5.141}$$

Our proof of the above inequality has some similarities to the proof of the additivity of the squashed entanglement of a channel [102] (there are, however, some key differences). Let $\hat{\theta}^{a_1,x_1}_{B_1E_1}$ and $\hat{\kappa}^{a_2,x_2}_{B_2E_2}$ be non-signaling extensions of $\hat{\theta}^{a_1,x_1}_{B_1}$ and $\hat{\kappa}^{a_2,x_2}_{B_2}$, respectively, and suppose that $|\hat{\theta}^{a_1,x_1}\rangle_{B_1E_1F_1}$ and $|\hat{\kappa}^{a_2,x_2}\rangle_{B_2E_2F_2}$ purify $\hat{\theta}^{a_1,x_1}_{B_1E_1}$ and $\hat{\kappa}^{a_2,x_2}_{B_2E_2}$, respectively. Consider the following states:

$$\rho_{X_1X_2\bar{A}_1\bar{A}_2B_1B_2E} := \sum_{x_1,x_2,a_1,a_2} p_{X_1X_2}(x_1,x_2)[x_1]\otimes[x_2]\otimes[a_1]\otimes[a_2]\otimes\hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2E}, \tag{5.142}$$

$$\omega_{X_1X_2\bar{A}_1\bar{A}_2B_1B_2E_1E_2F_1F_2} := \sum_{x_1,x_2,a_1,a_2} p_{X_1X_2}(x_1,x_2)[x_1]\otimes[x_2]\otimes[a_1]\otimes[a_2]\otimes\hat{\theta}^{a_1,x_1}_{B_1E_1F_1} \otimes\hat{\kappa}^{a_2,x_2}_{B_2E_2F_2}, \tag{5.143}$$

where $p_{X_1X_2}(x_1,x_2)$ is some probability distribution and $\operatorname{Tr}_E(\hat{\rho}^{a_1,a_2,x_1,x_2}_{B_1B_2E}) = \hat{\theta}^{a_1,x_1}_{B_1}\otimes\hat{\kappa}^{a_2,x_2}_{B_2}$. Consider that

$$\begin{aligned}
&\inf_{\rho_{\bar{A}_1\bar{A}_2X_1X_2B_1B_2E}} I(\bar{A}_1\bar{A}_2X_1X_2;B_1B_2|E)_\rho \\
&\quad\le I(\bar{A}_1\bar{A}_2X_1X_2;B_1B_2|E_1E_2)_\omega && (5.144)\\
&\quad= H(B_1B_2|E_1E_2)_\omega - H(B_1B_2|E_1E_2\bar{A}_1X_1\bar{A}_2X_2)_\omega && (5.145)\\
&\quad= H(B_1B_2|E_1E_2)_\omega + H(B_1B_2|F_1F_2\bar{A}_1X_1\bar{A}_2X_2)_\omega && (5.146)\\
&\quad\le H(B_1|E_1)_\omega + H(B_2|E_2)_\omega + H(B_1|F_1\bar{A}_1X_1)_\omega + H(B_2|F_2\bar{A}_2X_2)_\omega && (5.147)\\
&\quad= H(B_1|E_1)_\omega + H(B_2|E_2)_\omega - H(B_1|E_1\bar{A}_1X_1)_\omega - H(B_2|E_2\bar{A}_2X_2)_\omega && (5.148)\\
&\quad= I(X_1\bar{A}_1;B_1|E_1)_\omega + I(X_2\bar{A}_2;B_2|E_2)_\omega. && (5.149)
\end{aligned}$$

The first inequality follows because $\omega_{X_1X_2\bar{A}_1\bar{A}_2B_1B_2E_1E_2}$ is a particular non-signaling extension whereas $\rho_{X_1X_2\bar{A}_1\bar{A}_2B_1B_2E}$ is an arbitrary non-signaling extension. The first equality follows from the chain rule for conditional mutual information. Conditioned on $\bar{A}_1\bar{A}_2X_1X_2$, the state on $B_1E_1B_2E_2F_1F_2$ is pure, and so the second equality follows from the duality of conditional entropy. The first inequality is a consequence of the strong subadditivity of quantum entropy [17]. The third equality follows again from the duality of conditional entropy as well as the no-signaling condition. To see this for the entropy $H(B_1|F_1\bar{A}_1X_1)_\omega$, consider that this entropy is evaluated with respect to the following reduced state:

$$\operatorname{Tr}_{X_2\bar{A}_2B_2E_2F_2}\left(\sum_{x_1,x_2,a_1,a_2} p_{X_1X_2}(x_1,x_2)[x_1]\otimes[x_2]\otimes[a_1]\otimes[a_2]\otimes\hat{\theta}^{a_1,x_1}_{B_1E_1F_1}\otimes\hat{\kappa}^{a_2,x_2}_{B_2E_2F_2}\right)$$

$$= \sum_{x_1,x_2,a_1,a_2} p_{X_1X_2}(x_1,x_2)[x_1]\otimes[a_1]\otimes\hat{\theta}^{a_1,x_1}_{B_1E_1F_1}\otimes\operatorname{Tr}_{B_2E_2F_2}\{\hat{\kappa}^{a_2,x_2}_{B_2E_2F_2}\} \tag{5.150}$$

$$= \sum_{x_1,a_1} p_{X_1}(x_1)[x_1]\otimes[a_1]\otimes\hat{\theta}^{a_1,x_1}_{B_1E_1F_1}\otimes\operatorname{Tr}_{B_2}\left(\sum_{x_2} p_{X_2|X_1}(x_2|x_1)\sum_{a_2}\hat{\kappa}^{a_2,x_2}_{B_2}\right) \tag{5.151}$$

$$= \sum_{x_1,a_1} p_{X_1}(x_1)[x_1]\otimes[a_1]\otimes\hat{\theta}^{a_1,x_1}_{B_1E_1F_1}\otimes\operatorname{Tr}_{B_2}\left(\sum_{x_2} p_{X_2|X_1}(x_2|x_1)\kappa_{B_2}\right) \tag{5.152}$$

$$= \sum_{x_1,a_1} p_{X_1}(x_1)[x_1]\otimes[a_1]\otimes\hat{\theta}^{a_1,x_1}_{B_1E_1F_1}\otimes\operatorname{Tr}_{B_2}(\kappa_{B_2}) \tag{5.153}$$

$$= \sum_{x_1,a_1} p_{X_1}(x_1)[x_1]\otimes[a_1]\otimes\hat{\theta}^{a_1,x_1}_{B_1E_1F_1}. \tag{5.154}$$

In the above, the third equality is the critical one in which we have used the no-signaling constraint for the assemblage $\{\hat{\kappa}^{a_2,x_2}_{B_2}\}_{a_2,x_2}$, allowing for the effective removal of correlation between X_1 and X_2. Thus, the above analysis allows for seeing that the remaining state on $B_1E_1F_1$ conditioned on $\bar{A}_1$ and X_1 is independent of any of the second system. For the last equality, we employ the definition of conditional mutual information. Since the above development holds for all non-signaling extensions of the form in (5.143), we find that

$$\inf_{\rho_{\bar{A}_1\bar{A}_2X_1X_2B_1B_2E}} I(\bar{A}_1\bar{A}_2X_1X_2;B_1B_2|E_1)_\rho$$

$$\leq \inf_{\omega_{\bar{A}_1X_1B_1E_1}} I(\bar{A}_1X_1;B_1|E_1)_\omega + \inf_{\omega_{\bar{A}_2X_2B_2E_2}} I(\bar{A}_2X_2;B_2|E_2)_\omega \tag{5.155}$$

$$\leq \sup_{p_{X_1}}\inf_{\omega_{\bar{A}_1X_1B_1E_1}} I(\bar{A}_1X_1;B_1|E_1)_\omega + \sup_{p_{X_2}}\inf_{\omega_{\bar{A}_2X_2B_2E_2}} I(\bar{A}_2X_2;B_2|E_2)_\omega. \tag{5.156}$$

Since the above inequality holds for an arbitrary probability distribution $p_{X_1 X_2}$, we conclude that

$$\sup_{p_{X_1 X_2}} \inf_{\rho_{\bar{A}_1 \bar{A}_2 X_1 X_2 B_1 B_2 E}} I(\bar{A}_1 \bar{A}_2 X_1 X_2; B_1 B_2 | E)_\rho$$

$$\leq \sup_{p_{X_1}} \inf_{\omega_{\bar{A}_1 X_1 B_1 E_1}} I(\bar{A}_1 X_1; B_1 | E_1)_\omega + \sup_{p_{X_2}} \inf_{\omega_{\bar{A}_2 X_2 B_2 E_2}} I(\bar{A}_2 X_2; B_2 | E_2)_\omega, \quad (5.157)$$

which is equivalent to (5.141). ■

Monogamy of entanglement is a fundamental property of entanglement quantum states and is a statement regarding the extendibility of a quantum state. As an example, the maximally entangled state Φ_{AB} is not extendible and hence the systems A and B cannot be entangled with any other extension C. Monogamy of entanglement is reflected by entanglement measures such as squashed entanglement in the form of the following inequality:

$$E(A; BC)_\rho \geq E(A; B)_\rho + E(A; C)_\rho, \quad (5.158)$$

where E is the entanglement measure.

Monogamy of steering has also been explored in [103, 104]. We prove here that the restricted intrinsic steerability is monogamous in the following sense: for a tripartite state ρ_{ABC}, Alice and Charlie perform measurements on their systems and steer Bob's system. We see that their ability to steer Bob's system is restricted.

Proposition 57 (Monogamy) *Let $\{\hat{\rho}_B^{a,c,x_1,x_2}\}$ be an assemblage with classical inputs x_1 and x_2 for Alice and Charlie, respectively, and classical outputs a and c for Alice and Charlie, respectively, and obeying the following additional no-signaling con-*

straints:

$$\sum_c \hat{\rho}_B^{a,c,x_1,x_2} = \sum_c \hat{\rho}_B^{a,c,x_1,x_2'} := \hat{\theta}_B^{a,x_1} \qquad \forall x_2, x_2', \tag{5.159}$$

$$\sum_a \hat{\rho}_B^{a,c,x_1,x_2} = \sum_a \hat{\rho}_B^{a,c,x_1',x_2} := \hat{\kappa}_B^{c,x_2} \qquad \forall x_1, x_1', \tag{5.160}$$

such that the reduced assemblages are $\{\hat{\theta}_B^{a,x_1}\}_{a,x_1}$ *and* $\{\hat{\kappa}_B^{c,x_2}\}_{c,x_2}$*. Then the following monogamy inequality holds*

$$S^R(\bar{A}\bar{C};B)_{\hat{\rho}} \geq S^R(\bar{A};B)_{\hat{\theta}} + S^R(\bar{C};B)_{\hat{\kappa}}. \tag{5.161}$$

Proof. This proof follows from an application of the chain rule for conditional mutual information, much like the proof of monogamy for the squashed entanglement [105]. First, consider the following classical–quantum state:

$$\rho_{X_1X_2\bar{A}\bar{C}BE} := \sum_{x_1,x_2,a,c} p_{X_1}(x_1)p_{X_2}(x_2)[x_1] \otimes [x_2] \otimes [a] \otimes [c] \otimes \hat{\rho}_{BE}^{a,c,x_1,x_2}, \tag{5.162}$$

where $\hat{\rho}_{BE}^{a,c,x_1,x_2}$ is a non-signaling extension of $\hat{\rho}_B^{a,c,x_1,x_2}$. Let

$$\theta_{X_1\bar{A}BF} := \sum_{x_1,a} p_{X_1}(x_1)[x_1] \otimes [a] \otimes \hat{\theta}_{BF}^{a,x_1}, \tag{5.163}$$

$$\kappa_{X_2\bar{C}BG} := \sum_{x_2,a} p_{X_2}(x_2)[x_2] \otimes [c] \otimes \hat{\kappa}_{BG}^{c,x_2}, \tag{5.164}$$

where $\hat{\theta}_{BF}^{a,x_1}$ is a non-signaling extension of $\hat{\theta}_B^{a,x_1}$ and $\hat{\kappa}_{BG}^{c,x_2}$ is a non-signaling extension of $\hat{\kappa}_B^{c,x_2}$. Then we have from the chain rule for conditional mutual information that

$$I(X_1X_2\bar{A}\bar{C};B|E)_\rho$$

$$= I(X_1\bar{A};B|E)_\rho + I(X_2\bar{C};B|E\bar{A}X_1)_\rho \tag{5.165}$$

$$\geq \inf_{\theta_{X_1\bar{A}BF}} I(X_1\bar{A};B|E)_\theta + \inf_{\kappa_{X_2\bar{C}BG}} I(X_2\bar{C};B|G)_\kappa. \tag{5.166}$$

Since the above inequality holds for all non-signaling extensions $\rho_{X_1X_2\bar{A}\bar{C}BE}$, we conclude that

$$\inf_{\rho_{X_1X_2\bar{A}\bar{C}BE}} I(X_1X_2\bar{A}\bar{C};B|E)_\rho \ \geq \ \inf_{\theta_{X_1\bar{A}BF}} I(X_1\bar{A};B|E)_\theta \ + \ \inf_{\kappa_{X_2\bar{C}BG}} I(X_2\bar{C};B|G)_\kappa. \tag{5.167}$$

Optimizing the left-hand side with respect to product distributions, we find that

$$\sup_{p_{X_1},p_{X_2}} \inf_{\rho_{X_1X_2\bar{A}\bar{C}BE}} I(X_1X_2\bar{A}\bar{C};B|E)_\rho \geq \inf_{\theta_{X_1\bar{A}BF}} I(X_1\bar{A};B|E)_\theta + \inf_{\kappa_{X_2\bar{C}BG}} I(X_2\bar{C};B|G)_\kappa. \tag{5.168}$$

The development holds for any choice of distributions p_{X_1} and p_{X_2}, and so we conclude that

$$\sup_{p_{X_1},p_{X_2}} \inf_{\rho_{X_1X_2\bar{A}\bar{C}BE}} I(X_1X_2\bar{A}\bar{C};B|E)_\rho$$

$$\geq \sup_{p_{X_1}} \inf_{\theta_{X_1\bar{A}BF}} I(X_1\bar{A};B|E)_\theta + \sup_{p_{X_2}} \inf_{\kappa_{X_2\bar{C}BG}} I(X_2\bar{C};B|G)_\kappa \tag{5.169}$$

$$= S^R(\bar{A};B)_{\hat{\theta}} + S^R(\bar{C};B)_{\hat{\kappa}}. \tag{5.170}$$

Finally optimizing the left-hand side with respect to all input distributions $p_{X_1X_2}$, we conclude (5.161). ■

Here we establish the faithfulness of restricted intrinsic steerability.

Theorem 58 (Faithfulness of restricted intrinsic steerability) *For every assem-*

blage $\hat{\rho}_B^{a;x}$, the restricted intrinsic steerability $S^R(A;B)_{\hat{\rho}} = 0$, if and only if it is an LHS assemblage. Quantitatively, if $S^R(\bar{A};B)_{\hat{\rho}} \leq \varepsilon$, where $0 < \varepsilon^{\frac{1}{16}}|\mathcal{X}|^{\frac{1}{2}} < 1$, there exists an LHS assemblage $\sigma_{\bar{A}XB}$ such that

$$\sup_{p_X(x)} \|\rho_{\bar{A}XB} - \sigma_{\bar{A}XB}\|_1 \leq |\mathcal{X}| \left(\varepsilon^{1/4} + \frac{\varepsilon^{1/16}|\mathcal{X}|^{1/2}}{1-\varepsilon^{1/16}|\mathcal{X}|^{1/2}} + 4|\mathcal{X}|e^{-\frac{\varepsilon^{-1/4}}{3}} \right). \tag{5.171}$$

Proof. The forward direction ("if") follows from Proposition 53. We now give a proof for the reverse direction ("only if") of the theorem.

Let us first construct a proof strategy for a uniform probability distribution $p_X(x) = \frac{1}{|\mathcal{X}|}$, and then we generalize it to a proof for an arbitrary distribution $p_X(x)$. This proof shares some ideas from the proof for faithfulness of squashed entanglement [97].

Invoking Theorem 5.1 of [106], we know that there exists a recovery channel $\mathcal{R}_{XE\to\bar{A}XE}$ such that

$$\|\rho_{\bar{A}XBE} - \mathcal{R}_{XE\to\bar{A}XE}(\rho_{BE}\otimes\rho_X)\|_1 \leq \sqrt{I(\bar{A};B|EX)_\rho \ln 2} =: t, \tag{5.172}$$

$$\|\rho_{\bar{A}XBE} - \mathcal{R}_{X_2E\to\bar{A}_2X_2E}\circ \operatorname{Tr}_{\bar{A}_1X_1}(\rho_{\bar{A}_1X_1BE}\otimes\rho_{X_2})\|_1 \leq t, \tag{5.173}$$

where systems $\bar{A}_1$ and $\bar{A}_2$ are isomorphic to system $\bar{A}$, and systems X_1 and X_2 are isomorphic to X. In the above, we have invoked the no-signaling condition $I(X;BE)_\rho = 0$, which implies that ρ_{BE} and ρ_X are product as written. Now, let us apply this recovery channel again. We then have that

$$\begin{aligned}\|\mathcal{R}_{X_3E\to\bar{A}_3X_3E}\circ\operatorname{Tr}_{X_2\bar{A}_2}(\rho_{\bar{A}_2X_2BE}\otimes\rho_{X_3}) - \\ \bigcirc_{i=2}^{3}\mathcal{R}_{X_iE\to\bar{A}_iX_iE}\circ\operatorname{Tr}_{A_{i-1}X_{i-1}}(\rho_{\bar{A}_1X_1BE}\otimes\rho_{X_2}\otimes\rho_{X_3})\big\|_1 \leq t.\end{aligned} \tag{5.174}$$

which follows from the monotonicity of trace distance with respect to $\mathcal{R}_{X_3E\to\bar{A}_3X_3E}\circ\operatorname{Tr}_{X_2\bar{A}_2}$. Then, combining the above equation with (5.172) via the triangle inequality, we obtain

$$\left\|\rho_{\bar{A}XBE}-\bigcirc_{i=2}^{3}\mathcal{R}_{X_iE\to\bar{A}_iX_iE}\circ\operatorname{Tr}_{A_{i-1}X_{i-1}}(\rho_{\bar{A}_1X_1BE}\otimes\rho_{X_2}\otimes\rho_{X_3})\right\|_1\leq 2t. \tag{5.175}$$

For $j\in\{4,\ldots,n\}$, again apply the channels $\mathcal{R}_{XE\to\bar{A}_jX_jE}\circ\operatorname{Tr}_{\bar{A}_{j-1}X_{j-1}}$, along with the monotonicity of trace norm under quantum channels, combining the equations via the triangle inequality, to obtain the following inequality:

$$\left\|\rho_{\bar{A}XB}-\operatorname{Tr}_E\{\bigcirc_{i=2}^{j}\mathcal{R}_{X_iE\to\bar{A}_iX_iE}\circ\operatorname{Tr}_{A_{i-1}X_{i-1}}\left(\rho_{\bar{A}_1X_1BE}\otimes\rho_X^{\otimes j}\right)\}\right\|_1\leq nt. \tag{5.176}$$

The recovery channel $\mathcal{R}_{X_iE\to\bar{A}_iX_iE}$ can be taken as [107]

$$\begin{aligned}
\mathcal{R}_{XE\to\bar{A}XE}\left(\cdot\right)&=\rho_{\bar{A}XE}^{\frac{1}{2}+i\omega}\rho_{XE}^{-\frac{1}{2}-i\omega}\left(\cdot\right)\rho_{XE}^{-\frac{1}{2}+i\omega}\rho_{\bar{A}XE}^{\frac{1}{2}-i\omega}, &(5.177)\\
&=\sum_x|x\rangle\langle x|_X\otimes(\rho_{\bar{A}E}^x)^{\frac{1}{2}+i\omega}\rho_E^{-\frac{1}{2}+i\omega}\left(\cdot\right)\rho_E^{-\frac{1}{2}+i\omega}(\rho_{\bar{A}E}^x)^{\frac{1}{2}-i\omega}, &(5.178)
\end{aligned}$$

for some $\omega\in\mathbb{R}$. Let $\sigma_{\bar{A}^nX^nBE}$ denote the following state:

$$\begin{aligned}
\sigma_{\bar{A}^nX^nBE}&=\left(\mathcal{R}_{X_nE\to\bar{A}_nX_nE}\circ\ldots\circ\mathcal{R}_{X_1E\to\bar{A}_1X_1E}\right)\left(\sigma_{BE}\otimes\sigma_X^{\otimes n}\right) &(5.179)\\
&=\sum_{a^n,x^n}p_{X^n}(x^n)q_{\bar{A}^n|X^n}(a^n|x^n)\,|x^n\rangle\langle x^n|_{X^n}\otimes|a^n\rangle\langle a^n|_{\bar{A}^n}\otimes\sigma_{BE}^{a^n,x^n}. &(5.180)\\
\sigma_{\bar{A}^nX^nB}&=\operatorname{Tr}_E(\sigma_{\bar{A}^nX^nBE}) &(5.181)\\
&=\sum_{a^n,x^n}p_{X^n}(x^n)q_{\bar{A}^n|X^n}(a^n|x^n)\,|x^n\rangle\langle x^n|_{X^n}\otimes|a^n\rangle\langle a^n|_{\bar{A}^n}\otimes\sigma_B^{a^n,x^n}. &(5.182)\\
\sigma_{\bar{A}_iX_iB}&=\operatorname{Tr}_{A^{[n]\setminus\{i\}}X^{[n]\setminus\{i\}}}(\sigma_{\bar{A}^nX^nB}) &(5.183)\\
&=\sum_{a^n,x^n}p_{X^n}(x^n)q_{\bar{A}^n|X^n}(a^n|x^n)\,|x_i\rangle\langle x_i|_{X_i}\otimes|a_i\rangle\langle a_i|_{\bar{A}_i}\otimes\sigma_B^{a^n,x^n}, &(5.184)
\end{aligned}$$

where $A^{[n]\backslash\{i\}} = A_1A_2\ldots A_{i-1}A_{i+1}\ldots A_n$ and similarly $X^{[n]\backslash\{i\}} = X_1X_2\ldots X_{i-1}X_{i+1}$ $\ldots X_n$. Furthermore, $q_{\bar{A}^n|X^n}(a^n|x^n)$ is a probability distribution for a^n given x^n after the application of the recovery channels $\mathcal{R}_{X_iE\to\bar{A}_iX_iE}$. From (5.176), we obtain for all $i \in \{1, 2, \ldots, n\}$ that

$$\left\|\rho_{\bar{A}XB} - \sigma_{\bar{A}_iX_iB}\right\|_1 \leq nt. \tag{5.185}$$

The application of the recovery channels generates the data $(x_1, a_1), (x_2, a_2), \ldots,$ (x_n, a_n). The x_i correspond to the measurement choices, and the a_i correspond to the measurement outcomes. This data is called the “cheat sheet” and acts like a hidden-variable λ. The formulation of the cheat sheet is similar to the construction of a local hidden-variable model in [108].

We now devise an algorithm to generate $\tilde{a}$ from $\tilde{x}$ by using the cheat sheet. The generated state $\sigma_{\tilde{A}\tilde{X}B}$ is a local hidden state, with the cheat sheet as the hidden variable. We then prove that $\sigma_{\tilde{A}\tilde{X}B}$ is close to the original state $\rho_{\bar{A}XB}$.

Alice receives $\tilde{x}$. She searches for all the values of i for which $x_i = \tilde{x}$, and generates i uniformly at random

$$p_{I|\tilde{X}X^n}(i|\tilde{x}x^n) = \frac{1}{N(\tilde{x}|x^n)}\delta_{x_i\tilde{x}}, \tag{5.186}$$

where $\delta_{x_i\tilde{x}}$ is the Kronecker delta function and where $N(\tilde{x}|x^n)$ is the number of times that the letter $\tilde{x}$ appears in the sequence x^n. Then, she outputs $\tilde{a}$ with probability

$$p_{\tilde{A}|A^nI}(\tilde{a}|a^ni) = \delta_{\tilde{a},a_i}. \tag{5.187}$$

Therefore,

$$p_{\tilde{A}|\tilde{X}X^nA^n}(\tilde{a}|\tilde{x}x^na^n) = \sum_{i=1}^{n} p_{\tilde{A}|A^nIX^n\tilde{X}}(\tilde{a}|a^nix^n\tilde{x})p_{I|\tilde{X}X^nA^n}(i|\tilde{x}x^na^n) \tag{5.188}$$

$$= \sum_{i=1}^{n} p_{\tilde{A}|A^nI}(\tilde{a}|a^n i) p_{I|\tilde{X}X^n}(i|\tilde{x}x^n). \tag{5.189}$$

$$= \sum_{i=1}^{n} \frac{1}{N(\tilde{x}|x^n)} \delta_{\tilde{x}x_i} \delta_{\tilde{a}a_i}. \tag{5.190}$$

If $\tilde{x}$ does not belong to the sequence x^n, then she generates $\tilde{a}$ randomly. This sequence of actions can be expressed in terms of the following conditional probability distribution:

$$p_{\tilde{A}|\tilde{X}X^nA^n}(\tilde{a}|\tilde{x}, x^n, a^n) := \begin{cases} \frac{1}{|A|}, & \text{if } N(\tilde{x}|x^n) = 0 \\ \sum_{i=1}^{n} \frac{1}{N(\tilde{x}|x^n)} \delta_{\tilde{x},x_i} \delta_{a_i,\tilde{a}} & \text{else.} \end{cases} \tag{5.191}$$

It is easy to check that $\sum_{\tilde{a}} p_{\tilde{A}|\tilde{X}X^nA^n}(\tilde{a}|\tilde{x}, x^n, a^n) = 1$.

We now use the notion of robust typicality [109] for the analysis.

Definition 59 (Robust typicality [109]) *Let x^n be a sequence of elements drawn from a finite alphabet $\mathcal{X}$, and let $p(x)$ be a probability distribution on $\mathcal{X}$. Let $N(x|x^n)$ be the empirical distribution of x^n. Then the δ-robustly typical set $T_\delta^{X^n}$ for $\delta > 0$ is defined as*

$$T_\delta^{X^n} := \left\{ \forall x \in \mathcal{X}, \left| \frac{1}{n} N(x|x^n) - p_X(x) \right| \leq \delta p(x) \right\}. \tag{5.192}$$

The following result holds for $0 < \delta < 1$:

Property 60 *The probability of a sequence x^n to be in the robustly typical set is bounded from below as*

$$\Pr\left\{ X^n \in T_\delta^{X^n} \right\} \geq 1 - 2|\mathcal{X}| \exp^{-\frac{n\delta^2 \mu_X}{3}}, \tag{5.193}$$

where

$$\mu_X := \min_{x \in \mathcal{X}, p_X(x) > 0} p_X(x). \tag{5.194}$$

The state generated after the application of the algorithm in (5.191) is as follows:

$$\sigma_{\tilde{A}\tilde{X}B} = \sum_{\tilde{x},\tilde{a}} p_{\tilde{X}}(\tilde{x}) \, |\tilde{x}\rangle\langle\tilde{x}|_{\tilde{X}} \otimes \sum_{x^n, a^n} p_{\tilde{A}|\tilde{X}X^n A^n}(\tilde{a}|\tilde{x}, x^n, a^n) p_{X^n}(x^n) q_{\bar{A}^n|X^n}(a^n|x^n) \\ |\tilde{a}\rangle\langle\tilde{a}|_{\tilde{A}} \otimes \sigma_B^{a^n, x^n}. \tag{5.195}$$

Then, define the following sets:

- $S_1(\tilde{x})$: set of sequences x^n such that $\tilde{x} \in x^n$ and $x^n \in T_\delta^{X^n}$,
- $S_2(\tilde{x})$: set of sequences x^n such that $\tilde{x} \notin x^n$ and $x^n \in T_\delta^{X^n}$,
- S_3: set of sequences x^n such that $x^n \notin T_\delta^{X^n}$.

So we can write the state $\sigma_{\tilde{A}\tilde{X}B}$ as

$$\begin{aligned} \sigma_{\tilde{A}\tilde{X}B} =& \sum_{\tilde{x},\tilde{a}} p_{\tilde{X}}(\tilde{x}) \, |\tilde{x}\rangle\langle\tilde{x}|_{\tilde{X}} \otimes \Bigg(\sum_{x^n \in S_1(\tilde{x}), a^n} p(\tilde{a}|\tilde{x}, x^n, a^n) \, |\tilde{a}\rangle\langle\tilde{a}| \otimes q(a^n, x^n) \sigma_B^{a^n, x^n} + \\ & \sum_{x^n \in S_2(\tilde{x}), a^n} p(\tilde{a}|\tilde{x}, x^n, a^n) \, |\tilde{a}\rangle\langle\tilde{a}| \otimes q(a^n, x^n) \sigma_B^{a^n, x^n} + \\ & \sum_{x^n \in S_3, a^n} p(\tilde{a}|\tilde{x}, x^n, a^n) \, |\tilde{a}\rangle\langle\tilde{a}| \otimes q(a^n, x^n) \sigma_B^{a^n, x^n} \Bigg), && (5.196) \\ \sigma_{\tilde{A}\tilde{X}B} =& \, \sigma^{(1)}_{\tilde{A}\tilde{X}B} + \sigma^{(2)}_{\tilde{A}\tilde{X}B} + \sigma^{(3)}_{\tilde{A}\tilde{X}B}. && (5.197) \end{aligned}$$

From the triangle inequality, we obtain the following:

$$\left\|\rho_{\bar{A}\bar{X}B} - \sigma_{\tilde{A}\tilde{X}B}\right\|_1 \leq \left\|\rho_{\bar{A}\bar{X}B} - \sigma^{(1)}_{\tilde{A}\tilde{X}B}\right\|_1 + \left\|\sigma^{(2)}_{\tilde{A}\tilde{X}B}\right\|_1 + \left\|\sigma^{(3)}_{\tilde{A}\tilde{X}B}\right\|_1, \tag{5.198}$$

where

$$\begin{aligned}
\left\| \rho_{\bar{A}XB} - \sigma^{(1)}_{\tilde{A}\tilde{X}B} \right\|_1 &\leq \left\| \rho_{\bar{A}\bar{X}B} - \frac{1}{n}\sum_{i=1}^{n} \sigma_{\bar{A}_i X_i B} \right\|_1 + \left\| \frac{1}{n}\sum_{i=1}^{n} \sigma_{\bar{A}_i X_i B} - \sigma^{(1)}_{\tilde{A}\tilde{X}B} \right\|_1 && (5.199)\\
&\leq nt + \left\| \frac{1}{n}\sum_{i=1}^{n} \sigma_{\bar{A}_i X_i B} - \sigma^{(1)}_{\tilde{A}\tilde{X}B} \right\|_1 . && (5.200)
\end{aligned}$$

Let us analyze each term individually, beginning with

$$\begin{aligned}
\left\| \sigma^{(3)}_{\tilde{A}\tilde{X}B} \right\|_1 &= \left\| \sum_{\tilde{x},\tilde{a}} p_{\tilde{X}}(\tilde{x})\, |\tilde{x}\rangle\langle\tilde{x}|_{\tilde{X}} \otimes \sum_{x^n \in S_3, a^n} p(\tilde{a}|\tilde{x}, x^n, a^n)\, |\tilde{a}\rangle\langle\tilde{a}| \otimes q(a^n, x^n)\sigma_B^{a^n,x^n} \right\|_1 && (5.201)\\
&\leq \sum_{\tilde{x},\tilde{a}} p(\tilde{x}) \sum_{x^n \in S_3, a^n} p(x^n) q(a^n|x^n) p(\tilde{a}|\tilde{x}, x^n, a^n) \left\| |\tilde{x}\rangle\langle\tilde{x}| \otimes |\tilde{a}\rangle\langle\tilde{a}| \otimes \sigma_B^{a^n,x^n} \right\|_1 && (5.202)\\
&= \sum_{\tilde{x}} p(\tilde{x}) \sum_{x^n \in S_3} p(x^n) \sum_{a^n} q(a^n|x^n) \sum_{\tilde{a}} p(\tilde{a}|\tilde{x}, x^n, a^n) \leq \varepsilon_1, && (5.203)
\end{aligned}$$

where $\varepsilon_1 = 2|\mathcal{X}| \exp^{-\frac{n\delta^2 \mu_X}{3}}$. The first inequality follows from convexity of trace distance, and the second inequality follows from the definition of S_3 and (5.193).

Let us now consider $S_2(\tilde{x})$, that is, the set of sequences x^n such that $\tilde{x} \notin x^n$ and $x^n \in T_\delta^{X^n}$. From Definition 59, we know that for the robustly-typical set, the following condition holds

$$x^n : \forall x \in \mathcal{X}, \quad \left| \frac{1}{n} N(x|x^n) - p_X(x) \right| \leq \delta p_X(x). \tag{5.204}$$

For a robustly-typical sequence to have an empirical distribution $N(x|x^n) = 0$, it is required that $\delta \geq 1$. So, we restrict $\delta \in (0, 1)$. Thus, by the fact that $p_X(x) > 0$ for all $x \in \mathcal{X}$, it is impossible for $N(\tilde{x}|x^n) = 0$ and $x^n \in T_\delta^{X^n}$. That is,

$$\left\| \sigma^{(2)}_{\tilde{A}\tilde{X}B} \right\|_1 = 0. \tag{5.205}$$

Consider that

$$\begin{aligned}
&\sigma^{(1)}_{\tilde{X}\tilde{A}B} \\
&= \sum_{\tilde{x}} p(\tilde{x}) [\tilde{x}]_{\tilde{X}} \otimes \sum_{a^n, x^n \in S_1(\tilde{x}), \tilde{a}} \sum_{i=1}^{n} \frac{1}{N(\tilde{x}|x^n)} \delta_{a_i,\tilde{a}} \delta_{\tilde{x},x_i} [\tilde{a}]_{\tilde{A}} \otimes p_{X^n}(x^n) q_{A^n|X^n}(a^n|x^n) \sigma_B^{a^n,x^n}, \\
&\hspace{30em} (5.206) \\
&= \sum_{\tilde{x}} p(\tilde{x}) [\tilde{x}]_{\tilde{X}} \otimes \sum_{\tilde{a}} [\tilde{a}]_{\tilde{A}} \otimes \frac{1}{n} \sum_{i=1}^{n} \sum_{x^{[n]\setminus\{i\},\tilde{x}} \in S_1(\tilde{x}), a^{[n]\setminus\{i\}}} \frac{p_{\tilde{X}}(\tilde{x})}{N(\tilde{x}|x^n)/n} p_{X^{[n]\setminus\{i\}}}(x^{[n]\setminus\{i\}}|\tilde{x}) \\
&\quad q(\tilde{a}|x^{[n]\setminus\{i\},\tilde{x}}) q(a^{[n]\setminus\{i\}}|x^{[n]\setminus\{i\},\tilde{x}}\tilde{a}) \sigma_B^{a^{[n]\setminus\{i\}},x^{[n]\setminus\{i\}},\tilde{x},\tilde{a}}, \qquad (5.207)
\end{aligned}$$

where $x^{[n]\setminus\{i\},\tilde{x}}$ refers to a sequence x^n with $x_i = \tilde{x}$.

We now want to give an upper bound on the second term in (5.200):

$$\left\| \frac{1}{n} \sum_{i=1}^{n} \sigma_{\bar{A}_i X_i B} - \sigma^{(1)}_{\tilde{A}XB} \right\|_1, \tag{5.208}$$

where

$$\sigma_{\bar{A}_i X_i B} = \sum_{a^n, x^n} p_{X^n}(x^n) q_{\bar{A}^n|X^n}(a^n|x^n) \, |x_i\rangle\!\langle x_i|_{X_i} \otimes |a_i\rangle\!\langle a_i|_{\bar{A}_i} \otimes \sigma_B^{a^n,x^n}. \tag{5.209}$$

Let us define the following sets:

- $S_1(x_i)$: set of sequences x^n such that $x_i \in x^n$ and $x^n \in T_\delta^{X^n}$,
- $S_2(x_i)$: set of sequences x^n such that $x_i \notin x^n$ and $x^n \in T_\delta^{X^n}$,

- S_3: set of sequences x^n such that $x^n \notin T_\delta^{X^n}$.

Then,

$$\begin{aligned}
\sigma_{\bar{A}_i X_i B} = &\sum_{a^n, x^n \in S_1(x_i)} p_{X^n}(x^n) q_{\bar{A}^n|X^n}(a^n|x^n) \, |x_i\rangle\langle x_i|_{X_i} \otimes |a_i\rangle\langle a_i|_{\bar{A}_i} \otimes \sigma_B^{a^n,x^n} \\
&+ \sum_{a^n, x^n \in S_2(x_i)} p_{X^n}(x^n) q_{\bar{A}^n|X^n}(a^n|x^n) \, |x_i\rangle\langle x_i|_{X_i} \otimes |a_i\rangle\langle a_i|_{\bar{A}_i} \otimes \sigma_B^{a^n,x^n} \\
&+ \sum_{a^n, x^n \in S_3} p_{X^n}(x^n) q_{\bar{A}^n|X^n}(a^n|x^n) \, |x_i\rangle\langle x_i|_{X_i} \otimes |a_i\rangle\langle a_i|_{\bar{A}_i} \otimes \sigma_B^{a^n,x^n} \quad (5.210) \\
= &\ \sigma^{(1)}_{\bar{A}_i X_i B} + \sigma^{(2)}_{\bar{A}_i X_i B} + \sigma^{(3)}_{\bar{A}_i X_i B}. \quad (5.211)
\end{aligned}$$

Then, using the convexity of trace distance with (5.208) and typicality arguments similar to (5.203) and (5.205), we find that

$$\begin{aligned}
\left\| \frac{1}{n} \sum_{i=1}^{n} \sigma_{\bar{A}_i X_i B} - \sigma^{(1)}_{\bar{A} X B} \right\|_1 &\leq \frac{1}{n} \sum_{i=1}^{n} \left\| \sigma_{\bar{A}_i X_i B} - \sigma^{(1),i}_{\bar{A} X B} \right\|_1 \quad (5.212) \\
&\leq \frac{1}{n} \sum_{i=1}^{n} \left\| \sigma^{(1)}_{\bar{A}_i X_i B} - \sigma^{(1),i}_{\bar{A} X B} \right\|_1 + \varepsilon_1, \quad (5.213)
\end{aligned}$$

where

$$\begin{aligned}
&\sigma^{(1)}_{\bar{A}_i X_i B} = \sum_{x_i} p_{X_i}(x_i) \, [x_i]_{X_i} \otimes \sum_{a_i} [a_i]_{\bar{A}_i} \\
&\otimes \sum_{x^{[n]\setminus\{i\}}, x_i \in S_1(x_i), a^{[n]\setminus\{i\}}} p(x^{[n]\setminus\{i\}}|x_i) q(\tilde{a}|x^{[n]\setminus\{i\}}, x_i) q(a^{[n]\setminus\{i\}}|x^{[n]\setminus\{i\}}, x_i, \tilde{a}) \sigma_B^{a^{[n]\setminus\{i\}} x^{[n]\setminus\{i\}}, x_i, a_i}.
\end{aligned} \quad (5.214)$$

and

$$\sigma^{(1),i}_{\bar{A} X B} = \sum_{\tilde{x}} p(\tilde{x}) [\tilde{x}]_{\tilde{X}} \otimes \sum_{\tilde{a}} [\tilde{a}]_{\tilde{A}} \otimes \sum_{x^{[n]\setminus\{i\}}, \tilde{x} \in S_1(\tilde{x}), a^{[n]\setminus\{i\}}} \frac{p_{\tilde{X}}(\tilde{x})}{N(\tilde{x}|x^n)/n} p_{X^{[n]\setminus\{i\}}}$$

$$(x^{[n]\setminus\{i\}}|\tilde{x})q(\tilde{a}|x^{[n]\setminus\{i\},\tilde{x}}) \quad q(a^{[n]\setminus\{i\}}|x^{[n]\setminus\{i\},\tilde{x}}\tilde{a})\sigma_B^{a^{[n]\setminus\{i\}},x^{[n]\setminus\{i\}},\tilde{x},\tilde{a}}. \tag{5.215}$$

Invoking (5.204), we find that

$$\frac{1}{n}\sum_{i=1}^{n}\left\|\sigma^{(1)}_{\bar{A}_iX_iB} - \sigma^{(1),i}_{\bar{A}XB}\right\|_1 \leq \frac{\delta}{1-\delta}, \tag{5.216}$$

where $\delta \in (0,1)$. After combining (5.203), (5.205), (5.213), and (5.216), we obtain

$$\|\rho_{\bar{A}XB} - \sigma_{\tilde{A}\tilde{X}B}\|_1 \leq nt + \frac{\delta}{1-\delta} + 2\varepsilon_1. \tag{5.217}$$

Minimizing over all possible no-signaling extensions, as required by the definition, we find that

$$\|\rho_{\bar{A}XB} - \sigma_{\tilde{A}\tilde{X}B}\|_1 \leq n \inf_{\rho_{\bar{A}XBE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1. \tag{5.218}$$

Since $\rho_{\bar{A}XB}$ and $\sigma_{\bar{A}XB}$ are classical-quantum states with $p_X(x) = \frac{1}{|\mathcal{X}|}$, we obtain

$$\sum_x \left\|\rho^x_{\bar{A}B} - \sigma^x_{\tilde{A}B}\right\|_1 \leq |\mathcal{X}|\left(n \inf_{\rho_{\bar{A}XBE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1\right). \tag{5.219}$$

This implies that the following inequality holds for all $x \in \mathcal{X}$:

$$\left\|\rho^x_{\bar{A}B} - \sigma^x_{\tilde{A}B}\right\|_1 \leq |\mathcal{X}|\left(n \inf_{\rho_{\bar{A}XBE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1\right). \tag{5.220}$$

This means that we can average the above to get a bound for any arbitrary distribution $p(x)$ on x. Therefore, we can now relax the assumption of a uniform probability distribution, in order to obtain the following bound for an arbitrary probability distribution:

$$\sup_{p_X(x)} \|\rho_{\bar{A}BX} - \sigma_{\tilde{A}BX}\|_1 \leq |\mathcal{X}| \left(n \sup_{p_X(x)} \inf_{\rho_{\bar{A}XBE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1 \right), \tag{5.221}$$

which implies that

$$\sup_{p_X(x)} \|\rho_{\bar{A}BX} - \sigma_{\tilde{A}BX}\|_1 \leq |\mathcal{X}| \left(n\sqrt{S^R(\bar{A};B)_{\hat{\rho}} \ln 2} + \frac{\delta}{1-\delta} + 2\varepsilon_1 \right). \tag{5.222}$$

Given $S^R(\bar{A};B)_{\hat{\rho}} \leq \varepsilon$ (as required by the condition of faithfulness), choose $n = (1/\varepsilon)^{1/4}$, $\delta = \varepsilon^{1/16}|\mathcal{X}|^{1/2}$ (recall that we require $\delta \in (0,1)$). We know by the Chernoff bound [109] that $\varepsilon_1 = 2|\mathcal{X}|e^{-\frac{1}{3|\mathcal{X}|}\delta^2 n}$. Substituting these values, we find that

$$\|\rho_{\bar{A}BX} - \sigma_{\tilde{A}BX}\|_1 \leq |\mathcal{X}| \left(\varepsilon^{1/4} + \frac{\varepsilon^{1/16}|\mathcal{X}|^{1/2}}{1-\varepsilon^{1/16}|\mathcal{X}|^{1/2}} + 4|\mathcal{X}|e^{-\frac{\varepsilon^{-1/4}}{3}} \right). \tag{5.223}$$

This concludes the proof. ■

5.5. Open questions

In this section, we state two open questions regarding the properties of intrinsic steerability and outline the proof attempts by the author. Besides the open questions listed below, the calculation of intrinsic steerability for various assemblages is an interesting open question. These calculations could provide insights for other quantum information phenomena as in [110].

5.5.1. Continuity of restricted intrinsic steerability

Suppose that we are given two assemblages $\hat{\rho}^{a,x}$ and $\hat{\sigma}^{a,x}$ such that the following holds:

$$\frac{1}{2}\|\hat{\rho}^{a,x}_B - \hat{\sigma}^{a,x}_B\|_1 \leq \varepsilon. \tag{5.224}$$

We want to prove the following:

$$|S^R(A;B)_{\hat{\rho}} - S^R(A;B)_{\hat{\sigma}}| \overset{?}{\leq} g_1(\varepsilon) + g_2\left(\varepsilon \log d\right) \tag{5.225}$$

where $g_1(\varepsilon), g_2(\varepsilon \log d) \to 0$ as $\varepsilon \to 0$ and d is equal to the min $\left[\dim \mathcal{H}_{\bar{A}}, \dim(\mathcal{H}_B)\right]$.

Proof attempt: Let us choose a particular no-signaling extension $\hat{\rho}^{a,x}_{BE}$ of $\hat{\rho}^{a,x}_{B}$. Then can we construct a no-signaling extension $\hat{\sigma}^{a,x}_{BE}$ of $\hat{\sigma}^{a,x}_{B}$ such that

$$\frac{1}{2}\|\hat{\rho}^{a,x}_{BE} - \hat{\sigma}^{a,x}_{BE}\|_1 \overset{?}{\leq} \varepsilon_1, \tag{5.226}$$

where $\varepsilon_1 \to 0$ as $\varepsilon \to 0$? If such a construction is possible, then by continuity of CMI, we have

$$|I(\bar{A};B|XE)_{\hat{\rho}} - I(\bar{A};B|XE)_{\hat{\sigma}}| \leq f(\varepsilon_1), \tag{5.227}$$

where $f(\varepsilon_1) = 2\varepsilon_1 \log d + 2g(\varepsilon_1)$. Here, $g(\varepsilon) = (\varepsilon+1)\log(\varepsilon+1) - \varepsilon \log \varepsilon$. Because the above inequality holds for any extension of the assemblage $\hat{\rho}^{a,x}_{B}$, we obtain the continuity of restricted intrinsic steerability:

$$|S^R(\bar{A};B)_{\hat{\rho}} - S^R(\bar{A};B)_{\hat{\sigma}}| \leq f(\varepsilon). \tag{5.228}$$

The bottleneck of the proof is the statement (5.226). We cannot directly invoke Uhlmann's theorem to prove this statement.

A similar statement can be conjectured for continuity of intrinsic steerability.

5.5.2. Squashed entanglement not less than intrinsic steerability

We suspect that squashed entanglement of a quantum state will be greater than the intrinsic steerability of any assemblage obtained from this state. Consider a bipartite state ρ_{AB} and consider $\hat{\rho}^{a,x}_{B}$ as an arbitrary assemblage obtained from ρ_{AB}.

Then, we conjecture the following:

$$E_{\mathrm{sq}}(A;B)_\rho \overset{?}{\geq} S^R(\bar{A};B)_{\hat{\rho}}. \tag{5.229}$$

Consider a particular extension ρ_{ABE} of ρ_{AB}. Let $\rho_{\bar{A}BX}$ be a classical-quantum state associated with the assemblage $\hat{\rho}_B^{a,x}$. Let $\rho_{\bar{A}BXE}$ be a non-signaling extension of $\rho_{\bar{A}BX}$ obtained from ρ_{AB}. Then,

$$I(A;B|E)_\rho \geq I(\bar{A};B|EX)_\rho \tag{5.230}$$

$$I(A;B|E)_\rho \geq S^R(\bar{A};B)_{\hat{\rho}} \tag{5.231}$$

$$\inf_{\rho_{ABE}} I(A;B|E)_\rho \geq S^R(\bar{A};B)_{\hat{\rho}} \tag{5.232}$$

The first inequality follows from monotonicity of CMI under local operations and the chain rule. The second inequality follows because the non-signaling extension $\rho_{\bar{A}BXE}$ in (5.230) is a particular kind of non-signaling extension. The third inequality follows because ρ_{ABE} is a particular extension of ρ_{AB}. However, the definition of squashed entanglement contains a factor of half for normalization and that cannot be taken into account by the aforementioned steps.

Chapter 6
Intrinsic Non-Locality

In this chapter, we introduce intrinsic non-locality and quantum intrinsic non-locality as quantifiers of non-local distributions based on conditional mutual information. We prove that they fulfill several desirable properties, such as monotonicity under local operations and shared randomness, convexity, faithfulness, superadditivity, and additivity with respect to tensor products.

6.1. Definitions of quantifiers

6.1.1. Definition of intrinsic non-locality

To calculate the amount of non-locality present in the distribution $p(a,b|x,y)$, we introduce a function $N : p(a,b|x,y) \to \mathbb{R}_{\geq 0}$, which we call *intrinsic non-locality.* Consider a distribution $p(a,b|x,y) \in \mathbf{NS}$. Now embed the distribution $p(a,b|x,y)$ into a classical state as

$$\rho_{\bar{A}\bar{B}XY} := \sum_{a,b,x,y} p(x,y)p(a,b|x,y)\, [a\, b\, x\, y]_{\bar{A}\bar{B}XY}\,, \tag{6.1}$$

where $p(x,y)$ is a probability distribution for the measurement choices x and y. Consider a no-signaling extension $\rho_{\bar{A}\bar{B}XYE}$ of $\rho_{\bar{A}\bar{B}XY}$:

$$\rho_{\bar{A}\bar{B}XYE} := \sum_{a,b,x,y} p(x,y)\, [a\, b\, x\, y]_{\bar{A}\bar{B}XY} \otimes p(a,b|x,y)\rho_E^{a,b,x,y}, \tag{6.2}$$

such that $\mathrm{Tr}_E(\rho_{\bar{A}\bar{B}XYE}) = \rho_{\bar{A}\bar{B}XY}$, and the following no-signaling constraints hold:

$$\sum_a p(a,b|x,y)\rho_E^{a,b,x,y} = \sum_a p(a,b|x',y)\rho_E^{a,b,x',y} \quad \forall x,x' \in \mathcal{X}. \tag{6.3}$$

It is then easy to see that given the value in system Y, the state of systems X and systems $\bar{B}E$ is product. This is equivalent to the following constraint on conditional mutual information:

$$I(\bar{B}E;X|Y)_\rho = 0 \quad \forall p(x,y). \tag{6.4}$$

Similarly, the following no-signaling constraints hold

$$\sum_b p(a,b|x,y)\rho_E^{a,b,x,y} = \sum_b p(a,b|x,y')\rho_E^{a,b,x,y'} \quad \forall y,y' \in \mathcal{Y}. \tag{6.5}$$

It is easy to see that given the value in systems X, the state of systems Y and $\bar{A}E$ is product. This is equivalent to the following constraint on conditional mutual information

$$I(\bar{A}E;Y|X)_\rho = 0 \quad \forall p(x,y). \tag{6.6}$$

Finally, we obtain

$$\begin{aligned}\sum_{a,b} p(a,b|x,y)\rho_E^{a,b,x,y} &= \sum_{a,b} p(a,b|x',y)\rho_E^{a,b,x',y} && (6.7)\\ &= \sum_{a,b} p(a,b|x',y')\rho_E^{a,b,x',y'} \quad \forall x,x' \in \mathcal{X}, y,y' \in \mathcal{Y}. && (6.8)\end{aligned}$$

The first equality follows from (6.3), and the second equality follows from (6.5). This implies that the state of Eve's system is independent of the measurement choices, i.e., $I(XY;E)_\rho = 0$ for all $p(x,y)$. We can then quantify the amount of non-local distributions in the distribution $p(a,b|x,y)$ as $\inf_{\rho_{\bar{A}\bar{B}XYE}} I(\bar{A};\bar{B}|XYE)$, where the infimum is with respect to no-signaling extensions $\rho_{\bar{A}\bar{B}XYE}$ of the above form. Since Alice and Bob want to maximize the non-local distributions of the two black boxes, we maximize over input probability distributions $p(x,y)$, leading us to the following definition:

Definition 61 (Intrinsic non-locality) *The intrinsic non-locality of a distribution* $p(a,b|x,y) \in \mathbf{NS}$ *is defined as*

$$N(\bar{A};\bar{B})_p = \sup_{p(x,y)} \inf_{\rho_{\bar{A}\bar{B}XYE}} I(\bar{A};\bar{B}|XYE)_\rho, \tag{6.9}$$

where $\rho_{\bar{A}\bar{B}XYE}$ *is a no-signaling extension of the state* $\rho_{\bar{A}\bar{B}XY}$*, i.e., subject to the constraints in* (6.3) *and* (6.5)*.*

In the next chapter 7, we obtain an upper bound on the distillable key of a distribution. With that in mind, we can also think of the extension system as a system with an eavesdropper.

6.1.2. Definition of quantum intrinsic non-locality

We now introduce a function $N^Q : p(a,b|x,y) \to \mathbb{R}_{\geq 0}$, which we call *quantum intrinsic non-locality*, with $p(a,b|x,y) \in \mathbf{Q}$. As stated in Section 2.1.3., a distribution in the set $\mathbf{Q}$ arises from some underlying state ρ_{AB} and POVMs of Alice and Bob characterized by $\{\Lambda_x^a\}_a$ and $\{\Lambda_y^b\}_b$, respectively.[1] Now, consider a quantum state ρ_{ABE} such that $\mathrm{Tr}_E(\rho_{ABE}) = \rho_{AB}$. We call ρ_{ABE} an extension of the state ρ_{AB}. Then, one possible extension of the classical-classical state $\rho_{\bar{A}\bar{B}XY}$ as defined in (6.1) is

$$\begin{aligned} \rho_{\bar{A}\bar{B}XYE} &= \sum_{a,b,x,y} p(x,y)\, \mathrm{Tr}_{AB}\left[\left(\Lambda_x^a \otimes \Lambda_y^b \otimes I_E\right)\rho_{ABE}\right] [a\,b\,x\,y]_{\bar{A}\bar{B}XY}\,, & (6.10)\\ &= \sum_{a,b,x,y} p(x,y)p(a,b|x,y)\, [a\,b\,x\,y]_{\bar{A}\bar{B}XY} \otimes \rho_E^{a,b,x,y}, & (6.11) \end{aligned}$$

where $p(a,b|x,y)\rho_E^{a,b,x,y} := \mathrm{Tr}_{AB}\left[\left(\Lambda_x^a \otimes \Lambda_y^b \otimes I_E\right)\rho_{ABE}\right]$. By definition, this extension is also a no-signaling extension and is subjected to the constraints in (6.3) and (6.5).

[1] For certain quantum distributions, it is possible to pinpoint the underlying quantum state and POVMs up to certain isometries. See [111, 112] in this context.

We call extensions of the form in (6.10) as *quantum extensions.*

For $p \in \mathbf{Q}$, the set of no-signaling extensions of p is strictly larger than the set of quantum extensions. For example, in the CHSH game, a distribution $p(a, b|x, y)$ reaching the Tsirelson bound only admits a trivial quantum extension, i.e., with constant $\rho_E^{a,b,x,y}$ independent of a, b, x, and y. However, the no-signaling extensions of such a distribution are not extremal, as can be seen by writing $p(a, b|x, y)$ as a convex combination of a PR box (with necessarily constant $\rho_E^{a,b,x,y}$ as an extension) and a local box (where $\rho_E^{a,b,x,y}$ contains the local hidden variable).

Therefore, to consider the setting in which there is an underlying quantum model, we define *quantum intrinsic non-locality* as follows:

Definition 62 (Quantum intrinsic non-locality) *The quantum intrinsic non-locality of a distribution $p(a, b|x, y) \in \mathbf{Q}$ is defined as*

$$N^Q(\bar{A}; \bar{B})_p = \sup_{p(x,y)} \inf_{\rho_{\bar{A}\bar{B}XYE}} I(\bar{A}; \bar{B}|XYE)_\rho, \tag{6.12}$$

where $\rho_{\bar{A}\bar{B}XYE}$ is a quantum extension of the state $\rho_{\bar{A}\bar{B}XY}$, that is, subject to the constraints in (6.10)*.*

Proposition 63 *If $p(a, b|x, y) \in \mathbf{Q}$, then*

$$N(\bar{A}; \bar{B})_p \leq N^Q(\bar{A}; \bar{B})_p. \tag{6.13}$$

Proof. This follows from the observation that a quantum extension $\sigma_{\bar{A}\bar{B}XYE}$ of $\rho_{\bar{A}\bar{B}XY}$ is a particular kind of no-signaling extension. ■

In general, we expect the calculation of intrinsic non-locality to be hard. This again can be traced back to the problem of finding the appropriate extension system, since we do not know any bound on the dimension of the extension system.

6.1.3. Intrinsic non-locality of a PR box

In this section, we calculate the intrinsic non-locality of a PR box.

Proposition 64 *The intrinsic non-locality of a PR box is equal to* 1*, i.e.,* $N(\bar{A};\bar{B})_p = 1$*, where* p *is the distribution defined in* (2.13)*.*

Proof. Consider the state

$$\rho_{\bar{A}\bar{B}XY} := \sum_{a,b,x,y} p(x,y)p(a,b|x,y)\, [a\, b\, x\, y]_{\bar{A}\bar{B}XY}\,, \tag{6.14}$$

where $p(x,y)$ is an arbitrary probability distribution. Consider a no-signaling extension of the state

$$\rho_{\bar{A}\bar{B}XYE} := \sum_{a,b,x,y} p(x,y)p(a,b|x,y)\, [a\, b\, x\, y]_{\bar{A}\bar{B}XY} \otimes \rho_E^{a,b,x,y}. \tag{6.15}$$

The no-signaling constraints are

$$\sum_{a,b,y} p(a,b|x,y)\, [b\, x\, y]_{\bar{B}XY} \otimes \rho_E^{x,y,a,b} = \sum_{a,b,y} p(a,b|x',y)\, [b\, x'\, y]_{\bar{B}XY} \otimes \rho_E^{x',y,a,b}, \tag{6.16}$$

$$\sum_{b,a,x} p(a,b|x,y)\, [a\, x\, y]_{\bar{A}XY} \otimes \rho_E^{x,y,a,b} = \sum_{b,a,x} p(a,b|x,y')\, [a\, x\, y']_{\bar{A}XY} \otimes \rho_E^{x,y',a,b}. \tag{6.17}$$

From (2.13), and the no-signaling constraint in (6.16), we arrive at the following constraints on the possible states of Eve's system:

$$\begin{bmatrix} \rho_E^{0000} & 0 & 0 & 0 \\ 0 & \rho_E^{0011} & 0 & 0 \\ 0 & 0 & \rho_E^{0100} & 0 \\ 0 & 0 & 0 & \rho_E^{0111} \end{bmatrix} = \begin{bmatrix} \rho_E^{1000} & 0 & 0 & 0 \\ 0 & \rho_E^{1011} & 0 & 0 \\ 0 & 0 & \rho_E^{1110} & 0 \\ 0 & 0 & 0 & \rho_E^{1101} \end{bmatrix}. \tag{6.18}$$

In the matrices given above, the rows and columns are indexed by (y, b). The first matrix on the left corresponds to $x = 0$, and the second one on the right corresponds to $x = 1$. The constraints in (6.18) can also be written as

$$\begin{aligned} &1)\ \rho_E^{0000} = \rho_E^{1000}, && 2)\ \rho_E^{0011} = \rho_E^{1011}, \\ &3)\ \rho_E^{0100} = \rho_E^{1110}, && 4)\ \rho_E^{0111} = \rho_E^{1101}. \end{aligned} \tag{6.19}$$

Similarly, from (2.13), and the no-signaling constraint in (6.17), we arrive at the following constraints on the possible states of Eve's system:

$$\begin{bmatrix} \rho_E^{0000} & 0 & 0 & 0 \\ 0 & \rho_E^{0011} & 0 & 0 \\ 0 & 0 & \rho_E^{1000} & 0 \\ 0 & 0 & 0 & \rho_E^{1011} \end{bmatrix} = \begin{bmatrix} \rho_E^{0100} & 0 & 0 & 0 \\ 0 & \rho_E^{0111} & 0 & 0 \\ 0 & 0 & \rho_E^{1101} & 0 \\ 0 & 0 & 0 & \rho_E^{1110} \end{bmatrix}. \tag{6.20}$$

In the above block matrices, the rows and columns are indexed by (x, a). The first matrix on the left corresponds to $y = 0$, and the second one on the right corresponds to $y = 1$. The constraints in (6.20) can also be written as

$$\begin{aligned} &5)\ \rho_E^{0000} = \rho_E^{0100}, && 6)\ \rho_E^{0011} = \rho_E^{0111}, \\ &7)\ \rho_E^{1000} = \rho_E^{1101}, && 8)\ \rho_E^{0111} = \rho_E^{1101}. \end{aligned} \tag{6.21}$$

By following 1 → 7 → 4 → 6 → 2 → 8 → 3 → 5 → 1 in the above, we obtain $\rho_E^{x,y,a,b} = \rho_E^{x',y',a',b'} \quad \forall x, x', y, y' \in [s]$ and $a, a', b, b' \in [r]$. This implies that $\rho_{\bar{A}\bar{B}XY}$ has a trivial tensor-product no-signaling extension. Hence,

$$\begin{aligned} I(\bar{A}; \bar{B}|XYE)_\rho = I(\bar{A}; \bar{B}|XY)_\rho &= \sum_{x,y} p(x, y) I(\bar{A}; \bar{B})_{\rho^{x,y}} && (6.22) \\ &= \sum_{x,y} p(x, y) \left(H(\bar{A})_{\rho^{x,y}} - H(\bar{A}|\bar{B})_{\rho^{x,y}} \right) && (6.23) \\ &= 1. && (6.24) \end{aligned}$$

It is easy to check that given realizations of X, Y, the entropies $H(\bar{A}|\bar{B})_{\rho^{x,y}} = 0$ and $H(\bar{A})_{\rho^{x,y}} = 1$. ■

6.2. Properties of intrinsic non-locality and quantum intrinsic non-locality

In this section, we prove that intrinsic non-locality and quantum intrinsic non-locality are faithful, monotone with respect to local operations and shared randomness, superadditive, and additive with respect to tensor products of distributions. These are the properties that are desirable for a measure of Bell non-locality to possess, as discussed in Section 2.2.3.. We also prove that quantum intrinsic non-locality of a distribution is never larger than the restricted intrinsic steerability of an associated assemblage.

6.2.1. Monotonicity under free operations

We expect any quantifier of non-locality to be monotone with respect to local operations and shared randomness. That is, a free operation should not increase the amount of non-locality in the device. We state this in the following proposition:

Proposition 65 (Monotonicity of intrinsic non-locality) *Let $p_i(a,b|x,y)$ be a distribution, and let $p_f(a_f,b_f|x_f,y_f)$ be a distribution that results from the action of local operations and shared randomness on $p_i(a,b|x,y)$, so that we can write the final probability distribution as follows:*

$$p_f(a_f,b_f|x_f,y_f) := \sum_{a,b,x,y} O^{(L)}(a_f,b_f|a,b,x,y,x_f,y_f)\, p_i(a,b|x,y)\, I^{(L)}(x,y|x_f,y_f), \tag{6.25}$$

where $I^{(L)}(x,y|x_f,y_f)$ and $O^{(L)}(a_f,b_f|a,b,x,y,x_f,y_f)$ are local boxes as described in (2.25) *and* (2.26). *Then,*

$$N(\bar{A};\bar{B})_{p_i} \geq N(\bar{A}_f;\bar{B}_f)_{p_f}. \tag{6.26}$$

Proof. First, we embed $p_f(a_f,b_f|x_f,y_f)$ in a quantum state:

$$\rho_{\bar{A}_f\bar{B}_fX_fY_f} = \sum_{x_f,y_f,a_f,b_f} p(x_f,y_f)\, p_f(a_f,b_f|x_f,y_f)\ [x_f\, y_f\, a_f\, b_f]_{X_fY_f\bar{A}_f\bar{B}_f}\,, \tag{6.27}$$

where $p(x_f,y_f)$ is an arbitrary probability distribution for x_f,y_f. Then invoking (2.24), (2.25), and (2.26), we obtain

$$\rho_{\bar{A}_f\bar{B}_fX_fY_f} = \sum_{x_f,y_f,a_f,b_f} p(x_f,y_f) \sum_{a,b,x,y}\sum_{\lambda_2} p_{\Lambda_2}(\lambda_2)\, O_A(a_f|a,x_f,x,\lambda_2)\, O_B(b_f|b,y,y_f,\lambda_2)\times$$
$$p_i(a,b|x,y)\sum_{\lambda_1} p_{\Lambda_1}(\lambda_1)\, I_A(x|x_f,\lambda_1)\, I_B(y|y_f,\lambda_1)\, [x_f\, y_f\, a_f\, b_f]_{X_fY_f\bar{A}_f\bar{B}_f}\,. \tag{6.28}$$

An arbitrary extension of the state in (6.27) is given by

$$\rho_{\bar{A}_f\bar{B}_fX_fY_fE} = \sum_{x_f,y_f,a_f,b_f} p(x_f,y_f)\, p_f(a_f,b_f|x_f,y_f)\, [x_f\, y_f\, a_f\, b_f]_{X_fY_f\bar{A}_f\bar{B}_f} \otimes \rho_E^{a_f,b_f,x_f,y_f}. \tag{6.29}$$

A particular extension of the state in (6.27) is given by

$$\begin{aligned}\zeta_{\bar{A}_f\bar{B}_fX_fY_fE\Lambda_1\Lambda_2} = \sum_{x_f,y_f,a_f,b_f} p(x_f,y_f) \sum_{a,b,x,y} \sum_{\lambda_2} p_{\Lambda_2}(\lambda_2)\, O_A(a_f|a,x_f,x,\lambda_2)\times \\ O_B(b_f|b,y,y_f,\lambda_2)\, p_i(a,b|x,y)\times \\ \sum_{\lambda_1} p_{\Lambda_1}(\lambda_1) I_A(x|x_f,\lambda_1)\, I_B(y|y_f,\lambda_1)\, [x_f\, y_f\, a_f\, b_f]_{\bar{A}_f\bar{B}_fX_fY_f} \otimes \tau_E^{a,b,x,y} \otimes [\lambda_1\lambda_2]_{\Lambda_1\Lambda_2}.\end{aligned} \tag{6.30}$$

This in turn is a marginal of the following state:

$$\begin{aligned}\zeta_{\bar{A}_f\bar{B}_fX_fY_fE\Lambda_1\Lambda_2XY\bar{A}\bar{B}} = \sum_{x_f,y_f,a_f,b_f} p(x_f,y_f) \sum_{a,b,x,y} \sum_{\lambda_2} p_{\Lambda_2}(\lambda_2) O_A(a_f|a,x_f,x,\lambda_2) \\ O_B(b_f|b,y,y_f,\lambda_2)\times p_i(a,b|x,y) \sum_{\lambda_1} p_{\Lambda_1}(\lambda_1)\, I_A(x|x_f,\lambda_1)\, I_B(y|y_f,\lambda_1)\, [x_f\, y_f\, a_f\, b_f]_{X_fY_f\bar{A}_f\bar{B}_f} \\ \otimes \tau_E^{a,b,x,y} \otimes [\lambda_1\lambda_2]_{\Lambda_1\Lambda_2} \otimes [x\, y\, a\, b]_{XY\bar{A}\bar{B}}.\end{aligned} \tag{6.31}$$

Consider that

$$\inf_{\text{ext. in (6.29)}} I(\bar{A}_f;\bar{B}_f|X_fY_fE)_\rho \leq I(\bar{A}_f;\bar{B}_f|X_fY_fE\Lambda_1\Lambda_2)_\zeta \tag{6.32}$$

$$\leq I(\bar{A}X_fX\Lambda_2;\bar{B}Y_fY\Lambda_2|X_fY_fE\Lambda_1\Lambda_2)_\zeta \tag{6.33}$$

$$= I(\bar{A}X;\bar{B}Y|X_fY_fE\Lambda_1\Lambda_2)_\zeta \tag{6.34}$$

$$= I(\bar{A}X;\bar{B}Y|X_fY_fE\Lambda_1)_\zeta \tag{6.35}$$

$$
\begin{aligned}
&= I(\bar{A};\bar{B}|XYX_fY_fE\Lambda_1)_\zeta + I(X;\bar{B}|X_fY_fE\Lambda_1Y)_\zeta \\
&\quad + I(Y;\bar{A}|X_fY_fE\Lambda_1X)_\zeta + I(X;Y|X_fY_f\Lambda_1E)_\zeta .
\end{aligned}
\tag{6.36}
$$

The first inequality follows from considering a particular extension in (6.30). The second inequality follows from data processing of conditional mutual information. The second equality follows because $\zeta_{\bar{A}\bar{B}XYX_fY_fE\Lambda_1\Lambda_2} = \zeta_{\bar{A}\bar{B}XYX_fY_fE\Lambda_1} \otimes \zeta_{\lambda_2}$. The last equality follows from the chain rule for conditional mutual information. Now, let us consider each term in (6.36). By inspection,

$$
\begin{aligned}
\zeta_{\bar{A}\bar{B}XYX_fY_fE\Lambda_1} = \sum_{x_f,y_f} p(x_f,y_f) \sum_{a,b,x,y,\lambda} p(\lambda_1)p_i(a,b|x,y)p(x,y|x_f,y_f,\lambda_1) \\
[x_f\, y_f\, \lambda_1\, x\, y\, a\, b]_{X_fY_f\Lambda_1XY\bar{A}\bar{B}} \otimes \tau_E^{a,b,x,y}.
\end{aligned}
\tag{6.37}
$$

Upon re-arranging, we obtain

$$
\begin{aligned}
\zeta_{\bar{A}\bar{B}XYX_fY_fE\Lambda_1} = \sum_{x,y} p(x,y) \sum_{x_f,y_f,\lambda_1} p(x_f,y_f,\lambda_1|x,y)\, [x\, y\, x_f\, y_f\, \lambda_1]_{XYX_fY_f\Lambda_1} \otimes \\
\sum_{a,b} p_i(a,b|x,y)\tau_E^{a,b,x,y} \otimes [a\, b]_{\bar{A}\bar{B}}.
\end{aligned}
\tag{6.38}
$$

So, given X,Y, the states $\zeta_{\bar{A}\bar{B}E}^{x,y}$ and $\zeta_{X_fY_f\Lambda_1}^{x,y}$ are in tensor product. Therefore

$$
I(\bar{A};\bar{B}|XYX_fY_fE\Lambda_1)_\zeta = I(\bar{A};\bar{B}|XYE)_\zeta, \tag{6.39}
$$

where $\zeta_{\bar{A}\bar{B}XYE}$ is a no-signaling extension of $\rho_{\bar{A}\bar{B}XY}$. Now consider that

$\zeta_{XX_fYY_f\bar{B}E\Lambda_1}$

$$= \sum_{x,y,x_f,y_f,\lambda_1} p(x,y,x_f,y_f,\lambda_1)\, [x\, x_f\, y\, y_f\, \lambda_1]_{XX_fYY_f\Lambda_1} \otimes \sum_b p(b|y)\, \tau_E^{b,y} \otimes [b]_{\bar{B}}\,. \tag{6.40}$$

$$= \sum_y p(y)\, [y]_Y \otimes \sum_{x,x_f,y_f,\lambda_1} p(x_f,y_f,x,\lambda_1|y)\, [x\, x_f\, y\, y_f\, \lambda_1]_{XX_fYY_f\Lambda_1} \otimes \sum_b p(b|y)\, \tau_E^{b,y} \otimes [b]_{\bar{B}}\,. \tag{6.41}$$

Then, by inspection

$$I(X;\bar{B}|X_fY_fE\Lambda_1Y)_\zeta = 0. \tag{6.42}$$

Similarly, $I(Y;\bar{A}|Y_fX_fE\Lambda_1X)_\zeta = 0$.

Now, consider the term $I(X;Y|X_fY_fE\Lambda_1)_\zeta$, with

$$\zeta_{XYX_fY_fE\Lambda_1} := \sum_{x_f,y_f} p(x_f,y_f) \sum_{x,y,\lambda_1} p(x|x_f,\lambda_1)\, p(y|y_f,\lambda_1)[x\, y\, x_f\, y_f\, \lambda_1]_{XYX_fY_f\Lambda_1} \otimes \rho_E. \tag{6.43}$$

Here, X and Y are independent given X_f, Y_f, and Λ_1. Therefore, $I(X;Y|X_fY_fE\Lambda_1)_\zeta = 0$. Combining the above equations, we obtain

$$\inf_{\text{ext. in (6.29)}} I(\bar{A}_f;\bar{B}_f|X_fY_fE)_\rho \leq I(\bar{A};\bar{B}|XYE)_\zeta. \tag{6.44}$$

Since (6.44) is true for an arbitrary no-signaling extension of $\rho_{\bar{A}\bar{B}XY}$, the above inequality holds after taking the infimum over all possible no-signaling extensions $\zeta_{\bar{A}\bar{B}XYE}$.

Finally, we can take the supremum over all the measurement choices, and we find that

$$N(\bar{A}_f;\bar{B}_f)_{p_f} \leq N(\bar{A};\bar{B})_{p_i}. \tag{6.45}$$

This concludes the proof. ■

Proposition 66 (Monotonicity) *Let $p_i(a,b|x,y) \in \mathbf{Q}$, and let $p_f(a_f,b_f|x_f,y_f)$ re-*

sult from the action of local operations and shared randomness on $p_i(a, b|x, y)$. We can write the final probability distribution as follows:

$$p_f(a_f, b_f|x_f, y_f) := \sum_{a,b,x,y} O^{(L)}(a_f, b_f|a, b, x, y, x_f, y_f)\, p_i(a, b|x, y)\, I^{(L)}(x, y|x_f, y_f), \tag{6.46}$$

where $I^{(L)}(x, y|x_f, y_f)$ and $O^{(L)}(a_f, b_f|a, b, x, y, x_f, y_f)$ are local boxes as described in (2.25) *and* (2.26)*. Then,*

$$N^Q(\bar{A}; \bar{B})_{p_i} \geq N^Q(\bar{A}_f; \bar{B}_f)_{p_f}. \tag{6.47}$$

Proof. First, we embed $p_f(a_f, b_f|x_f, y_f)$ in a quantum state:

$$\rho_{\bar{A}_f\bar{B}_f X_f Y_f} = \sum_{x_f,y_f,a_f,b_f} p(x_f, y_f)\, p_f(a_f, b_f|x_f, y_f)\ [x_f\, y_f\, a_f\, b_f]_{X_f Y_f \bar{A}_f \bar{B}_f}\,, \tag{6.48}$$

where $p(x_f, y_f)$ is an arbitrary probability distribution for x_f, y_f. The set of quantum distributions $\mathbf{Q}$ is closed under the action of local operations and shared randomness, implying that $p_f(a_f, b_f|x_f, y_f) \in \mathbf{Q}$. Since $p_f(a_f, b_f|x_f, y_f)$ is also a quantum distribution, we know that there exists an underlying state σ_{AB} and POVMs $\left\{\Lambda_{x_f}^{a_f}\right\}_{a_f}$ and $\left\{\Lambda_{y_f}^{b_f}\right\}_{b_f}$, such that

$$p_f(a_f, b_f|x_f, y_f) = \operatorname{Tr}\left[\left(\Lambda_{x_f}^{a_f} \otimes \Lambda_{y_f}^{b_f}\right) \sigma_{AB}\right]. \tag{6.49}$$

An arbitrary quantum extension of the state in (6.48) is given by

$$\sigma_{\bar{A}_f\bar{B}_f X_f Y_f E} = \sum_{x_f,y_f,a_f,b_f} p(x_f, y_f)\, p_f(a_f, b_f|x_f, y_f)\, [x_f\, y_f\, a_f\, b_f]_{X_f Y_f \bar{A}_f \bar{B}_f} \otimes \sigma_E^{a_f,b_f,x_f,y_f}, \tag{6.50}$$

where

$$\sigma_E^{a_f,b_f,x_f,y_f} = \frac{1}{p_f(a_f,b_f|x_f,y_f)} \mathrm{Tr}_{AB}\left[\left(\Lambda_{x_f}^{a_f} \otimes \Lambda_{y_f}^{b_f} \otimes I_E\right) \sigma_{ABE}\right], \tag{6.51}$$

and σ_{ABE} is an extension of σ_{AB}. Now, we know that

$$p_f(a_f,b_f|x_f,y_f) := \sum_{a,b,x,y} O^{(L)}(a_f,b_f|a,b,x,y,x_f,y_f)\, p_i(a,b|x,y)\, I^{(L)}(x,y|x_f,y_f), \tag{6.52}$$

and that the distributions $I^{(L)}(x,y|x_f y_f)$ and $O^{(L)}(a_f,b_f|a,b,x,y,x_f,y_f)$ are local distributions. Therefore, there exist separable states ρ_{XY} and $\rho_{A_F B_F}$, along with the POVMs which result in the distributions $I^{(L)}$ and $O^{(L)}$. That is,

$$I^{(L)}(x,y|x_f,y_f) = \mathrm{Tr}\left[\left(\Lambda_{x_f}^{x} \otimes \Lambda_{y_f}^{y}\right) \rho_{XY}\right], \tag{6.53}$$

$$O^{(L)}(a_f,b_f|a,b,x,y,x_f,y_f) = \mathrm{Tr}\left[\left(\Lambda_{a,x_f,x}^{a_f} \otimes \Lambda_{b,b_f,y}^{b_f}\right) \rho_{A_F B_F}\right] \tag{6.54}$$

Furthermore, we know that the distribution $p_i(a,b|x,y)$ is a quantum distribution. Therefore, it has an underlying state ρ_{AB} and POVMs characterized by $\{\Lambda_x^a\}_a$ and $\{\Lambda_y^b\}_b$. Then

$$p(a_f,b_f|x_f,y_f) = \sum_{a,b,x,y} \mathrm{Tr}\left[\left(\Lambda_{a,x_f,x}^{a_f} \otimes \Lambda_{b,b_f,y}^{b_f} \otimes \Lambda_x^a \otimes \Lambda_y^b \otimes \Lambda_{x_f}^{x} \otimes \Lambda_{y_f}^{y}\right)\left(\rho_{A_F B_F} \otimes \rho_{AB} \otimes \rho_{XY}\right)\right]. \tag{6.55}$$

Since ρ_{XY} is a separable state, we can write it as $\rho_{XY} = \sum_{\lambda_1} p(\lambda_1)\rho_X^{\lambda_1} \otimes \rho_Y^{\lambda_1}$. Let $\rho_{XY\Lambda_1} = \sum_{\lambda_1} p(\lambda_1)\rho_X^{\lambda_1} \otimes \rho_Y^{\lambda_1} \otimes [\lambda_1]_{\Lambda_1}$ be a particular extension of ρ_{XY}. Similarly, let $\rho_{A_F B_F \Lambda_2}$ be an extension of $\rho_{A_F B_F}$ and ρ_{ABE} an extension of ρ_{AB}.

A particular quantum extension of the state in (6.48) is given by

$$\rho_{\bar{A}_f \bar{B}_f X_f Y_f E \Lambda_1 \Lambda_2} \\ = \sum_{x_f, y_f, a_f, b_f} p(x_f, y_f) p_f(a_f, b_f | x_f, y_f) \left[x_f, y_f, a_f, b_f \right]_{X_f Y_f A_f B_f} \otimes \rho_E^{a,b,x,y} \otimes [\lambda_1 \lambda_2]_{\Lambda_1 \Lambda_2}, \tag{6.56}$$

where

$$\rho_E^{a,b,x,y} = \frac{1}{p(a,b|x,y)} \operatorname{Tr}_{AB} \left[\left(\Lambda_x^a \otimes \Lambda_y^b \otimes I_E \right) \rho_{ABE} \right]. \tag{6.57}$$

Then it follows that

$$\rho_{\bar{A}_f \bar{B}_f X_f Y_f E \Lambda_1 \Lambda_2} = \sum_{x_f, y_f, a_f, b_f} p(x_f, y_f) \sum_{a,b,x,y} \sum_{\lambda_2} p_{\Lambda_2}(\lambda_2)\, O_A(a_f | a, x_f, x, \lambda_2) \times \\ O_B(b_f | b, y, y_f, \lambda_2)\, p_i(a,b|x,y) \times \\ \sum_{\lambda_1} p_{\Lambda_1}(\lambda_1) I_A(x|x_f, \lambda_1)\, I_B(y|y_f, \lambda_1) \left[x_f\, y_f\, a_f\, b_f \right]_{\bar{A}_f \bar{B}_f X_f Y_f} \otimes \rho_E^{a,b,x,y} \otimes [\lambda_1 \lambda_2]_{\Lambda_1 \Lambda_2}. \tag{6.58}$$

This in turn is a marginal of the following state:

$$\rho_{\bar{A}_f \bar{B}_f X_f Y_f E \Lambda_1 \Lambda_2 XY \bar{A} \bar{B}} = \sum_{x_f, y_f, a_f, b_f} p(x_f, y_f) \sum_{a,b,x,y} \sum_{\lambda_2} p_{\Lambda_2}(\lambda_2) O_A(a_f | a, x_f, x, \lambda_2) \\ O_B(b_f | b, y, y_f, \lambda_2) \times p_i(a,b|x,y) \sum_{\lambda_1} p_{\Lambda_1}(\lambda_1)\, I_A(x|x_f, \lambda_1)\, I_B(y|y_f, \lambda_1) \left[x_f\, y_f\, a_f\, b_f \right]_{X_f Y_f \bar{A}_f \bar{B}_f} \\ \otimes \rho_E^{a,b,x,y} \otimes [\lambda_1 \lambda_2]_{\Lambda_1 \Lambda_2} \otimes [x\, y\, a\, b]_{XY \bar{A} \bar{B}}. \tag{6.59}$$

Then, following arguments similar to those given in Proposition 65, we obtain $N^Q(\bar{A}_f; \bar{B}_f)_{p_f} \leq N^Q(\bar{A}; \bar{B})_{p_i}$. ■

6.2.2. Convexity

In this section, we prove that intrinsic non-locality and quantum intrinsic non-locality are convex. This statement physically means that Bell non-locality cannot increase when mixing two distributions.

Proposition 67 (Convexity of intrinsic non-locality) *Let $p(a,b|x,y)$ and $q(a,b|x,y)$ be two distributions, and let $\lambda \in [0,1]$. Let $t(a,b|x,y)$ be a mixture of the two distributions, defined as $t(a,b|x,y) = \lambda p(a,b|x,y) + (1-\lambda)\, q(a,b|x,y)$. Then*

$$N(\bar{A};\bar{B})_t \leq \lambda N(\bar{A};\bar{B})_p + (1-\lambda) N(\bar{A};\bar{B})_q. \tag{6.60}$$

Proof. First, we embed the distribution $t(a,b|x,y)$ in the following classical-classical state $\tau_{\bar{A}\bar{B}XY}$:

$$\tau_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,y)\, t(a,b|x,y) [x\, y\, a\, b]_{XY\bar{A}\bar{B}}, \tag{6.61}$$

where $p(x,y)$ is an arbitrary probability distribution. Similarly, embed $p(a,b|x,y)$ in $\rho_{\bar{A}\bar{B}XY}$ and $q(a,b|x,y)$ in $\gamma_{\bar{A}\bar{B}XY}$:

$$\rho_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,y)\, p(a,b|x,y)\, [x\, y\, a\, b]_{XY\bar{A}\bar{B}}\,, \tag{6.62}$$

$$\gamma_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,y)\, q(a,b|x,y)\, [x\, y\, a\, b]_{XY\bar{A}\bar{B}}\,. \tag{6.63}$$

Next, consider an arbitrary no-signaling extension of $\tau_{\bar{A}\bar{B}XY}$:

$$\tau_{\bar{A}\bar{B}XYE} := \sum_{x,y,a,b} p(x,y)\, t(a,b|x,y)\, [x\, y\, a\, b]_{XY\bar{A}\bar{B}} \otimes \tau_E^{a,b,x,y}. \tag{6.64}$$

Similarly, consider an arbitrary no-signaling extension of $\rho_{\bar{A}\bar{B}XY}$ and $\gamma_{\bar{A}\bar{B}XY}$:

$$\rho_{\bar{A}\bar{B}XYE} = \sum_{x,y,a,b} p(x,y)\, p(a,b|x,y)\, [x\, y\, a\, b]_{XY\bar{A}\bar{B}} \otimes \rho_E^{a,b,x,y}, \tag{6.65}$$

$$\gamma_{\bar{A}\bar{B}XYE} = \sum_{x,y,a,b} p(x,y)\, q(a,b|x,y)\, [x\, y\, a\, b]_{XY\bar{A}\bar{B}} \otimes \gamma_E^{a,b,x,y}. \tag{6.66}$$

Now, consider the following particular no-signaling extension of $\tau_{\bar{A}\bar{B}XY}$:

$$\zeta_{\bar{A}\bar{B}XYEE'} := \sum_{x,y,a,b} p(x,y)\, [x\, y]_{XY} \otimes \Big(\lambda\, p(a,b|x,y) \rho_E^{a,b,x,y} \otimes [0]_{E'} + (1-\lambda)\, q(a,b|x,y) \gamma_E^{a,b,x,y} \otimes [1]_{E'} \Big). \tag{6.67}$$

Then,

$$\inf_{\text{ext. in (6.64)}} I(\bar{A};\bar{B}|XYE)_\tau \leq I(\bar{A};\bar{B}|XYEE')_\zeta \tag{6.68}$$

$$= \lambda I(\bar{A};\bar{B}|XYE)_\rho + (1-\lambda) I(\bar{A};\bar{B}|XYE)_\gamma. \tag{6.69}$$

The first inequality follows from choosing a particular no-signaling extension. The equality follows from properties of conditional mutual information. Since this holds for all non-signaling extensions of the form in (6.65) and (6.66), we conclude that

$$\inf_{\text{ext. in (6.64)}} I(\bar{A};\bar{B}|XYE)_\zeta \leq \lambda \inf_{\text{ext. in (6.65)}} I(\bar{A};\bar{B}|XYE)_\rho + (1-\lambda) \inf_{\text{ext. in (6.66)}} I(\bar{A};\bar{B}|XYE)_\gamma. \tag{6.70}$$

Taking the supremum over all measurement choices, we find that

$$\sup_{p(x,y)} \inf_{\text{ext. in (6.64)}} I(\bar{A};\bar{B}|XYE)_{\zeta} \leq \lambda \sup_{p(x,y)} \inf_{\text{ext. in (6.65)}} I(\bar{A};\bar{B}|XYE)_{\rho} + \\ (1-\lambda) \sup_{p(x,y)} \inf_{\text{ext. in (6.66)}} I(\bar{A};\bar{B}|XYE)_{\gamma}. \quad (6.71)$$

This completes the proof. ■

In this proposition, we prove the convexity of quantum intrinsic non-locality. The proof is similar to Proposition 67, with the difference being in the choice of the extension system.

Proposition 68 (Convexity of quantum intrinsic non-locality) *Let $p(a,b|x,y)$ and $q(a,b|x,y)$ be distributions in $\mathbf{Q}$, and let $\lambda \in [0,1]$. Let $t(a,b|x,y)$ be a mixture of the distributions, defined as $t(a,b|x,y) = \lambda p(a,b|x,y) + (1-\lambda)\, q(a,b|x,y)$. Then*

$$N^{Q}(\bar{A};\bar{B})_t \leq \lambda N^{Q}(\bar{A};\bar{B})_p + (1-\lambda) N^{Q}(\bar{A};\bar{B})_q. \quad (6.72)$$

Proof. Since $\mathbf{Q}$ is a convex set [113], we know that $t(a,b|x,y) \in \mathbf{Q}$. First, we embed the distribution $t(a,b|x,y)$ in the following quantum state $\tau_{\bar{A}\bar{B}XY}$:

$$\tau_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,y)\, t(a,b|x,y) [x\, y\, a\, b]_{XY\bar{A}\bar{B}}, \quad (6.73)$$

where $p(x,y)$ is an arbitrary probability distribution. Similarly, embed $p(a,b|x,y)$ in $\rho_{\bar{A}\bar{B}XY}$ and $q(a,b|x,y)$ in $\gamma_{\bar{A}\bar{B}XY}$:

$$\rho_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,y)\, p(a,b|x,y)\, [x\, y\, a\, b]_{XY\bar{A}\bar{B}}\,, \quad (6.74)$$

$$\gamma_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,y)\, q(a,b|x,y)\, [x\, y\, a\, b]_{XY\bar{A}\bar{B}}\,. \quad (6.75)$$

Next, consider an arbitrary quantum extension of $\tau_{\bar{A}\bar{B}XY}$:

$$\tau_{\bar{A}\bar{B}XYE} := \sum_{x,y,a,b} p(x,y)\, t(a,b|x,y)\, [x\, y\, a\, b]_{XY\bar{A}\bar{B}} \otimes \tau_E^{a,b,x,y}. \tag{6.76}$$

Similarly, consider an arbitrary quantum extension of $\rho_{\bar{A}\bar{B}XY}$ and $\gamma_{\bar{A}\bar{B}XY}$:

$$\rho_{\bar{A}\bar{B}XYE} = \sum_{x,y,a,b} p(x,y)\, p(a,b|x,y)\, [x\, y\, a\, b]_{XY\bar{A}\bar{B}} \otimes \rho_E^{a,b,x,y}, \tag{6.77}$$

$$\gamma_{\bar{A}\bar{B}XYE} = \sum_{x,y,a,b} p(x,y)\, q(a,b|x,y)\, [x\, y\, a\, b]_{XY\bar{A}\bar{B}} \otimes \gamma_E^{a,b,x,y}. \tag{6.78}$$

Let ρ_{AB} be a quantum state that, along with the POVMs characterized by Λ_x^a and Λ_y^b, yield the distribution $p(a,b|x,y)$. Let ρ_{ABE} be an extension of ρ_{AB}. Similarly, let γ_{AB} be a quantum state that, along with the POVMs characterized by M_x^a and M_y^b, yield the distribution $q(a,b|x,y)$. Let γ_{ABE} be an extension of γ_{AB}. Then, a particular quantum state that realizes the distribution $t(a,b|x,y)$ is the following:

$$\tau_{ABA'B'} = \lambda\rho_{AB} \otimes |00\rangle\langle 00|_{A'B'} + (1-\lambda)\gamma_{AB} \otimes |11\rangle\langle 11|_{A'B'}\,, \tag{6.79}$$

$$t(a,b|x,y) = \operatorname{Tr}\left[\left(\Lambda_x^a \otimes \Lambda_y^b \otimes (|00\rangle\langle 00|_{A'B'}) + M_x^a \otimes M_y^b \otimes (|11\rangle\langle 11|_{A'B'})\right)(\tau_{ABA'B'})\right], \tag{6.80}$$

where it is understood that Alice is measuring σ_Z on her system A' and Bob is measuring σ_Z on B', in addition to the other measurements on their systems A and B. Now, consider the following extension of $\tau_{ABA'B'}$:

$$\tau_{ABA'B'EE'} = \lambda\rho_{ABE} \otimes |000\rangle\langle 000|_{A'B'E'} + (1-\lambda)\gamma_{ABE} \otimes |111\rangle\langle 111|_{A'B'E'}\,. \tag{6.81}$$

Furthermore, consider the following particular quantum extension of $\tau_{\bar{A}\bar{B}XY}$:

$$\zeta_{\bar{A}\bar{B}XYEE'} := \sum_{x,y,a,b} p(x,y)\,[x\,y]_{XY} \otimes \Big(\lambda\, p(a,b|x,y)\rho_E^{a,b,x,y} \otimes [0]_{E'} + (1-\lambda)\, q(a,b|x,y)\gamma_E^{a,b,x,y} \otimes [1]_{E'}\Big)\,. \tag{6.82}$$

Then following similar arguments given in the proof of Proposition 67, we obtain

$$N^Q(\bar{A};\bar{B})_t \leq \lambda N^Q(\bar{A};\bar{B})_p + (1-\lambda) N^Q(\bar{A};\bar{B})_q, \tag{6.83}$$

concluding the proof. ■

6.2.3. Superadditivity and additivity

In this section, we prove that the quantifiers are supperadditive and additive under independent distributions.

Proposition 69 (Superadditivity and additivity of intrinsic non-locality) *Let $p(a_1,a_2,b_1,b_2|x_1,x_2,y_1,y_2)$ be a distribution for which the following no-signaling constraints hold:*

$$\sum_{a_1} p(a_1,a_2,b_1,b_2|x_1,x_2,y_1,y_2) = \sum_{a_1} p(a_1,a_2,b_1,b_2|x_1',x_2,y_1,y_2)$$
$$\forall x_1',x_1,x_2,y_1,y_2 \in [s],\; a_2,b_1,b_2 \in [r],$$

$$\sum_{a_2} p(a_1,a_2,b_1,b_2|x_1,x_2,y_1,y_2) = \sum_{a_2} p(a_1,a_2,b_1,b_2|x_1,x_2',y_1,y_2)$$
$$\forall x_2',x_2,x_1,y_1,y_2 \in [s],\; a_1,b_1,b_2 \in [r],$$

$$\sum_{b_1} p(a_1,a_2,b_1,b_2|x_1,x_2,y_1,y_2) = \sum_{b_1} p(a_1,a_2,b_1,b_2|x_1,x_2,y_1',y_2)$$
$$\forall y_1',y_1,x_1,x_2,y_2 \in [s],\; a_1,a_2,b_2 \in [r],$$

$$\sum_{b_2} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) = \sum_{b_2} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2')$$
$$\forall y_2', y_2, x_2, y_1, x_1 \in [s],\ a_1, a_2, b_1 \in [r].$$

Let $t(a_1, b_1|x_1, y_1)$ and $r(a_2, b_2|x_2, y_2)$ be distributions corresponding to the marginals of $p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2)$. Then the intrinsic non-locality is super-additive, in the sense that

$$N(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2)_p \geq N(\bar{A}_1; \bar{B}_1)_t + N(\bar{A}_2; \bar{B}_2)_r. \tag{6.84}$$

If $p(a_1, b_1, a_2, b_2|x_1, x_2, y_1, y_2) = t(a_1, b_1|x_1, y_1)r(a_2, b_2|x_2, y_2)$, then the intrinsic non-locality is additive in the following sense:

$$N(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2)_p = N(\bar{A}_1; \bar{B}_1)_t + N(\bar{A}_2; \bar{B}_2)_r. \tag{6.85}$$

Proof. Consider the classical-classical state $\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1Y_1X_2Y_2}$ with the following arbitrary no-signaling extension:

$$\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2E} = \sum_{x_1,x_2,y_1,y_2,a_1,a_2,b_1,b_2} p(x_1, y_1, x_2, y_2)\, p(a_1, b_1, a_2, b_2|x_1, x_2, y_1, y_2)$$
$$[a_1\, b_1\, x_1\, y_1\, a_2\, b_2\, x_2\, y_2]_{\bar{A}_1\bar{B}_1X_1Y_1\bar{A}_2\bar{B}_2X_2Y_2} \otimes \rho_E^{a_1,b_1,x_1,y_1,a_2,b_2,x_2,y_2}, \tag{6.86}$$

where $p(x_1, x_2, y_1, y_2)$ is an arbitrary probability distribution. From the chain rule of mutual information and non-negativity of conditional mutual information, we obtain

$$I(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2|X_1X_2Y_1Y_2E)_\rho$$
$$= I(\bar{A}_1\bar{A}_2; \bar{B}_1|X_1Y_1X_2Y_2E) + I(\bar{A}_1\bar{A}_2; \bar{B}_2|EX_1Y_1X_2Y_2\bar{B}_1) \tag{6.87}$$

$$
\begin{aligned}
&= I(\bar{A}_1;\bar{B}_1|X_1Y_1X_2Y_2E)_\rho + I(\bar{A}_2;\bar{B}_1|EX_1Y_1X_2Y_2\bar{A}_1)_\rho \\
&\qquad + I(\bar{A}_1;\bar{B}_2|X_1Y_1X_2Y_2E\bar{B}_1) + I(\bar{A}_2;\bar{B}_2|X_1Y_1X_2Y_2E\bar{A}_1\bar{B}_1) \qquad (6.88) \\
&\geq I(\bar{A}_1;\bar{B}_1|X_1Y_1X_2Y_2E)_\rho + I(\bar{A}_2;\bar{B}_2|X_1Y_1X_2Y_2E\bar{A}_1\bar{B}_1)_\rho. \qquad (6.89)
\end{aligned}
$$

From the no-signaling constraints in the statement of the proposition and (6.86), we obtain

$$
\begin{aligned}
\rho_{\bar{A}_1\bar{B}_1X_1X_2Y_1Y_2E} = \sum_{a_1,b_1,x_1,x_2,y_1,y_2} & p(x_1,x_2,y_1,y_2)\,[a_1\,b_1\,x_1\,y_1\,x_2\,y_2]_{\bar{A}_1\bar{B}_2X_1Y_1X_2Y_2} \\
& \otimes p(a_1,b_1|x_1,y_1)\,\rho_E^{x_1,y_1,a_1,b_1}. \qquad (6.90)
\end{aligned}
$$

We first embed $t(a_1,b_1|x_1,y_1)$ in $\tau_{\bar{A}_1\bar{B}_1X_1Y_1E}$, and $r(a_2,b_2|x_2,y_2)$ in $\gamma_{\bar{A}_2\bar{B}_2X_2Y_2E}$ and consider the following arbitrary no-signaling extensions:

$$
\tau_{\bar{A}_1\bar{B}_1X_1Y_1E} := \sum_{x_1,y_1} p(x_1,y_1) \otimes \sum_{a_1,b_1} [x_1\,y_1\,a_1\,b_1]_{X_1Y_1\bar{A}_1\bar{B}_1} \otimes t(a_1,b_1|x_1,y_1)\tau_E^{a_1,b_1,x_1,y_1}, \qquad (6.91)
$$

$$
\gamma_{\bar{A}_2\bar{B}_2X_2Y_2E} := \sum_{x_2,y_2} p(x_2,y_2) \otimes \sum_{a_2,b_2} [x_2\,y_2\,a_2\,b_2]_{X_2Y_2\bar{A}_2\bar{B}_2} \otimes r(a_2,b_2|x_2,y_2)\gamma_E^{a_2,b_2,x_2,y_2}. \qquad (6.92)
$$

Since $\rho_{\bar{A}_1\bar{B}_1X_1Y_1X_2Y_2E}$ is a particular no-signaling extension of $\tau_{\bar{A}_1\bar{B}_1X_1Y_1}$ and $\rho_{\bar{A}_1\bar{B}_1\bar{A}_2\bar{B}_2X_1Y_1X_2Y_2E}$ is a particular no-signaling extension of $\gamma_{\bar{A}_2\bar{B}_2X_2Y_2}$, we obtain the following inequality:

$$
\begin{aligned}
& I(\bar{A}_1\bar{A}_2;\bar{B}_1\bar{B}_2|X_1X_2Y_1Y_2E)_\rho \\
&\quad \geq I(\bar{A}_1;\bar{B}_1|X_1Y_1X_2Y_2E)_\rho + I(\bar{A}_2;\bar{B}_2|X_1Y_1X_2Y_2E\bar{A}_1\bar{B}_1)_\rho \qquad (6.93)
\end{aligned}
$$

$$\geq \inf_{\text{ext. in (6.91)}} I(\bar{A}_1;\bar{B}_1|X_1Y_1E)_\tau + \inf_{\text{ext. in (6.92)}} I(\bar{A}_2;\bar{B}_2|X_2Y_2E\bar{A}_1\bar{B}_1)_\gamma. \tag{6.94}$$

Since (6.94) holds for an arbitrary no-signaling extension of ρ, we obtain

$$\inf_{\text{ext. in (6.86)}} I(\bar{A}_1\bar{A}_2;\bar{B}_1\bar{B}_2|X_1X_2Y_1Y_2E)_\rho \geq$$
$$\inf_{\text{ext. in (6.91)}} I(\bar{A}_1;\bar{B}_1|X_1Y_1E)_\tau + \inf_{\text{ext. in (6.92)}} I(\bar{A}_2;\bar{B}_2|X_2Y_2E\bar{A}_1\bar{B}_1)_\gamma \tag{6.95}$$

Since the above equation holds for arbitrary probability distributions, we can take a supremum over all probability distributions to obtain

$$\sup_{p(x_1,y_1)p(x_2,y_2)} \inf_{\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2E}} I(\bar{A}_1\bar{A}_2;\bar{B}_1\bar{B}_2|X_1X_2Y_1Y_2E)_\rho \geq$$
$$\sup_{p(x_1,y_1)} \inf_{\tau_{\bar{A}_1\bar{B}_1X_1Y_1E}} I(\bar{A}_1;\bar{B}_1|X_1Y_1E)_\tau + \sup_{p(x_2,y_2)} \inf_{\gamma_{\bar{A}_2\bar{B}_2X_2Y_2E}} I(\bar{A}_2;\bar{B}_2|X_2Y_2E)_\gamma. \tag{6.96}$$

Since we have considered a supremum over product probability distributions for the measurement choices on the LHS, we can relax this to consider the supremum over all probability distributions $p(x_1,y_1,x_2,y_2)$ of the measurement choices. This concludes the proof of (6.84).

Now we give a proof for additivity of intrinsic non-locality with respect to product probability distributions. Since intrinsic non-locality is super-additive, it is sufficient to prove the following sub-additivity property for product probability distributions:

$$N(\bar{A}_1\bar{A}_2;\bar{B}_1\bar{B}_2)_p \leq N(\bar{A}_1;\bar{B}_1)_t + N(\bar{A}_2;\bar{B}_2)_r. \tag{6.97}$$

Consider the following states

$$\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2} = \sum_{a_1,b_1,x_1,y_1,a_2,b_2,x_2,y_2} p(x_1,x_2,y_1,y_2)\, t(a_1,b_1|x_1,y_1)\, r(a_2,b_2|x_2,y_2) \\ [a_1\, b_1\, a_2\, b_2\, x_1\, x_2\, y_1\, y_2]_{\bar{A}_1\bar{B}_1\bar{A}_2\bar{B}_2X_1Y_1X_2Y_2}\,. \quad (6.98)$$

Consider an arbitrary extension of the state $\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2}$

$$\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2E} := \sum_{a_1,b_1,x_1,y_1,a_2,b_2,x_2,y_2} p(x_1,x_2,y_1,y_2)\, t(a_1,b_1|x_1,y_1)\, r(a_2,b_2|x_2,y_2) \\ [a_1\, b_1\, x_1\, y_1\, a_2\, b_2\, x_2\, y_2] \otimes \rho_E^{a_1,b_1,x_1,y_1,a_2,b_2,x_2,y_2}. \quad (6.99)$$

Now, consider a particular extension of the state $\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2}$:

$$\zeta_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2E_1E_2} := \sum_{a_1,b_1,x_1,y_1,a_2,b_2,x_2,y_2} p(x_1,x_2,y_1,y_2)\, t(a_1,b_1|x_1,y_1)\, r(a_2,b_2|x_2,y_2) \\ [a_1b_1a_2b_2x_1x_2y_1y_2]_{\bar{A}_1\bar{B}_1X_1Y_1\bar{A}_2\bar{B}_2X_2Y_2} \otimes \rho_{E_1}^{a_1,b_1,x_1,y_1} \otimes \rho_{E_2}^{a_2,b_2,x_2,y_2}. \quad (6.100)$$

Then, we have the following set of inequalities:

$$\inf_{\text{ext. in (6.99)}} I(\bar{A}_1\bar{A}_2;\bar{B}_1\bar{B}_2|X_1Y_1X_2Y_2E)_\rho$$

$$\leq I(\bar{A}_1\bar{A}_2;\bar{B}_1\bar{B}_2|X_1Y_1X_2Y_2E_1E_2)_\zeta \quad (6.101)$$

$$= I(\bar{A}_1;\bar{B}_1|X_1Y_1X_2Y_2E_1E_2)_\zeta + I(\bar{A}_2;\bar{B}_1|E_1E_2X_1Y_1X_2Y_2\bar{A}_1)_\zeta \\ + I(\bar{A}_1;\bar{B}_2|X_1Y_1X_2Y_2E_1E_2\bar{B}_1)_\zeta + I(\bar{A}_2;\bar{B}_2|X_1Y_1X_2Y_2E_1E_2\bar{A}_1\bar{B}_1)_\zeta \quad (6.102)$$

$$= I(\bar{A}_1;\bar{B}_1|X_1Y_1X_2Y_2E_1E_2)_\zeta + I(\bar{A}_2;\bar{B}_2|X_1Y_1X_2Y_2E_1E_2\bar{A}_1\bar{B}_1)_\zeta. \quad (6.103)$$

The first inequality follows from a particular choice of an extension. The first equality follows from the chain rule. For the second equality, observe the following:

$$
\begin{aligned}
& I(\bar{A}_2;\bar{B}_1|E_1E_2X_1Y_1X_2Y_2\bar{A}_1)_\zeta \\
& = H(\bar{A}_2|E_1E_2X_1Y_1X_2Y_2\bar{A}_1)_\zeta - H(\bar{A}_2|E_1E_2X_1Y_1X_2Y_2\bar{A}_1\bar{B}_1)_\zeta \qquad (6.104) \\
& = \sum_{x_1x_2y_1y_2} p(x_1,x_2,y_1,y_2)\left[H(\bar{A}_2|\bar{A}_1E_1E_2)_{\zeta^{x_1x_2y_1y_2}} - H(\bar{A}_2|\bar{A}_1E_1E_2\bar{B}_1)_{\zeta^{x_1x_2y_1y_2}}\right], \qquad (6.105)
\end{aligned}
$$

where

$$
\zeta^{x_1,x_2,y_1,y_2}_{\bar{A}_1\bar{A}_2E_1E_2} = \sum_{a_1} t(a_1|x_1)\,[a_1]_{\bar{A}_1}\otimes\rho^{a_1,x_1}_{E_1}\otimes\sum_{a_2} r(a_2|x_2)\,[a_2]_{\bar{A}_2}\otimes\rho^{a_2,x_2}_{E_2}, \qquad (6.106)
$$

$$
\zeta^{x_1,x_2,y_1,y_2}_{\bar{A}_1\bar{A}_2\bar{B}_1E_1E_2} = \sum_{a_1,b_1} t(a_1,b_1|x_1,y_1)\,[a_1b_1]_{\bar{A}_1\bar{B}_1}\otimes\rho^{a_1,x_1,b_1,y_1}_{E_1}\otimes\sum_{a_2} r(a_2|x_2)\,[a_2]_{\bar{A}_2}\otimes\rho^{a_2,x_2}_{E_2}. \qquad (6.107)
$$

Then, from (6.106) and (6.107), it follows that

$$
H(\bar{A}_2|\bar{A}_1E_1E_2)_{\zeta^{x_1x_2y_2y_2}} = H(\bar{A}_2|E_2)_{\zeta^{x_1x_2y_2y_2}}, \qquad (6.108)
$$

$$
H(\bar{A}_2|\bar{A}_1E_1E_2\bar{B}_1)_{\zeta^{x_1x_2y_2y_2}} = H(\bar{A}_2|E_2)_{\zeta^{x_1x_2y_2y_2}}. \qquad (6.109)
$$

This is equivalent to $I(\bar{A}_2;\bar{B}_1|E_1E_2X_1Y_1X_2Y_2\bar{A}_1)_\zeta = 0$.

Similarly, $I(\bar{A}_1;\bar{B}_2|E_1E_2X_1Y_1X_2Y_2\bar{B}_1)_\zeta = 0$. Then by inspection of (6.100), and from the no-signaling constraints, it follows that

$$
\inf_{\text{ext. in (6.99)}} I(\bar{A}_1\bar{A}_2;\bar{B}_1\bar{B}_2|X_1Y_1X_2Y_2E)_\rho \le I(\bar{A}_1;\bar{B}_1|X_1Y_1E_1)_\zeta + I(\bar{A}_2;\bar{B}_2|X_2Y_2E_2)_\zeta. \qquad (6.110)
$$

Since the above statement holds for an arbitrary no-signaling extension of the form in (6.99), it follows that

$$\inf_{\text{ext. in (6.99)}} I(\bar{A}_1\bar{A}_2;\bar{B}_1\bar{B}_2|X_1Y_1X_2Y_2E)_\rho$$
$$\leq \inf_{\text{ext. in (6.100)}} I(\bar{A}_1;\bar{B}_1|X_1Y_1E_1)_\zeta + \inf_{\text{ext. in (6.100)}} I(\bar{A}_2;\bar{B}_2|X_2Y_2E_2)_\zeta. \quad (6.111)$$

Since the above inequality holds for an arbitrary probability distribution $p(x_1, x_2, y_1, y_2)$, we find that

$$\sup_{p(x_1,x_2,y_1,y_2)} \inf_{\text{ext. in (6.99)}} I(\bar{A}_1\bar{A}_2;\bar{B}_1\bar{B}_2|X_1Y_1X_2Y_2E)_\rho$$
$$\leq \sup_{p(x_1,y_1)} \inf_{\text{ext. in (6.100)}} I(\bar{A}_1;\bar{B}_1|X_1Y_1E_1)_\zeta + \sup_{p(x_2,y_2)} \inf_{\text{ext. in (6.100)}} I(\bar{A}_2;\bar{B}_2|X_2Y_2E_2)_\zeta.$$
$$(6.112)$$

This conclues the proof. ■

Proposition 70 (Superadditivity and additivity of quantum intrinsic non-locality) *Let $p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2)$ be a quantum distribution that arises from a four-party state $\rho_{A_1A_2B_1B_2}$, and POVMs characterized by $\Lambda^{a_1}_{x_1}, \Lambda^{a_2}_{x_2}, \Lambda^{b_1}_{y_1}$, and $\Lambda^{b_2}_{y_2}$. Then the following no-signaling constraints hold:*

$$\sum_{a_1} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) = \sum_{a_1} p(a_1, a_2, b_1, b_2|x'_1, x_2, y_1, y_2)$$
$$\forall x'_1, x_1, x_2, y_1, y_2 \in [s],\ a_2, b_1, b_2 \in [r]$$
$$\sum_{a_2} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) = \sum_{a_2} p(a_1, a_2, b_1, b_2|x_1, x'_2, y_1, y_2)$$
$$\forall x'_2, x_2, x_1, y_1, y_2 \in [s],\ a_1, b_1, b_2 \in [r]$$
$$\sum_{b_1} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) = \sum_{b_1} p(a_1, a_2, b_1, b_2|x_1, x_2, y'_1, y_2)$$

$$\forall y_1', y_1, x_1, x_2, y_2 \in [s],\ a_1, a_2, b_2 \in [r]$$

$$\sum_{b_2} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) = \sum_{b_2} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2')$$

$$\forall y_2', y_2, x_2, y_1, x_1 \in [s],\ a_1, a_2, b_1 \in [r].$$

Let $t(a_1, b_1|x_1, y_1)$ and $r(a_2, b_2|x_2, y_2)$ be quantum distributions corresponding to the marginals of $p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2)$. Then the quantum intrinsic non-locality is super-additive, in the sense that

$$N^Q(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2)_p \geq N^Q(\bar{A}_1; \bar{B}_1)_t + N^Q(\bar{A}_2; \bar{B}_2)_r. \tag{6.113}$$

If $p(a_1, b_1, a_2, b_2|x_1, x_2, y_1, y_2) = t(a_1, b_1|x_1, y_1)r(a_2, b_2|x_2, y_2)$, then the quantum intrinsic non-locality is additive in the following sense:

$$N^Q(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2)_p = N^Q(\bar{A}_1; \bar{B}_1)_t + N^Q(\bar{A}_2; \bar{B}_2)_r. \tag{6.114}$$

Proof. The proof follows by using similar techniques as in the proof of Proposition 69, and by taking appropriate quantum extensions. ■

6.2.4. Quantum intrinsic non-locality and intrinsic steerability

Let ρ_{AB} be a quantum state, and let $p_{\bar{A}|X}\rho_B^{a,x}$ be an assemblage that arises from the quantum state ρ_{AB} and some measurement $\{\Lambda_a^x\}$.[2] We then prove that the intrinsic steerability of the assemblage $p_{\bar{A}|X}\rho_B^{a,x}$ is never smaller than the quantum intrinsic non-locality of all the bipartite distributions that can arise from this assemblage.

Proposition 71 *Let $p(a, b|x, y)$ be a quantum distribution that is obtained by performing a POVM $\left\{\Lambda_y^b\right\}_b$ on the assemblage $\{p_{\bar{A}|X}(a|x)\rho_B^{a,x}\}_{a,x}$. Then the quantum*

[2] From [29], it can be seen that given a bipartite assemblage, we can always find an underlying quantum state and measurements.

intrinsic non-locality of the distribution p *does not exceed the restricted intrinsic steerability of the assemblage* $\hat{\rho}$*. That is,*

$$N^Q(\bar{A};\bar{B})_p \leq S^R(\bar{A};B)_{\hat{\rho}}, \tag{6.115}$$

where we recall that $\hat{\rho}$ *is a shorthand to denote the assemblage.*

Proof. Let $p(a,b|x,y)$ be a quantum distribution that arises from the assemblage $p_{\bar{A}|X}(a|x)\rho_B^{a,x}$. That is,

$$p(a,b|x,y) = \operatorname{Tr}\left[\Lambda_y^b\left(p_{\bar{A}|X}(a|x)\rho_B^{a,x}\right)\right]. \tag{6.116}$$

Let $p_{\bar{A}|X}(a|x)\rho_{BE}^{a,x}$ be a particular no-signaling extension of $p_{\bar{A}|X}(a|x)\rho_B^{a,x}$. Then one possible no-signaling extension of $p(a,b|x,y)$ is

$$p(a,b|x,y)\rho_E^{a,x,b,y} = \operatorname{Tr}_B\left[\Lambda_y^b\left(p_{\bar{A}|X}(a|x)\rho_{BE}^{a,x}\right)\right]. \tag{6.117}$$

From [29], it follows that the above is also a quantum extension.

Let $p(x,y)$ be an arbitrary probability distribution. Let $p(a,b|x,y)$ be a distribution embedded in a classical-classical state $\rho_{\bar{A}\bar{B}XY}$ with the following particular no-signaling extension:

$$\rho_{\bar{A}\bar{B}XYE} := \sum_{a,b,x,y} p(x,y)p(a,b|x,y)\,[a\,b\,x\,y]_{\bar{A}\bar{B}XY} \otimes \rho_E^{a,b,x,y}, \tag{6.118}$$

and an arbitrary quantum extension:

$$\sigma_{\bar{A}\bar{B}XYE} := \sum_{a,b,x,y} p(x,y)p(a,b|x,y)\,[a\,b\,x\,y]_{\bar{A}\bar{B}XY} \otimes \sigma_E^{a,b,x,y}. \tag{6.119}$$

Similarly, let $\rho_{\bar{A}XB}$ be a state into which the assemblage $p_{\bar{A}|X}(a|x)\rho_B^{a,x}$ is embedded, and let $\rho_{\bar{A}XBE}$ be a particular extension, where

$$\rho_{\bar{A}BXE} = \sum_{a,x} p(x)p_{\bar{A}|X}(a|x)[a\,x]_{\bar{A}X} \otimes \rho_{BE}^{a,x}. \tag{6.120}$$

Let

$$\rho_{\bar{A}BXYE} = \sum_{a,x} p(x,y)p_{\bar{A}|X}(a|x)[a\,x]_{\bar{A}X} \otimes \rho_{BE}^{a,x}. \tag{6.121}$$

Then,

$$I(\bar{A};B|XE)_\rho = I(\bar{A};BY|XE)_\rho. \tag{6.122}$$

This follows from chain rule of conditional mutual information and inspection of (6.121). Observe that Bob can perform a local operation and transform the state $\rho_{\bar{A}BXYE}$ to $\rho_{\bar{A}\bar{B}XYE}$. Then, from the data-processing inequality, we find that

$$I(\bar{A};B|XE)_\rho \geq I(\bar{A};\bar{B}Y|XE)_\rho. \tag{6.123}$$

This means that for every no-signaling extension $\rho_{\bar{A}BXE}$ of the state $\rho_{\bar{A}BX}$ that encodes the assemblage $\rho_{\bar{A}|X}(a|x)\rho^{a,x}$, we can find a quantum extension $\rho_{\bar{A}\bar{B}XYE}$ of $\rho_{\bar{A}\bar{B}XY}$ that encodes the distribution $p(a,b|x,y)$ derived from the assemblage $p_{\bar{A}|X}(a|x)$ $\rho^{a,x}$, such that (6.123) is true. Therefore, we obtain the following:

$$\inf_{\text{ext in (6.120)}} I(\bar{A};B|XE)_\rho \geq \inf_{\text{ext. in (6.118)}} I(\bar{A};\bar{B}Y|XE)_\rho \tag{6.124}$$

$$\geq \inf_{\text{ext in (6.119)}} I(\bar{A};\bar{B}Y|XE)_\sigma. \tag{6.125}$$

This in turn implies that

$$S^R(\bar{A};B)_{\hat{\rho}} \geq N^Q(\bar{A};\bar{B})_p, \tag{6.126}$$

concluding the proof. ■

6.2.5. Faithfulness

Proposition 72 *Intrinsic non-locality and quantum intrinsic non-locality vanish for distributions having a local hidden-variable model; i.e., if $p(a,b|x,y) \in \mathbf{L}$, then $N(\bar{A};\bar{B})_p = 0$ and $N^Q(\bar{A};\bar{B})_p = 0$.*

Proof. Given $p(a,b|x,y) \in \mathbf{L}$, then we can write it as

$$p(a,b|x,y) = \sum_{\lambda} p(\lambda)\, p(a|x,\lambda)\, p(b|y,\lambda). \tag{6.127}$$

Embed this in a classical-classical state with $p(x,y)$ an arbitrary probability distribution over x,y:

$$\rho_{\bar{A}\bar{B}XY} = \sum_{a,b,x,y} p(x,y) \sum_{\lambda} p(\lambda)\, p(a|x,\lambda)\, p(b|y,\lambda)\, [a\, b\, x\, y]_{\bar{A}\bar{B}XY}\,. \tag{6.128}$$

Then, consider the following quantum extension

$$\rho_{\bar{A}\bar{B}XYE} := \sum_{a,b,x,y} p(x,y)\, [a\, b\, x\, y]_{\bar{A}\bar{B}XY} \otimes \sum_{\lambda} p(\lambda)\, p(a|x,\lambda)\, p(b|y,\lambda)\, [\lambda]_E. \tag{6.129}$$

Then, by inspection, $\bar{A}$ and $\bar{B}$ are independent given XYE. This implies that $\inf_{\rho_{\bar{A}\bar{B}XYE}} I(\bar{A};\bar{B}|XYE)_\rho = 0$. Since this equality holds for an arbitrary probability distribution $p(x,y)$, we can then conclude that $N^Q(\bar{A};\bar{B})_p = 0$. Then, by (6.13) we conclude that $N(\bar{A};\bar{B})_p = 0$. ■

We now state below in Theorem 73 that $N(\bar{A};\bar{B})_p = 0$ and $N^Q(\bar{A};\bar{B})_p = 0$ implies that $p \in \mathbf{L}$.

Theorem 73 (Faithfulness of intrinsic non-locality) *For every no-signaling or quantum distribution $p(a,b|x,y)$, the intrinsic non-locality $N(\bar{A};\bar{B})_p = 0$, if and only if it has a local hidden variable description. Quantitatively, if $N(\bar{A};\bar{B})_p \le \varepsilon$, where $0 < \varepsilon^{1/16}d^{1/2} < 1$, for $d = |\mathcal{X}| \cdot |\mathcal{Y}|$, then there exists a probability distribution $l(a,b|x,y)$ having a local hidden-variable description, such that*

$$\sup_{p_{XY}(x,y)} \|\rho_{\bar{A}X\bar{B}Y} - \gamma_{\bar{A}X\bar{B}Y}\|_1 \le d\left(\varepsilon^{1/4} + \frac{\varepsilon^{1/16}d^{1/2}}{1-\varepsilon^{1/16}d^{1/2}} + 4de^{-\frac{\varepsilon^{-1/4}}{3}}\right), \tag{6.130}$$

where $\rho_{\bar{A}X\bar{B}Y}$ correponds to the classical-classical state $p_{XY}(x,y)p(a,b|x,y)$ and $\gamma_{\bar{A}X\bar{B}Y}$ is the classical-classical state corresponding to $p_{XY}(x,y)l(a,b|x,y)$.

Proof. The proof closely follows the proof for faithfulness of restricted intrinsic steerability, given in the proof of Theorem 58. We first construct a strategy for $p_{XY}(x,y) = \frac{1}{|\mathcal{X}|}.\frac{1}{|\mathcal{Y}|}$ and then generalize it to an arbitrary distribution. Invoking [106], we know that there exists a recovery channel $\mathcal{R}_{XE\to\bar{A}XE}$ such that

$$\|\rho_{\bar{A}X\bar{B}YE} - \mathcal{R}_{XE\to\bar{A}XE}(\rho_{\bar{B}YE}\otimes\rho_X)\|_1 \le \sqrt{I(\bar{A};\bar{B}Y|XE)_\rho \ln 2} = t. \tag{6.131}$$

Since $I(\bar{B}E;X|Y)_\rho = 0$ from (6.4), and $p_{XY}(x,y) = \frac{1}{\mathcal{X}}.\frac{1}{\mathcal{Y}}$, we can write $\rho_{\bar{B}XYE} = \rho_{\bar{B}YE}\otimes\rho_X$. Following an argument similar to (5.173)–(5.176), we obtain the following inequality:

$$\|\rho_{\bar{A}\bar{B}XY} - \omega_{A_iX_iBY}\| \le nt, \tag{6.132}$$

where

$$\omega_{\bar{A}^nX^n\bar{B}YE} = \bigcirc_{i=1}^n \mathcal{R}_{X_iE\to\bar{A}_iX_iE}\left(\rho_{\bar{B}YE}\otimes\rho_X^{\otimes n}\right), \tag{6.133}$$

$$\omega_{\bar{A}_i X_i B Y} = \mathrm{Tr}_{E \bar{A}^{n/i} X^{n/i}} \left(\omega_{\bar{A}^n X^n \bar{B} Y E} \right). \tag{6.134}$$

Since the distributions $p_X(x)$ and $p_Y(y)$ are independent, we have

$$I(X^n; Y)_p = 0. \tag{6.135}$$

From the no-signaling constraints, we have

$$I(X^n Y; E)_\rho = 0. \tag{6.136}$$

This implies that

$$I(X^n E; Y)_\rho = I(X^n; Y)_\rho + I(E; Y | X^n)_\rho = 0. \tag{6.137}$$

Since the systems $\bar{A}^n X^n E$ of $\omega_{\bar{A}^n X^n \bar{B} Y E}$ are obtained from the application of the recovery channel on systems $X_n E$ of the state $\rho_{X_n Y E \bar{B}}$, we can use quantum data processing for mutual information to obtain the following inequality:

$$I(A^n X^n; Y)_\omega = 0. \tag{6.138}$$

This implies that

$$\omega_{\bar{A}^n X^n \bar{B} Y} = \sum_{x^n, a^n, y, b} p(x^n)\, q(a^n | x^n)\, p(y)\, q(b | a^n x^n y)\, [x^n\, a^n\, b\, y]_{X^n \bar{A}^n \bar{B} Y}. \tag{6.139}$$

Alice's strategy is exactly the same as before, and the following state is obtained after the application of the algorithm in (5.191):

$$\gamma_{\tilde{A}\tilde{X}\bar{B}Y} := \sum_{\tilde{x},\tilde{a},b,y} p_X(\tilde{x})\, |\tilde{x}\rangle\!\langle\tilde{x}|_{\tilde{X}} \otimes \sum_{x^n,a^n,} p_{\tilde{A}|\tilde{X}X^nA^n}(\tilde{a}|\tilde{x},x^n,a^n) p_{X^n}(x^n)$$
$$q_{A^n|X^n}(a^n|x^n)p(y)q(b|a^nx^ny)[\tilde{a}\,b\,y]_{\tilde{A}\bar{B}Y}. \quad (6.140)$$

Note that this state is a local-hidden-variable state. This construction of the local-hidden-variable state shares some similarities with [108]. By following the arguments given for the proof of faithfulness of intrinsic steerability, we obtain

$$\|\rho_{\bar{A}X\bar{B}Y} - \gamma_{\tilde{A}\tilde{X}\bar{B}Y}\|_1 \leq nt + \frac{\delta}{1-\delta} + 2\varepsilon_1. \quad (6.141)$$

This implies

$$\|\rho_{\bar{A}X\bar{B}Y} - \gamma_{\tilde{A}\tilde{X}\bar{B}Y}\|_1 \leq n \inf_{\rho_{\bar{A}X\bar{B}YE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1. \quad (6.142)$$

This implies

$$\sum_{a,b} |p(a,b|x,y) - l(a,b|x,y)| \leq |\mathcal{X}||\mathcal{Y}| \left(\inf_{\rho_{\bar{A}X\bar{B}YE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1 \right) \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}. \quad (6.143)$$

Now, using triangle inequality, we obtain the following for any arbitrary distribution $p(x,y)$:

$$\|\rho_{\bar{A}X\bar{B}Y} - \gamma_{\tilde{A}\tilde{X}\bar{B}Y}\|_1 \leq |\mathcal{X}||\mathcal{Y}| \left(\inf_{\rho_{\bar{A}X\bar{B}YE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1 \right). \quad (6.144)$$

This implies

$$\sup_{p_{XY}(x,y)} \|\rho_{\bar{A}X\bar{B}Y} - \gamma_{\tilde{A}\tilde{X}\bar{B}Y}\|_1 \leq |\mathcal{X}||\mathcal{Y}| \left(\sqrt{N(\bar{A};\bar{B})_p \ln 2} + \frac{\delta}{1-\delta} + 2\varepsilon_1 \right). \quad (6.145)$$

Given $N(\bar{A};\bar{B})_p \leq \varepsilon$ (as required by the condition of faithfulness), choose $n = (1/\varepsilon)^{1/4}$, $\delta = \varepsilon^{1/16}|\mathcal{X}|^{1/2}|\mathcal{Y}|^{1/2}$. This proof holds only if $\delta \in (0,1)$. We know by the Chernoff bound [109] that $\varepsilon_1 = 2|\mathcal{X}||\mathcal{Y}|e^{-\frac{1}{3|\mathcal{X}|\cdot|\mathcal{Y}|}\delta^2 n}$. Substituting these values, we obtain

$$\|\rho_{\bar{A}X\bar{B}Y} - \gamma_{\tilde{A}\tilde{X}\bar{B}Y}\|_1 \leq |\mathcal{X}| \cdot |\mathcal{Y}| \left(\varepsilon^{1/4} + \frac{\varepsilon^{1/16}|\mathcal{X}|^{1/2} \cdot |\mathcal{Y}|^{1/2}}{1 + \varepsilon^{1/16}|\mathcal{X}|^{1/2} \cdot |\mathcal{Y}|^{1/2}} + 4|\mathcal{X}| \cdot |\mathcal{Y}| e^{-\frac{\varepsilon^{-1/4}}{3}} \right). \tag{6.146}$$

This concludes the proof. ■

Corollary 74 (Faithfulness of quantum intrinsic non-locality) *For every quantum distribution $p(a,b|x,y)$, the quantum intrinsic non-locality $N^Q(\bar{A};\bar{B})_p = 0$, if and only if it has a local hidden variable description. Quantitatively, if $N^Q(\bar{A};\bar{B})_p \leq \varepsilon$, where $0 < \varepsilon^{1/16}d^{1/2} < 1$, for $d = |\mathcal{X}| \cdot |\mathcal{Y}|$, there exists a probability distribution $l(a,b|x,y)$ having a local hidden-variable description, such that*

$$\sup_{p_{XY}(x,y)} \|\rho_{\bar{A}X\bar{B}Y} - \gamma_{\bar{A}X\bar{B}Y}\|_1 \leq d \left(\varepsilon^{1/4} + \frac{\varepsilon^{1/16}d^{1/2}}{1 - \varepsilon^{1/16}d^{1/2}} + 4de^{-\frac{\varepsilon^{-1/4}}{3}} \right), \tag{6.147}$$

where $\rho_{\bar{A}X\bar{B}Y}$ correponds to the classical-classical state $p_{XY}(x,y)p(a,b|x,y)$ and $\gamma_{\bar{A}X\bar{B}Y}$ is the classical-classical state corresponding to $p_{XY}(x,y)l(a,b|x,y)$.

Proof. The if-part of the proof follows from Proposition 72. The only-if part follows from Proposition 63 and Theorem 73. ■

6.3. Open questions

- *Continuity of intrinsic non-locality*: Currently, it is not known if intrinsic non-locality is continuous along the lines discussed in Section 5.5.. Suppose that we have two probability distributions $p(a,b|x,y)$ and $q(a,b|x,y)$, such that their classical-classical states $\rho_{\bar{A}\bar{B}XY}$ and $\sigma_{\bar{A}\bar{B}XY}$ are ε close in trace norm, i.e.,

$$\frac{1}{2}\|\rho_{\bar{A}\bar{B}XY} - \sigma_{\bar{A}\bar{B}XY}\|_1 \leq \varepsilon, \tag{6.148}$$

where $\varepsilon > 0$. Then, for any no-signaling quantum extension $\rho_{\bar{A}\bar{B}XYE}$ of $\rho_{\bar{A}\bar{B}XY}$, can we find a no-signaling quantum extension $\sigma_{\bar{A}\bar{B}XYE}$ of $\sigma_{\bar{A}\bar{B}XY}$, such that

$$\frac{1}{2}\|\rho_{\bar{A}\bar{B}XYE} - \sigma_{\bar{A}\bar{B}XYE}\|_1 \overset{?}{\leq} f(\varepsilon), \tag{6.149}$$

where $f(\varepsilon)$ is a function of ε such that $f(\varepsilon) \to 0$ as $\varepsilon \to 0$? If so, then we can prove the continuity of quantum intrinsic non-locality.

- *Definitions for multipartite non-locality*: Multipartite non-locality deals with non-local correlations observed in n distant parties. One can define non-locality in a multipartite scenario in a variety of different ways, as discussed in [114]. An interesting question, from the point of view of information theory, is the quantification of multipartite non-locality. We expect that conditional mutual information will be useful in characterizing multipartite non-locality as well.

  For multipartite systems, the definition of conditional mutual information is a bit complex. There are at least two possible ways of defining CMI for multipartite scenarios, as discussed in [115, 116]. We could use both of these definitions. Another difficulty in defining CMI for multipartite scenarios will arise in defining various constraints on the no-signaling quantum eavesdropper. Although the definition seems complicated as of now, it possibly will be relevant in obtaining upper bounds for key rates in device-independent conference key distillation protocols [117].

Chapter 7
Upper Bounds

In Chapter 4, we discussed the need to obtain upper bounds on the distillable key of joint distributions and bipartite probability distributions. In this context, we discussed two results: intrinsic information, which is an upper bound on distillation key from joint probability distributions [9], and squashed entanglement, which is an upper bound on distillable key from bipartite states [11]. In this chapter, we introduce upper bounds on secret key agreement capacities of assemblages $\hat{\rho}_B^{a,x}$ and conditional probability distributions $p(a,b|x,y)$. To this end, we use quantities introduced in Chapters 5 and 6, along with their properties. These upper bounds are then shown to be upper bounds on key rates in DI-QKD and 1S-DI-QKD.

In general, in the device-independent literature or the one-sided device independent literature, several prior works have devised lower bounds on the key rates for particular protocols. To test the efficacy of these protocols, one calculates the rates that would be obtained for an honest device [39, 66, 67]. These protocols then give lower bounds on the secret key rates that can be extracted from an honest device. However, one can always ask if these lower bounds are "good enough." Can some other protocol perform better, hence giving better key rates for the honest device? In this context, it becomes imperative to introduce tight upper bounds.

For this, we can use the upper bounds on the secret key agreement capacities from assemblages and probability distributions. We calculate these bounds for honest devices, which are then compared to the lower bounds on the rates calculated for protocols for these honest devices. Ideally, we want the gap between the lower bounds and upper bounds to be small.

We can formulate the above question in a different manner.

Suppose that the correlations generated from a device are characterized by a distribution $p(a,b|x,y)$ or an assemblage $\hat{\rho}_B^{a,x}$. We then pose the following question:

> Given a device characterized by $p(a,b|x,y)$ or an assemblage $\hat{\rho}_B^{a,x}$, what is a non-trivial upper bound on the secret key rate that can be extracted from this device by using any possible protocol?

We calculate the upper bounds for an i.i.d. device, which means that in each round of the protocol, the device considered is characterized by the distribution $p(a,b|x,y)$ or an assemblage $\hat{\rho}_B^{a,x}$. The inputs of the device in a particular round can be correlated with the input of the device in other rounds. The assumption that the device is characterized by an i.i.d. correlation is not a drawback since we are interested in determining upper bounds on secret key rates here. For general attacks, the key rates tend to those of collective attacks for sufficiently large n.

7.1. Upper bounds based on CMI

7.1.1. Upper bounds on secret-key-agreement capacity of conditional probability distributions

In device-independent quantum key distribution, we assume the presence of an eavesdropper who obtains all of the classical data communicated between Alice and Bob during the protocol. Furthermore, the system held by the eavesdropper can have joint correlations with the systems held by Alice and Bob. Let Alice and Bob share a quantum correlation $p(a,b|x,y)$ as defined in (2.9). Let the correlation shared between Alice, Bob, and Eve be defined by an extension $p(a,b|x,y)\rho_E^{a,b,x,y}$. If $p(a,b|x,y)\rho_E^{a,b,x,y}$ has an underlying quantum strategy as described in (6.10), then we call the eavesdropper a quantum Eve. If $p(a,b|x,y)\rho_E^{a,b,x,y}$ only fulfills the constraints given in (6.3) and (6.5), then we call the eavesdropper a no-signaling Eve.

- **No-signaling eavesdropper**

We first define the secret-key-agreement capacity of a conditional probability distribution. Let $n \in \mathbb{Z}^+$, $R \geq 0$, and $\varepsilon \in [0,1]$. Let $p(a,b|x,y)$ be a correlation of

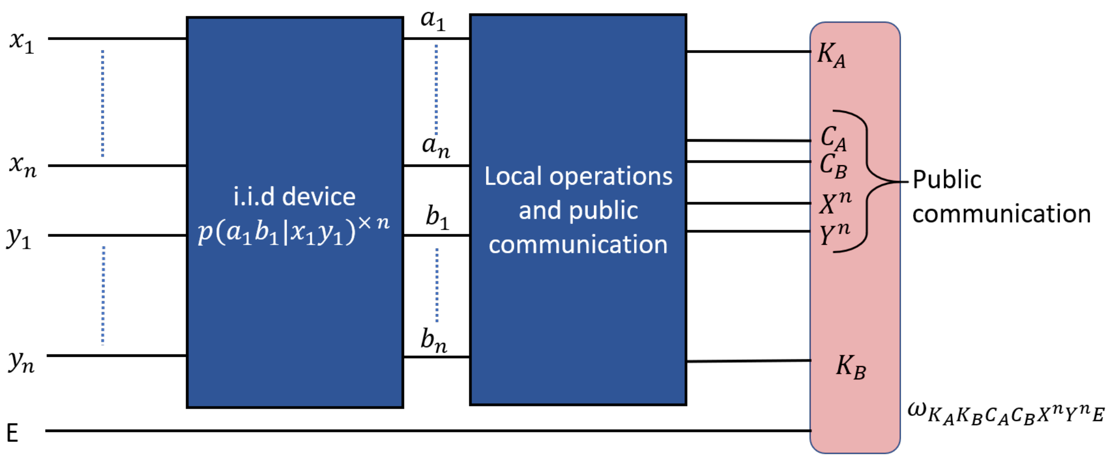
x_1
x_n
y_1
y_n
E
i.i.d device
$p(a_1 b_1|x_1 y_1)^{\times n}$
a_1
a_n
b_1
b_n
Local operations and public communication
K_A
C_A
C_B
X^n
Y^n
K_B
Public communication
$\omega_{K_A K_B C_A C_B X^n Y^n E}$

Figure 7.1. A generic device-independent quantum key distribution protocol.

the device shared between Alice and Bob. We define an (n, R, ε) device-independent secret-key-agreement protocol as follows:

- Alice and Bob select the inputs x^n and y^n to their devices according to $p_{X^nY^n}(x^n, y^n)$. The device is used n times, and the distribution $p_{X^nY^n}(x^n, y^n)$ is independent of Eve. Alice inputs x_i and obtains the output a_i. Bob inputs y_i and obtains the output b_i, where $i \in \{1, \ldots, n\}$. The input and output distributions are embedded in the state $\sigma_{\bar{A}^n\bar{B}^nX^nY^n}$, where

$$\sigma_{\bar{A}^n\bar{B}^nX^nY^n} := \sum_{x^n,y^n,a^n,b^n} p_{X^nY^n}(x^n, y^n)p^n(a^n, b^n|x^n, y^n)[a^nb^nx^ny^n]_{\bar{A}^n\bar{B}^nX^nY^n}, \tag{7.1}$$

and $p^n(a^n, b^n|x^n, y^n)$ is the i.i.d. extension of $p(a, b|x, y)$. The joint state held by Alice, Bob, and Eve is a no-signaling extension $\sigma_{\bar{A}^n\bar{B}^nX^nY^nE}$ of $\sigma_{\bar{A}^n\bar{B}^nX^nY^n}$.

- Alice and Bob perform local operations and public communication, with C_A denoting the classical register communicated from Alice to Bob, $\bar{C}_A$ is a classical register held by Eve that is a copy of C_A , the classical register C_B is communicated from Bob to Alice, and $\bar{C}_B$ is a classical register held by Eve that is a copy of C_B. This protocol yields a state $\omega_{K_A K_B E \bar{C}_A \bar{C}_B X^n Y^n}$ that satisfies

$$\frac{1}{2}\left\|\omega_{K_A K_B E X^n Y^n \bar{C}_A \bar{C}_B} - \overline{\Phi}_{K_A K_B} \otimes \omega_{E X^n Y^n \bar{C}_A \bar{C}_B}\right\|_1 \leq \varepsilon, \tag{7.2}$$

for all no-signaling extensions, where

$$\overline{\Phi}_{K_A K_B} = \frac{1}{2^{nR}} \sum_{k=1}^{2^{nR}} |kk\rangle\langle kk|_{K_A K_B} . \tag{7.3}$$

A rate R is achievable for a device characterized by p if there exists an $(n, R-\delta, \varepsilon)$ device-independent protocol for all $\varepsilon \in (0,1)$, $\delta > 0$, and sufficiently large n. The device-independent secret key agreement capacity $DI(p)$ of the device characterized by p is defined as the supremum of all achievable rates.

We now prove that the secret-key-agreement capacity of a conditional distribution is bounded from above by the intrinsic non-locality.

Theorem 75 *The intrinsic non-locality $N(\bar{A};\bar{B})_p$ is an upper bound on the device-independent secret-key-agreement capacity of a device characterized by p and sharing no-signaling correlations with an eavesdropper:*

$$DI(p) \leq N(\bar{A};\bar{B})_p. \tag{7.4}$$

Proof. For an arbitrary (n, R, ε) protocol, consider that

$$\begin{aligned}
nR &= I(K_A;K_B|EX^nY^n\bar{C}_A\bar{C}_B)_{\bar{\Phi}\otimes\omega} && (7.5)\\
&\leq I(K_A;K_B|EX^nY^n\bar{C}_A\bar{C}_B)_{\omega} + \varepsilon' && (7.6)\\
&\leq I(M_AC_BC_A;M_BC_BC_A|EX^nY^n\bar{C}_A\bar{C}_B)_{\tau} + \varepsilon' && (7.7)\\
&= I(M_AC_A;M_BC_B|EX^nY^n\bar{C}_A\bar{C}_B)_{\tau} + \varepsilon' && (7.8)
\end{aligned}$$

$$\leq I(M_A C_A; M_B C_B | E X^n Y^n)_\tau + \varepsilon' \tag{7.9}$$

$$\leq I(\bar{A}^n; \bar{B}^n | E X^n Y^n)_\sigma + \varepsilon', \tag{7.10}$$

where

$$\varepsilon' = nR\varepsilon + 2\left[(1+\varepsilon)\log(1+\varepsilon)) - \varepsilon \log \varepsilon\right]. \tag{7.11}$$

In the above equations, $\sigma_{X^n \bar{A}^n \bar{B}^n Y^n}$ is the classical-classical state obtained from the device after Alice and Bob enter in the measurement inputs. Alice, Bob, and Eve hold a no-signaling extension $\sigma_{X^n \bar{A}^n \bar{B}^n Y^n E}$. Alice performs a local operation $\mathcal{L}_A$ to obtain M_A and C_A. She communicates C_A to Bob, and Eve also obtains a copy $\bar{C}_A$ of the classical communication. Similarly, Bob performs a local operation $\mathcal{L}_B$ to obtain M_B and C_B. He communicates C_B to Alice, and Eve also obtains a copy $\bar{C}_B$ of the classical communication. Alice then performs a local operation $\mathcal{D}_A$ on M_A, C_B, and C_A to obtain K_A, while Bob performs a local operation $\mathcal{D}_B$ on M_B, C_A, and C_B to obtain K_B. For a pictorial representation of the above description, refer to Figure 7.1.

The first inequality follows from the uniform continuity of conditional mutual information [118, Proposition 1]. The second inequality follows from data processing. The second equality and third inequality follow from the chain rule of conditional mutual information, as well as the fact that $\bar{C}_A$ is a classical copy of C_A and $\bar{C}_B$ is a classical copy of C_B. The last inequality follows from data processing for conditional mutual information. Since the above inequality holds for an arbitrary no-signaling extension of $\sigma_{\bar{A}^n \bar{B}^n X^n Y^n}$, we find that

$$nR \leq \inf_{\sigma_{\bar{A}^n \bar{B}^n X^n Y^n E}} I(\bar{A}^n; \bar{B}^n | X^n Y^n E)_\sigma + \varepsilon'. \tag{7.12}$$

This implies that

$$nR \leq N(\bar{A}^n; \bar{B}^n)_p + \varepsilon'. \tag{7.13}$$

By the assumption that the device is i.i.d., we can invoke the additivity of intrinsic non-locality from Proposition 69 to obtain

$$(1-\varepsilon)R \leq N(\bar{A}; \bar{B})_p + 2\left[(1+\varepsilon)\log(1+\varepsilon)) - \varepsilon \log \varepsilon\right]/n. \tag{7.14}$$

Taking the limit as $n \to \infty$ and $\varepsilon \to 0$ then leads to $DI(p) \leq N(\bar{A}; \bar{B})_p$. ■

- **Quantum eavesdropper**

Now, let us consider a class of device-independent protocols in which the eavesdropper is restricted by quantum mechanics. These models have previously been studied in [39, 66]. The general form of a device-independent protocol with a quantum eavesdropper remains the same except that we now consider a quantum extension (6.10) of the state in (7.1). We then arrive at the following theorem:

Theorem 76 *The quantum intrinsic non-locality $N^Q(\bar{A}; \bar{B})_p$ is an upper bound on the device-independent secret-key-agreement capacity of a device characterized by p and sharing quantum correlations with an eavesdropper:*

$$DI(p) \leq N^Q(\bar{A}; \bar{B})_p. \tag{7.15}$$

Proof. The proof of the theorem is similar to that of Theorem 75. ■

We should explicitly point out that the general form for protocols that we consider allows both Alice and Bob to exchange public classical information. Therefore, the upper bounds via intrinsic non-locality and quantum intrinsic non-locality hold for two-way error correction as well. It has been observed in device-dependent QKD that two-way error-correcting protocols surpass the threshold of one-way error correction protocols [119, 120, 75]. This question has only recently been explored in DI-QKD in [121]. Therefore, it is possible that the upper bound via the intrinsic non-locality will not be tight for the existing DI-QKD protocols [39, 66], which consider only one-way error correction.

Another point to make is that in the protocols we consider, Alice and Bob announce their measurement choices. That is, X and Y are known to Eve. The secret key is extracted from $\bar{A}$ and $\bar{B}$. There are certain protocols in the device-independent literature where the outputs $\bar{A}$ and $\bar{B}$ are broadcast and the local randomness variables X and Y are the basis of the key [122] (note that [123] introduced this concept in the device-dependent QKD literature). For such DI-QKD protocols, our upper bounds do not hold.

- **Other works**

Bounds on device-independent QKD protocols based on certain states were also previously discussed in [124].

There is yet another way to model a no-signaling adversary in the device-independent secret distillation protocols, which has been considered in [79]. This model is set in "box world," in which each player including the eavesdropper has a set of possible inputs and outputs. Therefore, it becomes natural to model the joint system with a conditional probability distribution $P_{ABE|XYZ}$. In [61], the authors introduced squashed non-locality to provide an upper bound on key rates of device-independent protocols with the aforementioned model of the eavesdropper. This is in contrast to the model that we consider where the eavesdropper is a quantum no-signaling adversary but is not equipped with a set of measurement choices.

7.1.2. Upper bounds on secret-key-agreement capacity of assemblages.

In this section, we consider upper bounds on secret-key agreement capacity of assemblages.

Let $n \in \mathbb{Z}^+$, $R \geq 0$, and $\varepsilon \in [0,1]$. We define an (n, R, ε) one-sided device-independent secret-key-agreement protocol for an assemblage $\hat{\rho} := \{p_{A|X}(a|x)\rho_B^{a,x}\}_{a,x}$ as follows:

- Alice gives input x^n to get an output a^n. The assemblage shared by Alice and Bob is then

$$\rho_{\bar{A}^n X^n B^n} := \sum_{x^n, a^n} p_{X^n}(x^n) p_{A^n|X^n}(a^n|x^n)\, [x^n, a^n]_{X^n A^n} \otimes \rho_{B^n}^{a^n, x^n}, \tag{7.16}$$

  where $\{p_{A^n|X^n}(a^n|x^n)\rho_{B^n}^{a^n,x^n}\}_{a^n,x^n}$ is an i.i.d. extension of the assemblage $\{p_{A|X}(a|x)\rho_B^{a,x}\}_{a,x}$. Alice, Bob, and Eve hold a no-signaling extension of the above assemblage:

$$\rho_{\bar{A}^n X^n B^n E} := \sum_{x^n, a^n} p_{X^n}(x^n) p_{A^n|X^n}(a^n|x^n)\, [x^n, a^n]_{X^n A^n} \otimes \rho_{B^n E}^{a^n, x^n}. \tag{7.17}$$

- Bob inputs y_i and obtains the output b_i, where $i \in \{1, \ldots, n\}$. Let the measurement corresponding to y^n be a set $\{Y_{b^n}^n\}_{b^n}$ of measurement operators, such that $\sum_{b^n} (Y_{b^n}^n)^\dagger Y_{b^n}^n = I$. The state shared between Alice, Bob and Eve is then $\sigma_{\bar{A}^n X^n \bar{B}^n Y^n E}$:

$$\sigma_{\bar{A}^n X^n Y^n \bar{B}^n E} := \sum_{x^n, a^n} p_{X^n}(x^n) p_{\bar{A}^n|X^n}(a^n|x^n)\, [x^n, a^n]_{X^n \bar{A}^n} \otimes \sum_{y^n b^n} p_{Y^n}(y^n) [y^n]_{Y^n} \otimes \\ (Y_{b^n} \rho_{B^n E}^{a^n, x^n} (Y_{b^n})^\dagger). \tag{7.18}$$

- Alice and Bob perform local operations and public communication, with C_A being the classical register communicated from Alice to Bob, $\bar{C}_A$ is a classical register held by Eve that is a copy of C_A, the classical register C_B is communicated from Bob to Alice, and $\bar{C}_B$ is a classical register held by Eve that is a copy of C_B. This protocol yields a state $\omega_{K_A K_B E \bar{C}_A \bar{C}_B X^n Y^n}$ that satisfies

  $$\frac{1}{2}\left\|\omega_{K_A K_B E X^n Y^n \bar{C}_A \bar{C}_B} - \overline{\Phi}_{K_A K_B} \otimes \omega_{E X^n Y^n \bar{C}_A \bar{C}_B}\right\|_1 \leq \varepsilon, \tag{7.19}$$

  for all no-signaling extensions, where

  $$\overline{\Phi}_{K_A K_B} = \frac{1}{2^{nR}} \sum_{k=1}^{2^{nR}} |kk\rangle\langle kk|_{K_A K_B}. \tag{7.20}$$

A rate R is achievable for a device characterized by $\hat{\rho}$ if there exists an $(n, R-\delta, \varepsilon)$ one-sided device-independent protocol for all $\varepsilon \in (0,1)$, $\delta > 0$, and sufficiently large n. The one-sided device-independent capacity $SDI(\hat{\rho})$ of the device characterized by $\hat{\rho}$ is defined as the supremum of all achievable rates for $\hat{\rho}$.

Theorem 77 *The restricted intrinsic steerability $S^R(\bar{A};\bar{B})_{\hat{\rho}}$ is an upper bound on the one-sided device-independent secret-key-agreement capacity $SDI(\hat{\rho})$ of a device characterized by $\hat{\rho}$:*

$$SDI(\hat{\rho}) \leq S^R(\bar{A};B)_{\hat{\rho}}. \tag{7.21}$$

Proof. For obtaining the upper bound in the one-sided device-independent setting, we continue from (7.10) as follows:

$$\begin{aligned} nR &\leq I(\bar{A}^n;\bar{B}^n Y^n|EX^n)_\sigma - I(\bar{A}^n;Y^n|EX^n)_\sigma + \varepsilon' &(7.22)\\ &\leq I(\bar{A}^n;\bar{B}^n Y^n|EX^n)_\sigma + \varepsilon' &(7.23)\end{aligned}$$

$$\leq I(\bar{A}^n; B^n|EX^n)_\rho + \varepsilon', \tag{7.24}$$

The first inequality follows from the chain rule of conditional mutual information. The last inequality follows from data processing. Since the above inequality holds for an arbitrary no-signaling extension of $\rho_{\bar{A}^n X^n B^n}$, we obtain

$$nR \leq \inf_{\rho_{\bar{A}^n X^n B^n E}} I(\bar{A}^n; B^n|X^n E)_\rho + \varepsilon'. \tag{7.25}$$

This implies that

$$nR \leq S^R(\bar{A}^n; B^n)_{\hat{\rho}} + \varepsilon'. \tag{7.26}$$

Since we assume an i.i.d. device, we find by applying the additivity of restricted intrinsic steerability (Proposition 56) that

$$(1-\varepsilon)R \leq S^R(\bar{A}; B)_{\hat{\rho}} + 2\left[(1+\varepsilon)\log(1+\varepsilon)) - \varepsilon \log \varepsilon\right]/n. \tag{7.27}$$

Taking the limit as $n \to \infty$ and $\varepsilon \to 0$ then leads to the desired inequality $SDI(\hat{\rho}) \leq S^R(\bar{A}; B)_{\hat{\rho}}$. ■

In the following proposition, $K_D(\rho_{AB})$ refers to the distillable key of the state ρ_{AB}. For the exact definition, please refer to Definition 8 of [5].

Proposition 78 *Let ρ_{AB} be a bipartite state, $\hat{\rho}_B^{a,x}$ an assemblage resulting from the action of a POVM on Alice's system, and $p(a,b|x,y)$ a quantum correlation resulting from the action of an additional POVM on Bob's system. Then, the device-independent secret-key-agreement capacity of the quantum correlation p does not exceed the one-sided device-independent secret-key-agreement capacity of $\hat{\rho}$, which in* *turn does not exceed the distillable key of the state ρ_{AB}:*

$$DI(p) \leq SDI(\hat{\rho}) \leq K(\rho_{AB}). \tag{7.28}$$

Proof. The proof is a consequence of the following observation: the DI secret key distillation protocol is a special case of the SDI secret key distillation protocol with the measurements on Bob's side corresponding to i.i.d. measurements. Similarly, the SDI secret-key-agreement protocol is a special case of a secret-key-agreement protocol acting on the state ρ_{AB} with the local operations on Alice's side consisting of i.i.d. measurements. ■

7.1.3. Examples

In this section, we showcase the upper bounds for specific examples. We consider a characterization of honest devices and calculate upper bounds on intrinsic steerability and intrinsic non-locality for these devices. Then, we consider specific protocols, calculate the lower bounds on secret-key rates that one would obtain from these devices, and compare them with the upper bounds obtained above.

- **Device-independent protocol**

We now consider a device that is characterized by the correlation p, which has the following quantum strategy: Alice and Bob share a two-qubit isotropic state $\omega^{p}_{AB} = (1-p)\Phi_{AB} + p\pi_A \otimes \pi_B$, where $\Phi_{AB} = \frac{1}{2}\sum_{i,j=0}^{1} |ii\rangle\langle jj|$, and π denotes the maximally mixed state. This state arises from sending one share of Φ_{AB} through a depolarizing channel. Alice's measurement choices x_0, x_1, and x_2 correspond to σ_z, $\frac{\sigma_z+\sigma_x}{\sqrt{2}}$, and $\frac{\sigma_z-\sigma_x}{\sqrt{2}}$, respectively. Bob's measurement choices y_1 and y_2 correspond to σ_z and σ_x, respectively. The correlation resulting from this setup is then $p(a,b|x,y)$, with x taking values from $\{x_0, x_1, x_2\}$, the variable y taking values from $\{y_1, y_2\}$, and $a, b \in \{0,1\}$ being the measurement results. A specific device-independent protocol was studied in [39], which was then used to obtain a lower bound on the key rate from the above specified correlation. In this protocol, the rounds in which Alice and Bob input x_0 and y_1, respectively, correspond to σ_z measurements. Since the measurements are performed in the same basis, the measurement outcomes of these rounds form the basis of the raw key. The rounds in which Alice and Bob choose from $\{x_1, x_2\}$ and $\{y_1, y_2\}$ are used for checking the violation of the CHSH inequality. This choice of testing rounds and raw key rounds needs to be random. Also, the testing rounds are just a fraction of the total number of rounds.

The secret-key rate in a device-independent protocol is bounded from above as follows (Theorem 76):

$$R \leq \sup_{p(x,y)} \inf_{\rho_{\bar{A}\bar{B}XYE}} \sum_{x,y} p_{XY}(x,y) I(\bar{A};\bar{B}|E)_{\rho^{x,y}}, \tag{7.29}$$

where

$$\rho^{x,y}_{\bar{A}\bar{B}E} = \sum_{a,b} p(a,b|x,y)\, |ab\rangle\!\langle ab|_{\bar{A}\bar{B}} \otimes \rho_E^{a,b,x,y}. \tag{7.30}$$

The idea is now to consider some quantum extension of the probability distribution obtained from the black box, and then bound the quantum intrinsic non-locality from above.

The technique presented below is similar to the technique used in [125] to obtain upper bounds on the squashed entanglement of a depolarizing channel. An isotropic state is Bell local if $p \geq 1 - \frac{1}{\sqrt{2}}$ [126]. This implies that the quantum intrinsic non-locality of a correlation derived from ω^p_{AB} is equal to zero for $p \geq 1 - \frac{1}{\sqrt{2}}$ (Proposition 72). For $\epsilon \leq p \leq 1 - \frac{1}{\sqrt{2}}$, we can write the probability distribution $q_{\omega^p}(a,b|x,y)$ obtained from ω^p_{AB} as a convex combination of probability distributions obtained from

ω^{ϵ} and $\omega^{1-1/\sqrt{2}}$. That is, for some $0 \leq \alpha \leq 1$, we have

$$q_{\omega^p}(a,b|x,y) = (1-\alpha(\epsilon))q_{\omega^\epsilon}(a,b|x,y) + \alpha(\epsilon)q_{\omega^{1-1/\sqrt{2}}}(a,b|x,y). \tag{7.31}$$

By simple algebra, we obtain

$$\alpha(\epsilon) = \frac{p-\epsilon}{1-\frac{1}{\sqrt{2}}-\epsilon}. \tag{7.32}$$

Equation (7.31) can be written as

$$q_{\omega^p}(a,b|x,y) = (1-\alpha(\epsilon))q_{\omega^\epsilon}(a,b|x,y) + \alpha(\epsilon)\sum_{\lambda} p(\lambda)q_{\omega^{1-1/\sqrt{2}}}(a,|x,\lambda)q_{\omega^{1/\sqrt{2}}}(b,|y,\lambda). \tag{7.33}$$

Then, from convexity of quantum intrinsic non-locality (Proposition 68), we obtain

$$N^Q(\bar{A};\bar{B})_{q_{\omega^p}} \leq (1-\alpha(\epsilon))N^Q(\bar{A};\bar{B})_{q_{\omega^\epsilon}}. \tag{7.34}$$

Since the above equation is true for all α, we find that

$$N^Q(\bar{A};\bar{B})_{q_{\omega^p}} \leq \min_{0\leq\epsilon\leq p}(1-\alpha(\epsilon))N^Q(\bar{A};\bar{B})_{q_{\omega^\epsilon}}. \tag{7.35}$$

This implies that

$$N^Q(\bar{A};\bar{B})_{q_{\omega^p}} \leq \min_{0\leq\epsilon\leq p}(1-\alpha(\epsilon)) \sup_{p(x,y)} \inf_{\rho_{\bar{A}\bar{B}XYE}(\epsilon)} \sum_{x,y} p(x,y) I(\bar{A};\bar{B}|E)_{\rho^{x,y}_{\bar{A}\bar{B}E}(\epsilon)}, \tag{7.36}$$

where q^{ϵ}_{ω} is encoded in $\rho_{\bar{A}\bar{B}XY}(\epsilon)$ with $\rho_{\bar{A}\bar{B}XYE}(\epsilon)$ as the quantum extension. Let us

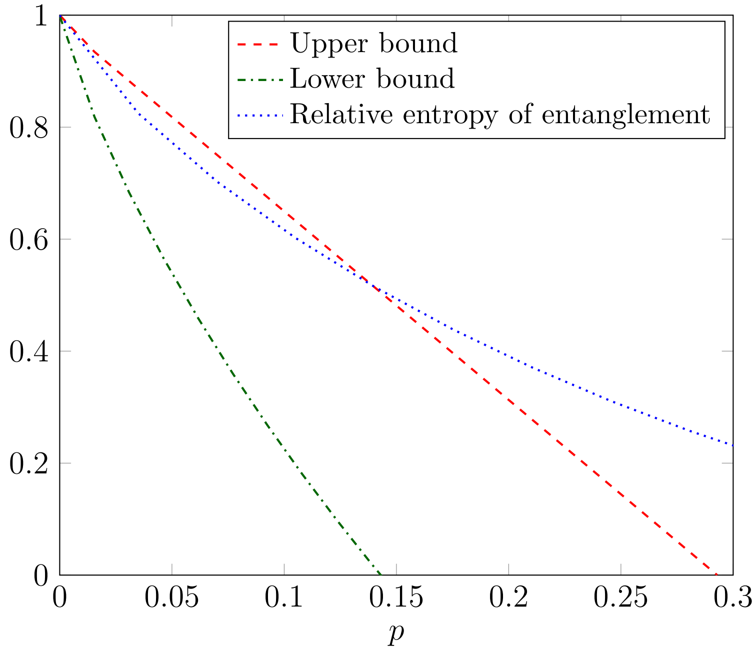

Figure 7.2. In this figure, we plot the upper bound in (7.38) and the lower bound from [39] for the device-independent protocol described in Section 7.1.3.. The relative entropy of entanglement of a qubit-qubit isotropic state is given in [44]. For further explanation of this plot, see the next section.

choose a trivial extension of the state $\rho^{x,y}_{\bar{A}\bar{B}}(\epsilon)$. It is easy to see that

$$I(\bar{A};\bar{B})_{\rho^{0,1}_{\bar{A}\bar{B}}(\epsilon)} \geq I(\bar{A};\bar{B})_{\rho^{x,y}_{\bar{A}\bar{B}}(\epsilon)} \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}. \tag{7.37}$$

Therefore,

$$R \leq \min_{0\leq\epsilon\leq p}(1-\alpha(\epsilon))I(\bar{A};\bar{B})_{\rho^{0,1}_{\bar{A}\bar{B}}(\epsilon)} = \min_{0\leq\epsilon\leq p}(1-\alpha(\epsilon))\left(\frac{2-\epsilon}{2}\log_2(2-\epsilon)+\frac{\epsilon}{2}\log_2\epsilon\right). \tag{7.38}$$

We plot this upper bound in Figure 7.2, and we interpret it and explain the relative entropy of entanglement bound in the next subsection.

The lower bound used in Figure 7.2 was obtained in [39]. The proof of the lower bound begins by first invoking the Devetak-Winter formula [2], as discussed in Chapter 3. Then, they introduced novel techniques to obtain an upper bound on the Holevo information term in the Devatak-Winter formula in terms of the CHSH violation. They also proved that the lower bound is tight for one-way classical communication protocols.

- **One-sided device-independent protocol**

Let us now consider an assemblage $\hat{\rho}(p)$ that is generated from an isotropic state, with $x_0 = \sigma_z$ and $x_1 = \sigma_x$. Then

$$\begin{aligned}\rho_{X\bar{A}B}(p) = &\frac{1}{4}\left(|0\rangle\langle 0|_X \otimes [|0\rangle\langle 0|_{\bar{A}} \otimes ((1-p)\,|0\rangle\langle 0|_B + p\pi_B)]\right)\\ &+\frac{1}{4}\left(|0\rangle\langle 0|_X \otimes [|1\rangle\langle 1|_{\bar{A}} \otimes ((1-p)\,|1\rangle\langle 1|_B + p\pi_B)]\right)\\ &+\frac{1}{4}\left(|1\rangle\langle 1|_X \otimes [|0\rangle\langle 0|_{\bar{A}} \otimes ((1-p)\,|+\rangle\langle +|_B + p\pi_B)]\right)\\ &+\frac{1}{4}\left(|1\rangle\langle 1|_X \otimes [|1\rangle\langle 1|_{\bar{A}} \otimes ((1-p)\,|-\rangle\langle -|_B + p\pi_B)]\right).\end{aligned} \tag{7.39}$$

If $p \geq 1/2$, it is known that $\rho_{X\bar{A}B}$ is unsteerable [26], and therefore intrinsic steerability is equal to zero for $p \geq \frac{1}{2}$ ([12, Proposition 7]). For $\epsilon \leq p \leq \frac{1}{2}$, we can write the state $\rho_{X\bar{A}B}(p)$ as a convex combination of states $\rho_{X\bar{A}B}(\epsilon)$ and $\rho_{X\bar{A}B}(\frac{1}{2})$. That is, for some $0 \leq \alpha \leq 1$

$$\rho_{X\bar{A}B}(p) = (1-\alpha)\rho_{X\bar{A}B}(\epsilon) + \alpha\rho_{X\bar{A}B}\left(\tfrac{1}{2}\right). \tag{7.40}$$

Then, by simple algebra we obtain

$$\alpha(\epsilon) = \frac{p-\epsilon}{\frac{1}{2}-\epsilon}. \tag{7.41}$$

From convexity of restricted intrinsic steerability (Proposition 55), we obtain

$$S(\bar{A};B)_{\hat{\rho}(p)} \leq S(\bar{A};B)_{\hat{\rho}(\epsilon)}. \tag{7.42}$$

Following the same argument as before, we obtain

$$S^R(\bar{A};B)_{\hat{\rho}(p)} \leq \min_{0\leq\epsilon\leq p}(1-\alpha(\epsilon)) \sup_{p_X(x)} \inf_{\rho_{\bar{A}BXE(\epsilon)}} \sum_{p_X(x)} p_X(x) I(\bar{A};B|E)_{\rho_{\bar{A}BE}(\epsilon)}. \tag{7.43}$$

Let us now choose a trivial extension of the assemblage. It is easy to see that

$$I(\bar{A};B)_{\rho^0(\epsilon)} = I(\bar{A};B)_{\rho^1(\epsilon)} \tag{7.44}$$

$$= 1 + \left(\tfrac{\epsilon}{2}\right)\log\left(\tfrac{\epsilon}{2}\right) + \left(1-\tfrac{\epsilon}{2}\right)\log\left(1-\tfrac{\epsilon}{2}\right). \tag{7.45}$$

We therefore obtain

$$S^R(\bar{A};B)_{\rho} = \min_{0\leq\epsilon\leq p}(1-\alpha(\epsilon))\left(1 + \left(\tfrac{\epsilon}{2}\right)\log\left(\tfrac{\epsilon}{2}\right) + \left(1-\tfrac{\epsilon}{2}\right)\log\left(1-\tfrac{\epsilon}{2}\right)\right). \tag{7.46}$$

We plot this bound in Figure 7.3.

Due to the fact that squashed entanglement is an upper bound on the rate at which secret key can be distilled from an isotropic state [11, 127], as well as the above protocols being particular protocols for secret key distillation, squashed entanglement is also an upper bound on the rate at which secret key can be distilled in one-sided device-independent and device-independent protocols. However, the upper bound on squashed entanglement of an isotropic state that we obtain after choosing the extension as given in [125] is greater than the bound obtained on restricted intrinsic steerability of the assemblage considered above. Therefore, we do not plot the squashed-entanglement bounds in Figures 7.2 or 7.3. For the same reason given above, the relative entropy of entanglement is also an upper bound on the rate at which secret key can be distilled in one-sided device-independent and device-independent protocols [5]. The relative entropy of entanglement of qubit-qubit isotropic states has been

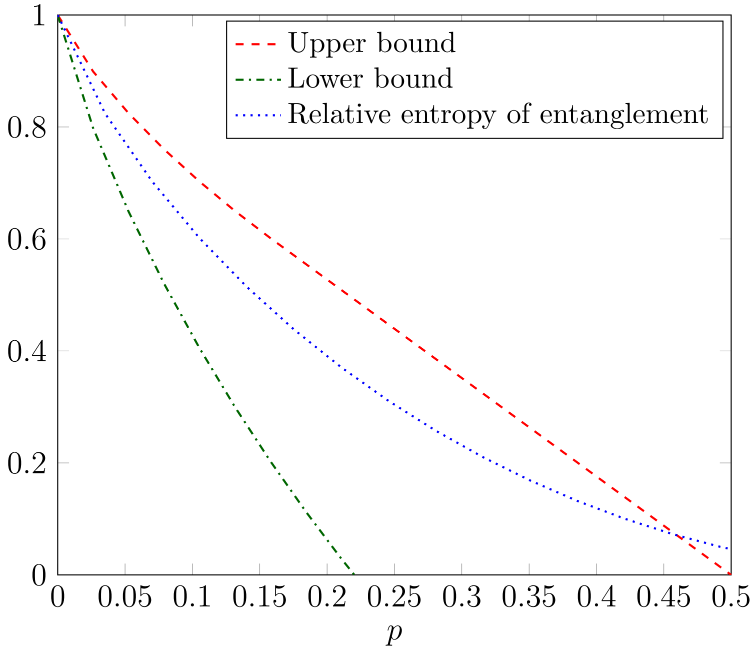

Figure 7.3. In this figure, we plot the upper bound in (7.46) and the lower bound from [38] for the one-sided device-independent protocol described in Section 7.1.3.. The relative entropy of entanglement of a qubit-qubit isotropic state is given in [44].

calculated in [44], which we plot in the above figures. This bound performs better than intrinsic non-locality and intrinsic steerability in certain regimes. This suggests that it might be worthwhile to explore if relative entropy of steering [37, 55] and relative entropy of non-locality [60] would be useful as upper bounds for one-sided device-independent and device-independent quantum key distribution, respectively. Another possible reason for this is that intrinsic non-locality is not a function of the particular Bell inequality being invoked in the protocol.

The bounds that we obtain do not closely match the lower bounds obtained from prior literature. One reason for this discrepancy can be traced back to the following question: is a violation of Bell inequality or steering inequality sufficient for security in DI-QKD and 1S-DI-QKD? Since our measure is faithful, it is equal to zero if and only if there is no violation of steering inequality or Bell inequality. However, the lower bounds hit zero at a lower value of p than expected from the faithfulness condition. Another possible reason for the discrepancy has been discussed in Section 7.1.3., pertaining to two-way error correction that is allowed in the protocols considered above. Another point to note here is that the aforementioned protocol relies on the violation of the CHSH inequality, while the bounds that we have are for protocols that violate other inequalities as well.

7.1.4. Open questions

We know from Chapter 4 that modifications to intrinsic information along the lines of [92] and [93] give better upper bounds on the secret-key-agreement capacity of a joint probability distribution. One important area of investigation here is to explore if such modifications of squashed entanglement, intrinsic steerability, and intrinsic non-locality give better upper bounds for different settings of QKD.

7.2. Relative entropy bounds

In this section, we explore the connection between measures based on relative entropy and secret-key-agreement capacity. In this direction, a remarkable result was proven in [25], which shows that the relative entropy of entanglement is an upper bound on the distillable key of a bipartite state ρ_{AB}. This then points to the following direction: can relative entropy of steering and Bell non-locality be proven to be upper bounds on the distillable key from assemblages and conditional probability distributions, respectively? This question has still not been resolved.

We then discuss techniques introduced in [5], which make it possible to establish a connection between relative entropy of entanglement with upper bounds on the distillable key of bipartite states. We also discuss the difficulties encountered with proving relative entropy of steering and nonlocality as upper bounds on secret-key-agreement capacity of assemblages and conditional probability distributions, respectively.

7.2.1. Relative entropy of entanglement and secret key distillation

Two major conceptual insights of [5] were instrumental in proving relative entropy of entanglement as an upper bound on the distillable key of bipartite states. The first one was the introduction of *private states*. The second one was showing the equivalence between distillation of ρ_{AB} and distillation of secret keys from ψ^{ρ}_{ABE}, where ψ^{ρ}_{ABE} is a purification of ρ_{AB}.

Due to the monogamy of entanglement, we know that the maximally entangled state is in tensor product with any other party. If Alice and Bob measure this state in the same basis, then they obtain a secure key. That is,

- Alice and Bob share a maximally entangled state Φ_{AB}. This implies that the quantum system of any eavesdropper is in a tensor product with Φ_{AB}.

- Alice and Bob perform measurements on their respective systems. If they perform measurements in the same basis, then they obtain the following ideal secret key:

$$\rho_{\bar{A}\bar{B}E} = \frac{1}{d}\sum_{i=1}^{d} |ii\rangle\langle ii|_{\bar{A}\bar{B}} \otimes \rho_E. \tag{7.47}$$

Is the maximally entangled state (up to local isometries) the only state from which we can extract an ideal secret key? One of the main insights of [5] was to answer this question with the introduction of private states, defined as follows:

$$\rho_{ABA'B'} = U^t_{AA'BB'}(\Phi_{AB} \otimes \rho_{A'B'})U^{t\dagger}_{AA'BB'}, \tag{7.48}$$

where $\rho_{A'B'}$ is an arbitrary bipartite state and $U^t_{AA'BB'} = \sum_{i,j=0}^{d-1} |ij\rangle\langle ij|_{AB} \otimes U^{i,j}_{A'B'}$, with each $U^{ij}_{A'B'}$ a unitary operator. An ideal secret key can be extracted from an arbitrary state $\sigma_{AA'BB'}$ if and only if it can be expressed as a private state; see Theorem 2 of [5]. This shows that the unit of secrecy is a private state instead of a maximally entangled state. A maximally entangled state is thus a specific private state.

Now, in a secret key distillation protocol, we have three parties: Alice, Bob, and Eve. These three parties share a purification ψ^ρ_{ABE} of the state ρ_{AB}, which is the state shared between Alice and Bob. Alice and Bob perform an LOPC operation to distill out a secret key, as discussed in Section 4.1.. The distillable key of ρ_{AB} is denoted by $K_D(\psi^\rho_{ABE})$.

A possible key distillation protocol is as follows: first, distill out a private state from ρ_{AB} using LOCC operations and then perform measurements on the private state to distill a secret key. This is possible since we know that a private state contains a secret key that can be accessed by measurements on the A and B systems. Let us denote the optimal rate of private-state distillation from ρ_{AB} as $C_D(\rho_{AB})$. For the exact definition, see [5]. It is then simple to see $C_D(\rho_{AB}) \leq K_D(\psi^{\rho}_{ABE})$.

Another insight of [5] was to prove that $C_D(\rho_{AB}) = K_D(\psi^{\rho}_{ABE})$. To this end, they proved that given an LOPC protocol, which extracts a secret key from a pure state ψ^{ρ}_{ABE}, one can define a coherent version of this LOPC protocol. The coherent version of the LOPC protocol acts on a pure state ψ_{ABE} and gives $\psi_{AA'BB'E}$, such that tracing out the A' and B' systems gives an ideal secret key. Tracing out Eve's system in $\psi_{AA''BB'E}$ gives the private state $\rho_{AA'BB'}$. Thus to every LOPC protocol, which gives an ideal key, they defined a coherent LOPC protocol. The removal of Eve's system from the coherent LOPC protocol defines an LOCC protocol with the output as a private state. This equivalence then proved that $C_D(\rho_{AB}) = K_D(\psi^{\rho}_{ABE})$. Therefore, to obtain upper bounds on the distillable key of ρ_{AB}, one can upper bound the rate of private-state distillation instead. This upper bound is obtained in terms of the relative entropy of entanglement.

The definition of relative entropy of entanglement (REE) does not make any explicit reference to an eavesdropper system. This is in contrast to squashed entanglement, where one can think of the conditioning system as an eavesdropper's system. The lack of reference to an eavesdropper's system in REE was the main bottleneck in proving upper bounds on distillable key in terms of relative entropy of entanglement. This was resolved in [5] via the introduction of private states and the coherent version of LOPC protocols.

7.2.2. Open question: relative entropy of Bell non-locality and secret key distillation

In Section 2.2., we discussed a method of defining monotones in resource theories based on relative entropy. Using that formalism, we can define relative entropy of non-locality [60] as follows:

Definition 79 (Relative entropy of non-locality) *The relative entropy of Bell* *non-locality of the conditional probability distribution $p(a,b|x,y)$ is defined as*

$$R^L(p(a,b|x,y)) = \sup_{p_{XY}(x,y)} \inf_{q(a,b|x,y)\in\mathbf{L}} D\left(p(x,y)p(a,b|x,y) \| p(x,y)q(a,b|x,y)\right). \quad (7.49)$$

We would like to address the following question: Is relative entropy of Bell non-locality an upper bound on the distillable key of a probability distribution? That is,

$$R^L(p(a,b|x,y)) \overset{?}{\geq} K_D(p(a,b|x,y)). \quad (7.50)$$

As far as the author is aware, this is still an unresolved question. In this context, first we need to identify the unit of privacy in quantum distributions. This unit of privacy should satisfy two main criteria:

- For some x and y, $p(a=i, b=i|x,y) = \frac{1}{d}$, where $i \in \{0, \dots d-1\}$.
- This probability distribution should be in tensor product with any quantum extension of the eavesdropper.

Next, let Alice and Bob share a device characterized by a quantum distribution $q(a,b|x,y)$, with the eavesdropper holding any quantum extension. Then, the distillable key via an LOPC operation is $K_D(q(a,b|x,y))$. Let the distillation of $p(a,b|x,y)$ from $q(a,b|x,y)$, with a local operation and shared randomness operation, be denoted by $C_D(q(a,b|x,y))$. Then, is $C_D(q(a,b|x,y)) \overset{?}{=} K_D(q(a,b|x,y))$? Addressing these two questions will be helpful in proving that the relative entropy of non-locality is an upper bound on distillable key.

7.2.3. Open question: relative entropy of steering and secret key distillation from assemblages

We can define restricted relative entropy of steering [1] [55] as follows:

[1] Relative entropy of steering, which is a monotone under 1W-LOCC, has been introduced in [37, 55].

Definition 80 (Restricted relative entropy of steering) *Let* $\{\hat{\rho}_B^{a,x}\}_{a,x}$ *be an assemblage. Then the restricted relative entropy of steering is given by*

$$R_S^R(\overline{A};B)_{\hat{\rho}} := \sup_{p_X} \inf_{\{\hat{\sigma}_B^{a,x}\}_{a,x} \in \text{LHS}} D(\rho_{X\overline{A}B} \| \sigma_{X\overline{A}B}), \tag{7.51}$$

where

$$\rho_{X\overline{A}B} := \sum_{x,a} p_X(x)|x\rangle\langle x|_X \otimes |a\rangle\langle a|_{\overline{A}} \otimes \hat{\rho}_B^{a,x}, \tag{7.52}$$

$$\sigma_{X\overline{A}B} := \sum_{x,a} p_X(x)|x\rangle\langle x|_X \otimes |a\rangle\langle a|_{\overline{A}} \otimes \hat{\sigma}_B^{a,x}. \tag{7.53}$$

We would now like to address the following question: Is the restricted relative entropy of steering an upper bound on the distillable key of an assemblage? That is

$$R_S^R(\bar{A};B)_{\hat{\rho}} \overset{?}{\geq} K_D(\hat{\rho}_B^{a,x}). \tag{7.54}$$

The bottleneck is again the lack of reference to an explicit Eavesdropper system in the definition of restricted relative entropy of steering.

Chapter 8
Future Directions and Open Questions

In this thesis, we discussed conditional mutual information quantifiers for quantum steering and non-locality, which have been inspired by intrinsic information [9] and squashed entanglement [10]. Subsequently, we proved various properties of the quantifiers, such as faithfulness, convexity, monotonicity under the free operations, superadditivity, and additivity under tensor products. We then used these properties to prove that restricted intrinsic steerability and quantum intrinsic non-locality are upper bounds on secret-key rates in device-independent secret-key-agreement protocols. Then, we showcased these bounds for particular examples. We also discussed various open questions regarding the properties of these quantities and upper bounds on secret-key rates in device-independent secret-key-agreement protocols. Now, we discuss various future directions.

We discussed in Chapter 3, the close connection between various correlations and different settings in QKD. Several works, such as [128, 129, 130], have explored other correlations in quantum mechanics besides entanglement, steering, and non-locality. This introduction of various other resources has expanded the toolkit available for quantum key distribution theorists. A unified framework to describe these correlations was introduced in [131]. We first give an overview of this unification. Motivated by this development, we describe various possible settings in quantum key distribution.

8.1. Possible settings for quantum key distribution

In this formalism, we consider channels $\mathcal{N}$ that accept two inputs and give two outputs. Alice's input is denoted by $\mathcal{X}$ and output by $\mathcal{A}$. Bob's input is denoted by $\mathcal{Y}$ and output by $\mathcal{B}$. These two-input, two-output channels can be described as $\mathcal{N} : \mathcal{S}(\mathcal{H}_X \otimes \mathcal{H}_Y) \to \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, with one input-output pair associated with Alice's lab and the other input-output pair associated with Bob's lab. The inputs can be either a classical random variable, a quantum state, or a trivial input. By a trivial input, we mean that the channel's output is independent of the input. We also impose a no-signaling constraint on the channel, which implies that the input on Alice's side cannot influence Bob's output and vice versa. That is,

$$\mathcal{N}_{\mathcal{X}\mathcal{Y}\rightarrow\mathcal{B}}\left(\rho\otimes\sigma\right)=\mathrm{Tr}\left[\rho\right]\mathcal{N}_{\mathcal{Y}\rightarrow B}\left(\sigma\right), \tag{8.1}$$

where ρ and σ are *any* inputs to the channel, and $\mathcal{N}_{\mathcal{X}\mathcal{Y}\rightarrow\mathcal{B}}=\mathrm{Tr}_{\mathcal{A}}\left[\mathcal{N}_{\mathcal{X}\mathcal{Y}\rightarrow\mathcal{B}\mathcal{A}}\right]$ is the reduced channel. A similar no-signaling condition can be imposed on $\mathcal{N}_{\mathcal{X}\mathcal{Y}\rightarrow\mathcal{A}}$. As an example, suppose that Alice and Bob receive an unknown bipartite state. Then, we can map the generation of this bipartite state to a channel with trivial inputs and two quantum outputs. The output state is independent of the inputs of the channel. Bipartite quantum channels have been studied in [132, 133] and in the references therein. Various information-theoretic upper bounds on the entanglement and secret-key-agreement capacities of a bipartite channel were considered in [134, 135, 136] .

To connect this framework with quantum key distribution, we can think of the aforementioned channel as describing a device that Alice and Bob have in their laboratories. The inputs and outputs describe the interaction that Alice and Bob have with this device. For example, trivial inputs imply that the output of the device is independent of the input. If the channel has quantum outputs, this means that the quantum state shared between Alice and Bob is not characterized. If the channel has classical outputs, then we can assume that the measurement devices with Alice and Bob are untrusted.

The trivial inputs are denoted by t, classical input and output by c, and quantum input and output by q. To describe the channel we use $df-gh$, where dh are inputs to the channel, gh are the outputs of the channel, $d,h\in\{t,c,q\}$, and $g,f\in\{c,q\}$.

Let us now consider different resources in this model. First, we start with two trivial inputs and two quantum outputs: a $tt - qq$ channel. Since the channel is not trusted, the output of the channel, which is a bipartite state, is also not trusted. This corresponds to the setting of trusted quantum key distribution, in which Alice and Bob receive an uncharacterized state from a source. Once Alice and Bob have the untrusted state, they can perform trusted measurements on this state to obtain secret keys.

Next, we consider a $ct - cq$ channel, which means that Alice inputs a classical input and obtains a classical output. Also, Bob obtains a quantum output. This corresponds to the setting of a steering assemblage, in which Alice's device is untrusted, Bob's system is untrusted, and Bob's measurement device is trusted. This resource is relevant for one-sided device-independent quantum key distribution.

Next, we consider a $cc - cc$ channel, which means that Alice and Bob's inputs, as well as outputs, are classical. We can relate this to the scenario in which Alice and Bob's preparation, as well measurement device, is untrusted. This corresponds precisely to the scenario of device-independent quantum key distribution with the resource being a Bell non-local box.

Now let us consider another resource wherein Alice and Bob input quantum states to an untrusted channel and obtain classical outputs. The channel is described by $qq - cc$ channel. This channel corresponds to the scenario in which Alice and Bob trust the state preparation; however, they do not trust the measurement procedure. This corresponds to the scenario of measurement device-independent quantum key distribution, which has been studied in [86, 87].

Next, we consider the following scenario: Alice inputs a quantum state while Bob's input is trivial. Alice's output is classical, and Bob's output is quantum. This setting corresponds to a $qt - cq$ channel, in which Alice's state preparation is trusted, but her measurement device is not trusted. On Bob's side, the state preparation is untrusted because it is an output of an untrusted channel, while Bob's measurement devices are trusted. The resource in this scenario is called a teleportage, introduced in [128]. As far as the author is aware, such a setting of quantum key distribution has not yet been considered in the literature, thus remaining unexplored.

The next setting that one can consider is as follows: Alice's input and output are classical. Bob's input is quantum, and his output is classical. The channel is described as a $cq - cc$ channel. This corresponds to the setting in which Alice does not trust her preparation and measurement device, while Bob does not trust his measurement device. This is in contrast to a one-sided device-independent protocol in which Bob trusts the measurement but not the preparation. As far as the author is aware, such a setting of quantum key distribution has not yet been considered in the literature. We summarize the contents of this section in Table 8.1.

8.2. Measurement-device-independent QKD

In this section, we first introduce the basic components of a measurement-device-independent quantum key distribution protocol, as discussed in [86, 87]. We then introduce a conditional mutual information-based function, which could be an upper bound on the key rate that can be extracted in this setting.

- Alice and Bob prepare the states $\rho_{AA'}$ and $\rho_{BB'}$. They send the A' and B' systems to a third party Charlie over untrusted quantum channels $\mathcal{N}_1$ and $\mathcal{N}_2$.
- Charlie performs a quantum instrument $\Pi_{A'B' \to EC_1C_2}$, where C_1 and C_2 are the classical outputs, and E is a quantum output.
- Charlie publicly communicates the classical outputs C_1 and C_2 to Alice and Bob. He keeps the system E with himself.
- Alice and Bob receive the classical outputs and perform local quantum channels

Table 8.1. Various settings in quantum key distribution. In this table, ✓ implies that the device is trusted and ✗ implies that the device is untrusted.

Object	Alice's preparation device	Alice's measurement device	Bob's preparation device	Bob's measurement device	Setting
Quantum state $tt - qq$	N/A	✓	N/A	✓	Trusted QKD [3]
Assemblage $ct - cq$	✗	✗	N/A	✓	1S-DI-QKD [67]
Non-local box $cc - cc$	✗	✗	✗	✗	DI-QKD [23]
Distributed measurement $qq - cc$	✓	✗	✓	✗	MDI-QKD [86, 87]
Teleportage $qt - cq$	✓	✗	N/A	✓	Unexplored setting
MDI steering assemblage $cq - cc$	✗	✗	✓	✗	Unexplored setting

and measurements on their systems to obtain the classical bits $\bar{A}$ and $\bar{B}$.

- These steps are repeated n times. At the end of these n rounds, Alice has $\bar{A}^n$ and Bob has $\bar{B}^n$.

- Alice and Bob then perform error correction and privacy amplification using public authenticated channels to extract a secret key. The state at the end of the protocol is $\sigma_{K_AK_BE^nC_1C_2E'}$, where E' captures all the classical information that Eve gathers because of Alice and Bob's public communication.

- For an (n, R, ε) protocol, the following inequality holds

$$\frac{1}{2}\|\sigma_{K_AK_BEC_1C_2E'} - \overline{\Phi}_{K_AK_B} \otimes \sigma_{EC_1C_2E'}\|_1 \leq \varepsilon, \tag{8.2}$$

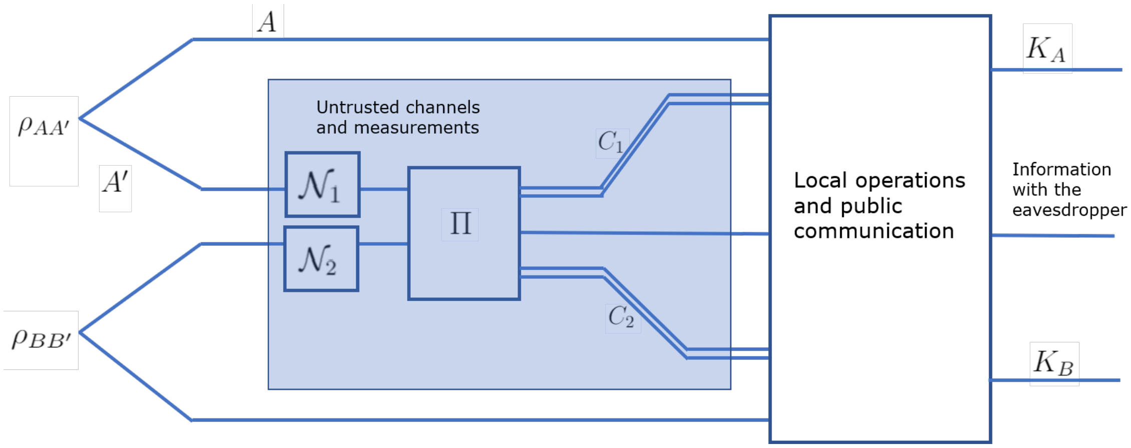
A
$\rho_{AA'}$
A'
$\rho_{BB'}$
Untrusted channels and measurements
$\mathcal{N}_1$
$\mathcal{N}_2$
Π
C_1
C_2
Local operations and public communication
K_A
Information with the eavesdropper
K_B

Figure 8.1. Pictorial representation of a measurement-device-independent protocol.

where $\varepsilon \geq 0$, and $\bar{\Phi}_{K_AK_B} = \frac{1}{d}\sum_{j=0}^{d-1} |jj\rangle\!\langle jj|_{K_AK_B}$.

A rate R is achievable for $(\mathcal{N}_1, \mathcal{N}_2, \Pi)$ if there exists an $(n, R-\delta, \varepsilon)$ measurement-device-independent protocol for all $\varepsilon \in (0,1)$, $\delta > 0$, and sufficiently large n. The measurement-device-independent secret-key-agreement capacity of $\mathrm{MDI}(\mathcal{N}_1, \mathcal{N}_2, \Pi)$ is defined as the supremum of all achievable rates. For a pictorial depiction of this protocol, see Figure 8.1.

In this setting, we trust the state preparation part. That is, we suppose that we have a characterization of $\rho_{AA'}$ and $\rho_{BB'}$. We also suppose that Alice and Bob's laboratory, where they prepare the states $\rho_{AA'}$ and $\rho_{BB'}$, are completely shielded from any side-channel attacks and do not leak out any information to the eavesdropper. Charlie can be completely untrusted and hence can be an eavesdropper. Therefore, the quantum instrument implemented by Charlie is completely untrusted. This protocol is resilient towards any side-channel attacks on the measurement devices.

8.2.1. CMI-based measure for MDI-QKD

We want to quantify the correlations generated between the systems A' and B', which are not shared with C_1, C_2, and E. Now, recall that $\rho_{AA'}$ is in tensor product with $\rho_{BB'}$. To ensure that there are some correlations existing between the systems

$\rho_{AA'}$ and $\rho_{BB'}$, the channels $\mathcal{N}_1$ and $\mathcal{N}_2$ should not be entanglement breaking. The measurement Π should also generate entanglement between systems A and B. The measure we introduce here is a function of the channel that connects Alice-Charlie and Bob-Charlie, and the measurement Π, which is employed by Charlie. We now introduce the quantifier for distributed measurement, $\mathrm{MDI}(\mathcal{N}_1, \mathcal{N}_2, \Pi)$, as follows:

$$\mathrm{MDI}(\mathcal{N}_1, \mathcal{N}_2, \Pi) = \frac{1}{2} \sup_{\rho_{AA'}, \rho_{BB'}} \inf_{\mathcal{S}_{E\to E'}} I(A;B|C_1C_2E')_\rho, \tag{8.3}$$

where

$$\rho_{ABEC_1C_2} = \mathcal{S}_{E\to E_1} \left(\mathrm{id}_{AB} \otimes \Pi_{A'B'\to C_1C_2E}\right) \left(\mathrm{id} \otimes \mathcal{N}_1 \otimes \mathrm{id} \otimes \mathcal{N}_2\right) \left(\rho_{AA'} \otimes \rho_{BB'}\right), \tag{8.4}$$

and $\mathcal{S}_{E\to E'}$ is a quantum channel implement by Charlie.

To make $\mathrm{MDI}(\mathcal{N}_1, \mathcal{N}_2, \Pi)$ independent of the states that Alice and Bob input to the unknown channels, we take the supremum over $\rho_{AA'}$ and $\rho_{BB'}$. That is, we suppose that Alice and Bob input a state that maximizes the possible correlations given the channels $\mathcal{N}_1$, $\mathcal{N}_2$, and the instrument Π. We also have an infimum over quantum channels that can be performed by the eavesdropper or Charlie. This is required because CMI is not monotone under the actions of a local channel on the conditioning system. The factor of 1/2 is required due to the normalization condition.

An MDI-QKD protocol should give a secure key irrespective of the channels $\mathcal{N}_1, \mathcal{N}_2$, or the measurement Π implemented by an eavesdropper. In most MDI-QKD protocols, Charlie implements a Bell measurement. While, in reality, it is hard to implement a Bell measurement perfectly, any noise due to the implementation is attributed to the eavesdropper/Charlie. This is precisely the reason to call this protocol measurement device-independent. The channels from Alice-Charlie or Bob-

Charlie are also untrusted, and all the noise is attributed to the eavesdropper. Again, one wants to check to the efficacy of the protocols. In this context, having upper bounds will be useful. One can calculate the upper bounds for a particular model of $\mathcal{N}_1$ and $\mathcal{N}_2$, and the instrument Π that we think Charlie would implement. We can calculate this bound for expected noise models and compare these to the bounds that we obtain from a particular protocol.

Proving that this quantity is an upper bound for collective attacks is an open question and currently under consideration.

References

[1] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[2] Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 461(2053):207–235, 2005.

[3] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public-key distribution and coin tossing. *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India*, pages 175–179, 1984.

[4] Peter W. Shor and John Preskill. Simple Proof of Security of the BB84 Quantum Key Distribution Protocol. *Physical Review Letters*, 85(2):441–444, 2000.

[5] Karol Horodecki, Michal Horodecki, Pawe Horodecki, and Jonathan Oppenheim. General paradigm for distilling classical key from quantum states. *IEEE Transactions on Information Theory*, 55(4):1898–1929, April 2009. arXiv:quant-ph/0506189.

[6] Masahiro Takeoka, Saikat Guha, and Mark M. Wilde. The squashed entanglement of a quantum channel. *IEEE Transactions on Information Theory*, 60(8):4987–4998, August 2014. arXiv:1310.0129.

[7] Stefano Pirandola, Riccardo Laurenza, Carlo Ottaviani, and Leonardo Banchi. Fundamental limits of repeaterless quantum communications. *Nature Communications*, 8:15043, 2017.

[8] Mark M. Wilde, Marco Tomamichel, and Mario Berta. Converse bounds for private communication over quantum channels. *IEEE Transactions on Information Theory*, 63(3):1792–1817, 2017.

[9] Ueli M. Maurer and Stefan Wolf. Unconditionally secure key agreement and the intrinsic conditional information. *IEEE Transactions on Information Theory*, 45(2):499–514, March 1999.

[10] Matthias Christandl and Andreas Winter. Squashed entanglement: An additive entanglement measure. *Journal of Mathematical Physics*, 45(3):829–840, November 2004. arXiv:quant-ph/0308088.

[11] Matthias Christandl, Artur Ekert, Michał Horodecki, Paweł Horodecki, Jonathan Oppenheim, and Renato Renner. Unifying Classical and Quantum

Key Distillation. In *Theory of Cryptography*, pages 456–478. Springer Berlin Heidelberg, 2007. arXiv:quant-ph/0608199.

[12] Eneet Kaur, Xiaoting Wang, and Mark M. Wilde. Conditional mutual information and quantum steering. *Physical Review A*, 96(2):022332, August 2017. arXiv:1612.03875.

[13] Eneet Kaur, Mark M Wilde, and Andreas Winter. Fundamental limits on key rates in device-independent quantum key distribution. *New Journal of Physics*, 22(2):023039, February 2020.

[14] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[15] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, Second edition, 2017.

[16] N. J. Cerf and C. Adami. Negative entropy and information in quantum mechanics. *Physical Review Letters*, 79:5194–5197, December 1997.

[17] Elliott H. Lieb and Mary Beth Ruskai. Proof of the strong subadditivity of quantum-mechanical entropy. *Journal of Mathematical Physics*, 14(12):1938–1941, 1973.

[18] Hisaharu Umegaki. Conditional expectation in an operator algebra. iv. entropy and information. *Kodai Math. Sem. Rep.*, 14(2):59–85, 1962.

[19] Fumio Hiai and Dénes Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physics*, 143(1):99–114, 1991.

[20] Göran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40(2):147–151, 1975.

[21] A. Uhlmann. The "transition probability" in the state space of a $*$-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.

[22] Christopher A. Fuchs and Jeroen Van De Graaf. Cryptographic distinguishability measures for quantum mechanical states. *IEEE Transactions on Information Theory*, pages 45–1216, 1999.

[23] Artur K. Ekert. Quantum cryptography based on Bell's theorem. *Physical Review Letters*, 67:661–663, August 1991.

[24] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum metrology. *Physical Review Letters*, 96:010401, January 2006.

[25] Ryszard Horodecki, Pawel Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews of Modern Physics*, 81(2):865–942, June 2009. arXiv:quant-ph/0702225.

[26] Howard M. Wiseman, S. J. Jones, and Andrew C. Doherty. Steering, entanglement, nonlocality, and the Einstein-Podolsky-Rosen paradox. *Physical Review Letters*, 98(14):140402, May 2007. arXiv:quant-ph/0612147.

[27] Ivan Supic and Matty J Hoban. Self-testing through EPR-steering. *New Journal of Physics*, 18(7):075006, July 2016.

[28] Alexandru Gheorghiu, Elham Kashefi, and Petros Wallden. Robustness and device independence of verifiable blind quantum computing. *New Journal of Physics*, 17(8):083040, August 2015.

[29] Ana Belén Sainz, Nicolas Brunner, Daniel Cavalcanti, Paul Skrzypczyk, and Tamás Vértesi. Postquantum steering. *Physical Review Letters*, 115:190403, November 2015.

[30] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Reviews of Modern Physics*, 86(2):419–478, April 2014. arXiv:1303.2849.

[31] Daniel Rorhlich and Sandu Popescu. Quantum nonlocality as an axiom. *Foundations of Physics*, 24(3):379–385, March 1994. arXiv:quant-ph/9508009.

[32] Eric Chitambar and Gilad Gour. Quantum resource theories. *Reviews of Modern Phsyics*, 91:025001, April 2019.

[33] Andreas Winter and Dong Yang. Operational resource theory of coherence. *Physical Review Letters*, 116:120404, 2016.

[34] Alexander Streltsov, Gerardo Adesso, and Martin B. Plenio. Colloquium: Quantum coherence as a resource. *Reviews of Modern Phsyics*, 89:041003, October 2017.

[35] Xin Wang and Mark M. Wilde. Resource theory of asymmetric distinguishability. *Physical Review Research*, 1:033170, December 2019.

[36] Julio I. de Vicente. On nonlocality as a resource theory and nonlocality measures. *Journal of Physics A: Mathematical and Theoretical*, 47:424017, October 2014. arXiv:1401.6941.

[37] Rodrigo Gallego and Leandro Aolita. Resource theory of steering. *Physical Review X*, 5(4):041008, October 2015. arXiv:1409.5804.

[38] Cyril Branciard, Eric G. Cavalcanti, Stephen P. Walborn, Valerio Scarani, and Howard M. Wiseman. One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering. *Physical Review A*, 85:010301, January 2012. arXiv:1109.1435.

[39] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani. Device-Independent Security of Quantum Cryptography against Collective Attacks. *Physical Review Letters*, 98(23):230501, June 2007. arXiv:quant-ph/0702152.

[40] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54(5):3824–3851, November 1996. arXiv:quant-ph/9604024.

[41] Charles H. Bennett, David P. DiVincenzo, Christopher A. Fuchs, Tal Mor, Eric Rains, Peter W. Shor, John A. Smolin, and William K. Wootters. Quantum nonlocality without entanglement. *Physical Review A*, 59(2):1070–1091, 1999.

[42] Reinhard F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Physical Review A*, 40(8):4277–4281, October 1989.

[43] Michał Horodecki and Paweł Horodecki. Reduction criterion of separability and limits for a class of distillation protocols. *Physical Review A*, 59:4206–4216, June 1999. arXiv:quant-ph/9708015.

[44] Vlatko Vedral, Martin B. Plenio, M. A. Rippin, and Peter L. Knight. Quantifying entanglement. *Physical Review Letters*, 78(12):2275–2279, March 1997. arXiv:quant-ph/9702027.

[45] Guifré Vidal and Rolf Tarrach. Robustness of entanglement. *Physical Review A*, 59:141–155, January 1999.

[46] G. Vidal and R. F. Werner. Computable measure of entanglement. *Physical Review A*, 65:032314, February 2002.

[47] M. B. Plenio. Logarithmic negativity: A full entanglement monotone that is not convex. *Physical Review Letters*, 95:090503, August 2005.

[48] Yichen Huang. Computing quantum discord is NP-complete. *New Journal of Physics*, 16(3):033027, 2014.

[49] Leonid Gurvits. Classical deterministic complexity of edmonds problem and quantum entanglement. In *Proceedings of the thirty-fifth ACM symposium on Theory of computing - STOC'03*. ACM Press, 2003.

[50] Sevag Gharibian. Strong NP-hardness of the quantum separability problem. *Quantum Information & Computation*, 10(3&4):343–360, 2010.

[51] Erwin Schrödinger. Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(4):555–563, October 1935.

[52] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47(10):777–780, May 1935.

[53] Paul Skrzypczyk, Miguel Navascués, and Daniel Cavalcanti. Quantifying Einstein-Podolsky-Rosen steering. *Physical Review Letters*, 112:180404, May 2014.

[54] Marco Piani and John Watrous. Necessary and sufficient quantum information characterization of Einstein-Podolsky-Rosen steering. *Physical Review Letters*, 114:060404, February 2015.

[55] Eneet Kaur and Mark M Wilde. Relative entropy of steering: on its definition and properties. *Journal of Physics A: Mathematical and Theoretical*, 50(46):465301, November 2017. arXiv:1612.07152.

[56] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195–200, November 1964.

[57] Manuel Forster, Severin Winkler, and Stefan Wolf. Distilling nonlocality. *Physical Review Letters*, 102(12):120401, March 2009. arXiv:0809.3173.

[58] Manuel Forster and Stefan Wolf. Bipartite units of nonlocality. *Physical Review A*, 84(4):042112, October 2011. arXiv:0808.0651.

[59] Rodrigo Gallego and Leandro Aolita. Nonlocality free wirings and the distinguishability between Bell boxes. *Physical Review A*, 95(3):032118, March 2017. arXiv:1611.06932.

[60] Wim van Dam, Richard D. Gill, and Peter D. Grunwald. The statistical strength of nonlocality proofs. *IEEE Transactions on Information Theory*, 51(8):2812–2835, August 2005. arXiv:quant-ph/0307125.

[61] Marek Winczewski, Tamoghna Das, and Karol Horodecki. Upper bounds on secure key against non-signaling adversary via non-signaling squashed secrecy monotones. March 2019. arXiv:1903.12154.

[62] Eleni Diamanti, Hoi-Kwong Lo, Bing Qi, and Zhiliang Yuan. Practical challenges in quantum key distribution. *npj Quantum Information*, 2(1):16025, 2016.

[63] Hoi-Kwong Lo, Marcos Curty, and Kiyoshi Tamaki. Secure quantum key distribution. *Nature Photonics*, 8(8):595–604, 2014.

[64] Valerio Scarani, Helle Bechmann-Pasquinucci, Nicolas J. Cerf, Miloslav Dušek, Norbert Lütkenhaus, and Momtchil Peev. The security of practical quantum key distribution. *Reviews of Modern Phsyics*, 81:1301–1350, September 2009.

[65] Douglas Stebila, Michele Mosca, and Norbert Lütkenhaus. The case for quantum key distribution. In *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 283–296. Springer Berlin Heidelberg, 2010.

[66] Rotem Arnon-Friedman, Frédéric Dupuis, Omar Fawzi, Renato Renner, and Thomas Vidick. Practical device-independent quantum cryptography via entropy accumulation. *Nature Communications*, 9(1), 2018.

[67] Marco Tomamichel and Renato Renner. Uncertainty relation for smooth entropies. *Physical Review Letters*, 106:110506, March 2011. arXiv:1009.2015.

[68] Frédéric Grosshans, Nicolas J. Cerf, Jérôme Wenger, Rosa Tualle-Brouri, and Philippe Grangier. Virtual entanglement and reconciliation protocols for quantum cryptography with continuous variables. *Quantum Information and Computation*, 3:535–552, 2003.

[69] Renato Renner, Nicolas Gisin, and Barbara Kraus. Information-theoretic security proof for quantum-key-distribution protocols. *Physical Review A*, 72:012332, July 2005.

[70] H. Bechmann-Pasquinucci and N. Gisin. Incoherent and coherent eavesdropping in the six-state protocol of quantum cryptography. *Physical Review A*, 59:4238–4248, June 1999.

[71] Charles H. Bennett. Quantum cryptography using any two nonorthogonal states. *Physical Review Letters*, 68:3121–3124, May 1992.

[72] Michal Horodecki, Peter W. Shor, and Mary Beth Ruskai. Entanglement breaking channels. *Reviews in Mathematical Physics*, 15(6):629–641, 2003. arXiv:quant-ph/0302031.

[73] Nicolas Gisin and Stefan Wolf. Linking classical and quantum key agreement: Is there "bound information"? In *Advances in Cryptology — CRYPTO 2000*, pages 482–500. Springer Berlin Heidelberg, 2000.

[74] Nicolas Gisin, Renato Renner, and Stefan Wolf. Bound information: The classical analog to bound quantum entanglement. In *European Congress of Mathematics*, pages 439–447. Birkhäuser Basel, 2001.

[75] Sumeet Khatri and Norbert Lütkenhaus. Numerical evidence for bound secrecy from two-way postprocessing in quantum key distribution. *Physical Review A*, 95:042320, April 2017. arXiv:1612.07734.

[76] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15):880–884, October 1969.

[77] N. David Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Physical Review Letters*, 65:1838–1840, October 1990.

[78] Andrea Coladangelo. Generalization of the Clauser-Horne-Shimony-Holt inequality self-testing maximally entangled states of any local dimension. *Physical Review A*, 98:052115, November 2018.

[79] Jonathan Barrett, Lucien Hardy, and Adrian Kent. No signaling and quantum key distribution. *Physical Review Letters*, 95:010503, June 2005. arXiv:quant-ph/0405101.

[80] Lluís Masanes. Universally composable privacy amplification from causality constraints. *Physical Review Letters*, 102:140501, April 2009. arXiv:0807.2158.

[81] Lluís Masanes, Renato Renner, Matthias Christandl, Andreas Winter, and Jonathan Barrett. Full security of quantum key distribution from no-signaling constraints. *IEEE Transactions on Information Theory*, 60(8):4973–4986, August 2014. arXiv:quant-ph/0606049.

[82] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, 2015.

[83] Lynden K. Shalm, Evan Meyer-Scott, Bradley G. Christensen, Peter Bierhorst, Michael A. Wayne, Martin J. Stevens, Thomas Gerrits, Scott Glancy, Deny R. Hamel, Michael S. Allman, Kevin J. Coakley, Shellee D. Dyer, Carson Hodge, Adriana E. Lita, Varun B. Verma, Camilla Lambrocco, Edward Tortorici, Alan L. Migdall, Yanbao Zhang, Daniel R. Kumor, William H. Farr, Francesco Marsili, Matthew D. Shaw, Jeffrey A. Stern, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Thomas Jennewein, Morgan W. Mitchell, Paul G. Kwiat, Joshua C. Bienfang, Richard P. Mirin, Emanuel Knill, and Sae Woo Nam. Strong loophole-free test of local realism. *Physical Review Letters*, 115:250402, December 2015.

[84] Marissa Giustina, Marijn A. M. Versteegh, Sören Wengerowsky, Johannes Handsteiner, Armin Hochrainer, Kevin Phelan, Fabian Steinlechner, Johannes Kofler, January-Åke Larsson, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Morgan W. Mitchell, Jörn Beyer, Thomas Gerrits, Adriana E. Lita, Lynden K. Shalm, Sae Woo Nam, Thomas Scheidl, Rupert Ursin, Bernhard Wittmann, and Anton Zeilinger. Significant-Loophole-Free test of Bell's Theorem with Entangled Photons. *Physical Review Letters*, 115:250401, December 2015.

[85] Marcin Pawłowski and Nicolas Brunner. Semi-device-independent security of one-way quantum key distribution. *Physical Review A*, 84:010302, July 2011.

[86] Hoi-Kwong Lo, Marcos Curty, and Bing Qi. Measurement-Device-Independent Quantum Key Distribution. *Physical Review Letters*, 108:130503, 2012.

[87] Samuel L. Braunstein and Stefano Pirandola. Side-Channel-Free Quantum Key Distribution. *Physical Review Letters*, 108:130502, 2012.

[88] Marco Tomamichel, Serge Fehr, Jkedrzej Kaniewski, and Stephanie Wehner. One-sided device-independent QKD and position-based cryptography from monogamy games. In *Advances in Cryptology – EUROCRYPT 2013*, pages 609–625. Springer Berlin Heidelberg, 2013.

[89] Matthias Christandl, Robert König, Graeme Mitchison, and Renato Renner. One-and-a-Half Quantum de Finetti theorems. *Communications in Mathematical Physics*, 273(2):473–498, 2007.

[90] Matthias Christandl, Robert König, and Renato Renner. Postselection technique for quantum channels with applications to quantum cryptography. *Physical Review Letters*, 102:020504, January 2009.

[91] Rotem Arnon-Friedman, Renato Renner, and Thomas Vidick. Non-signaling parallel repetition using de Finetti reductions. *IEEE Transactions on Information Theory*, 62(3):1440–1457, 2016.

[92] Renato Renner and Stefan Wolf. New Bounds in Secret-Key Agreement: The Gap between Formation and Secrecy Extraction. In *Advances in Cryptology — EUROCRYPT 2003*, pages 562–577. Springer Berlin Heidelberg, May 2003.

[93] Amin A. Gohari and Venkat Anantharam. Information-theoretic key agreement of multiple terminals–Part 1. *IEEE Transactions on Information Theory*, 56(8):3973–3996, July 2010.

[94] Robert R. Tucci. Separability of density matrices and conditional information transmission. quant-ph/0005119v1.

[95] Robert R. Tucci. Entanglement of distillation and conditional mutual information. 2002. arXiv:quant-ph/0202144.

[96] M. Christandl, R. Renner, and S. Wolf. A property of the intrinsic mutual information. In *IEEE International Symposium on Information Theory, 2003. Proceedings.*, pages 258–258, June 2003.

[97] Ke Li and Andreas Winter. Squashed entanglement, k -extendibility, quantum markov chains, and recovery maps. *Foundations of Physics*, 48(8):910–924, 2018. arXiv:1410.4184.

[98] Fernando G. S. L. Brandão, Matthias Christandl, and Jon Yard. Faithful Squashed Entanglement. *Communications in Mathematical Physics*, 306(3):805–830, 2011.

[99] Matthias Christandl, Norbert Schuch, and Andreas Winter. Entanglement of the antisymmetric state. *Communications in Mathematical Physics*, 311(2):397–422, March 2012.

[100] Patrick Hayden, Richard Jozsa, Denes Petz, and Andreas Winter. Structure of states which satisfy strong subadditivity of quantum entropy with equality. *Communications in Mathematical Physics*, 246(2):359–374, April 2004. arXiv:quant-ph/0304007.

[101] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. *Physical Review A*, 53(4):2046–2052, April 1996. arXiv:quant-ph/9511030.

[102] Masahiro Takeoka, Saikat Guha, and Mark M. Wilde. Fundamental rate-loss tradeoff for optical quantum key distribution. *Nature Communications*, 5:5235, October 2014. arXiv:1504.06390.

[103] Antony Milne, Sania Jevtic, David Jennings, Howard Wiseman, and Terry Rudolph. Quantum steering ellipsoids, extremal physical states and monogamy. *New Journal of Physics*, 16(8):083017, August 2014. arXiv:1403.0418.

[104] Margaret D. Reid. Monogamy inequalities for the EPR paradox and quantum steering. *Physical Review A*, 88(6):062108, December 2013. arXiv:1310.2729.

[105] Masato Koashi and Andreas Winter. Monogamy of quantum entanglement and other correlations. *Physical Review A*, 69(2):022309, February 2004. arXiv:quant-ph/0310037.

[106] Omar Fawzi and Renato Renner. Quantum conditional mutual information and approximate Markov chains. *Communications in Mathematical Physics*, 340(2):575–611, September 2015. arXiv:1410.0664.

[107] Mark M. Wilde. Recoverability in quantum information theory. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 471(2182), October 2015. arXiv:1505.04661.

[108] Barbara M. Terhal, Andrew C. Doherty, and David Schwab. Symmetric Extensions of Quantum States and Local Hidden Variable Theories. *Physical Review Letters*, 90(15):157903, April 2003. arXiv:quant-ph/0210053.

[109] Alon Orlitsky and James R. Roche. Coding for computing. *IEEE Transactions on Information Theory*, 47(3):903–917, March 2001.

[110] M. Christandl and A. Winter. Uncertainty, monogamy, and locking of quantum correlations. *IEEE Transactions on Information Theory*, 51(9):3159–3165, 2005.

[111] Tzyh Haur Yang and Miguel Navascués. Robust self-testing of unknown quantum systems into any entangled two-qubit states. *Physical Review A*, 87:050102, May 2013. arXiv:1210.4409.

[112] Dominic Mayers and Andrew Yao. Self testing quantum apparatus. *Quantum Info. Comput.*, 4(4):273–286, 2004. arXiv:quant-ph/0307205.

[113] Itamar Pitowsky. The range of quantum probability. *Journal of Mathematical Physics*, 27(6):1556–1565, 1986.

[114] Jean-Daniel Bancal, Jonathan Barrett, Nicolas Gisin, and Stefano Pironio. Definitions of multipartite nonlocality. *Physical Review A*, 88:014102, July 2013.

[115] Satosi Watanabe. Information theoretical analysis of multivariate correlation. *IBM Journal of Research and Development*, 4(1):66–82, 1960.

[116] Te Sun Han. Linear dependence structure of the entropy space. *Information and Control*, 29(4):337–368, 1975.

[117] Jérémy Ribeiro, Gláucia Murta, and Stephanie Wehner. Fully device-independent conference key agreement. *Physical Review A*, 97:022307, February 2018.

[118] Maxim E. Shirokov. Tight continuity bounds for the quantum conditional mutual information, for the Holevo quantity and for capacities of quantum channels. *Journal of Mathematical Physics*, 58:102202, September 2017. arXiv:1512.09047.

[119] Joonwoo Bae and Antonio Acín. Key distillation from quantum channels using two-way communication protocols. *Physical Review A*, 75:012334, January 2007. arXiv:quant-ph/0610048.

[120] Shun Watanabe, Ryutaroh Matsumoto, Tomohiko Uyematsu, and Yasuhito Kawano. Key rate of quantum key distribution with hashed two-way classical communication. *Physical Review A*, 76:032312, September 2007. arXiv:0705.2904.

[121] Ernest Y.-Z. Tan, Charles C.-W. Lim, and Renato Renner. Advantage distillation for device-independent quantum key distribution. March 2019. arXiv:1903.10535.

[122] Ramij Rahaman, Matthew G. Parker, Piotr Mironowicz, and Marcin Pawłowski. Device-independent quantum key distribution based on measurement inputs. *Physical Review A*, 92:062304, December 2015. arXiv:1308.6447.

[123] Valerio Scarani, Antonio Acín, Grégoire Ribordy, and Nicolas Gisin. Quantum cryptography protocols robust against photon number splitting attacks for weak laser pulse implementations. *Physical Review Letters*, 92:057901, February 2004. arXiv:quant-ph/0211131.

[124] Karol Horodecki and Gláucia Murta. Bounds on quantum nonlocality via partial transposition. *Physical Review A*, 92(1):010301, July 2015. arXiv:1407.6999.

[125] Kenneth Goodenough, David Elkouss, and Stephanie Wehner. Assessing the performance of quantum repeaters for all phase-insensitive Gaussian bosonic channels. *New Journal of Physics*, 18:063005, June 2016. arXiv:1511.08710.

[126] Ryszard Horodecki, Pawel Horodecki, and Michal Horodecki. Violating Bell inequality by mixed spin-1/2 states: necessary and sufficient condition. *Physics Letters A*, 200(5):340–344, February 1995.

[127] Mark M. Wilde. Squashed entanglement and approximate private states. *Quantum Information Processing*, 15(11):4563–4580, November 2016. arXiv:1606.08028.

[128] Matty J Hoban and Ana Belén Sainz. A channel-based framework for steering, non-locality and beyond. *New Journal of Physics*, 20(5):053048, May 2018.

[129] Francesco Buscemi. All entangled quantum states are nonlocal. *Physical Review Letters*, 108:200401, May 2012.

[130] Eric G. Cavalcanti, Michael J. W. Hall, and Howard M. Wiseman. Entanglement verification and steering when Alice and Bob cannot be trusted. *Physical Review A*, 87:032306, 2013.

[131] Denis Rosset, David Schmid, and Francesco Buscemi. Characterizing nonclassicality of arbitrary distributed devices. November 2020. arXiv:1911.12462.

[132] Stefan Bäuml, Siddhartha Das, Xin Wang, and Mark M. Wilde. Resource theory of entanglement for bipartite quantum channels. July 2019. arXiv:1907.04181.

[133] Gilad Gour and Carlo Maria Scandolo. The entanglement of a bipartite channel. July 2019. arXiv:1907.02552.

[134] Siddhartha Das. *Bipartite quantum interactions: entangling and information processing abilities, LSU Dissertation*. PhD thesis, Department of Physics and Astronomy, LSU, 2019. arXiv:1901.05895.

[135] Siddhartha Das, Stefan Bäuml, and Mark M. Wilde. Entanglement and secret-key-agreement capacities of bipartite quantum interactions and read-only memory devices. *Physical Review A*, 101:012344, January 2020.

[136] Stefan Bäuml, Siddhartha Das, and Mark M. Wilde. Fundamental limits on the capacities of bipartite quantum interactions. *Physical Review Letters*, 121:250504, December 2018.

Vita

Eneet Kaur was born in the year 1991. She passed the AISSE examination (Xth board) from Mata Jai Kaur Public School, Delhi and CBSE 12th Examination (XIIth board) from KRM, Vikaspuri, Delhi. She enrolled for B.Sc. Physics in Hansraj College, Delhi University in May 2009, and received her Bachelors Degree in 2012. She then started Masters in Physics in IIT Roorkee in May 2012. She worked at Raman Research Institute, Bangalore as a Temporary Project Assistant from May 2014-2015. Later, she joined the Department of Physics and Astronomy at Louisiana State University, Baton Rouge for Ph.D. starting Fall 2015. She plans to graduate in May 2020. She has accepted a postdoctoral position in the group of Norbert Lütkenhaus at IQC Waterloo.