1-1-2020

Fundamental limits on key rates in device-independent quantum key distribution

Eneet Kaur
Louisiana State University

Mark M. Wilde
Louisiana State University

Andreas Winter
Universitat Autònoma de Barcelona

Recommended Citation

Kaur, E., Wilde, M., & Winter, A. (2020). Fundamental limits on key rates in device-independent quantum key distribution. *New Journal of Physics, 22* (2) https://doi.org/10.1088/1367-2630/ab6eaa

PAPER • OPEN ACCESS

Fundamental limits on key rates in device-independent quantum key distribution

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
24 August 2019

REVISED
2 January 2020

ACCEPTED FOR PUBLICATION
22 January 2020

PUBLISHED
27 February 2020

PAPER

Fundamental limits on key rates in device-independent quantum key distribution

Eneet Kaur[1], **Mark M Wilde**[1] **and Andreas Winter**[2]

[1] Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803, United States of America
[2] Departament de Física: Grup d'Informació Quántica, Universitat Autónoma de Barcelona, E-08193, Bellaterra (Barcelona), Spain

E-mail: mwilde@lsu.edu

Keywords: device independent quantum key distribution, intrinsic non-locality, one-sided device independent quantum key distribution, intrinsic steerability, Bell non-locality

Abstract

In this paper, we introduce intrinsic non-locality and quantum intrinsic non-locality as quantifiers for Bell non-locality, and we prove that they satisfy certain desirable properties such as faithfulness, convexity, and monotonicity under local operations and shared randomness. We then prove that intrinsic non-locality is an upper bound on the secret-key-agreement capacity of any device-independent protocol conducted using a device characterized by a correlation p, while quantum intrinsic non-locality is an upper bound on the same capacity for a correlation arising from an underlying quantum model. We also prove that intrinsic steerability is faithful, and it is an upper bound on the secret-key-agreement capacity of any one-sided-device-independent protocol conducted using a device characterized by an assemblage $\hat{\rho}$. Finally, we prove that quantum intrinsic non-locality is bounded from above by intrinsic steerability.

1. Introduction

In principle quantum key distribution (QKD) [BB84, Eke91, SBPC+09] provides unconditional security [May01, SP00, LCT14] for establishing secret key at a distance. In the standard QKD setting, Alice and Bob (two spatially separated parties) trust the functioning of their devices. That is, it is assumed that they know the ensemble of states that their sources are preparing and the measurements that their devices are performing. However, this is a very strong set of assumptions.

It is possible to consider other scenarios in which the trust assumptions are relaxed while still obtaining unconditional security. When one of the devices is untrusted, the protocol is referred to as one-sided-device-independent (SDI) QKD [TR11, BCW+12]. If both the devices are untrusted, then we are dealing with the scenario of device-independent (DI) QKD [MY98, ABG+07, VV14, AFDF+18].

It is interesting to note that the three scenarios of QKD mentioned above are in correspondence with a hierarchy of quantum correlations [WJD07]. The standard QKD approach requires that Alice and Bob share entanglement [HHHH09] or that they are connected by a channel that can preserve entanglement. In SDI-QKD, a requirement for Alice and Bob to generate secret key is that their systems violate a steering inequality [BCW+12, CS17]. For DI-QKD, Alice and Bob's systems should violate a Bell inequality [CHSH69, ABG+07, BCP+14].

In this paper, we establish upper bounds on secret-key rates that are achievable with DI-QKD and SDI-QKD. To this end, we first introduce intrinsic non-locality and quantum intrinsic non-locality as quantifiers of non-local correlations. 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. We also provide a proof for faithfulness of restricted intrinsic steerability, a quantifier of quantum steering introduced in [KWW17], thus solving an open question from [KWW17].

Next, we consider a device that is characterized by a correlation p, and we allow Alice and Bob to perform local operations and public communication on its inputs and outputs (this contains the parameter estimation, error correction, and privacy amplification) to extract a secret key from this device. Then, we prove that intrinsic non-locality is an upper bound on the rate at which secret key can be extracted from this device, such that the secret key is protected from a third party possessing an arbitrary no-signaling extension of the correlation, as well as copies of all of the classical data publicly exchanged in the protocol. We do the same for quantum intrinsic non-locality and a third party possessing an arbitrary quantum extension of the correlation.

We then consider a device that is characterized by an assemblage $\hat{\rho}$ and prove that restricted intrinsic steerability is an upper bound on the rate at which secret key can be extracted from this device, such that the secret key is protected from a third party possessing an arbitrary no-signaling extension of the assemblage (as considered in [KWW17]), as well as copies of all of the classical data publicly exchanged during the protocol.

The present work is inspired by [MW99], which introduced intrinsic information and proved that it is an upper bound on the distillable secret key for Alice and Bob protected from an adversary Eve, such that Alice has access to a random variable X, Bob to a random variable Y, and Eve to Z, such that the joint distribution is P_{XYZ}. Later, [CW04], taking inspiration from the underlying idea of intrinsic information, defined squashed entanglement as a quantum version of the former, which turns out to be an entanglement measure with many desirable properties. (See also [Tuc02] for discussions related to squashed entanglement.) The squashed entanglement was later established as an upper bound on the distillable secret key of a bipartite quantum state [CEH+07] (see also [Wil16] in this context). The squashed entanglement of a channel was later defined and proved to be an upper bound on the secret-key-agreement capacity of a quantum channel [TGW14, Wil16]. Both the intrinsic steerability from [KWW17], intrinsic non-locality, and the quantum intrinsic non-locality defined here are strongly related to these previous quantities. It is fair to say that intrinsic non-locality is closest in spirit to [MW99], in that, it is defined entirely in terms of classical random variables accessible to Alice and Bob.

This paper is structured as follows: we recall the definition of restricted intrinsic steerability in section 2. We then introduce intrinsic non-locality and quantum intrinsic non-locality, and we analyze its mathematical properties in section 3. In section 4, we provide a proof for the faithfulness of intrinsic steerability. Section 5 provides a proof for the faithfulness of intrinsic non-locality. We prove upper bounds on secret-key-agreement capacities for device-independent and one-SDI protocols in section 6. In section 7, we showcase our bounds for some specific examples, thus obtaining explicit bounds on the secret-key rate that can be obtained from specific bipartite correlations studied in the device-independent literature. We end with section 8, where we conclude and discuss some open questions.

Note: after posting the first version of this paper to the arXiv, we became aware of related results presented in [WDH19]. In [WDH19], squashed non-locality was introduced as a measure of non-locality. This quantity was then proven to be an upper bound on device-independent secret key rates that are secure against a no-signaling adversary with classical inputs and outputs. This is the scenario in which an untrusted no-signaling device is shared by the honest parties (as in our model), while the inputs and outputs of an adversary Eve are assumed to satisfy only the no-signaling conditions. The latter is more 'liberal' than the models considered in our work.

2. Restricted intrinsic steerability

In this section, we recall the definition of restricted intrinsic steerability, which was introduced in [KWW17]. We begin by recalling the notion of an assemblage. Let ρ_{AB} be a bipartite quantum state shared by Alice and Bob. Suppose that Alice performs a measurement labeled by $x \in \mathcal{X}$, with $\mathcal{X}$ denoting a finite set of quantum measurement choices, and she gets a classical output $a \in \mathcal{A}$, with $\mathcal{A}$ denoting a finite set of measurement outcomes. An *assemblage* [Pus13] 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 $\hat{\rho}_B^{a,x} := p_{\bar{A}|X}(a|x)\rho_B^{a,x}$. 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:

$$\rho_{X\bar{A}B} := \sum_{a,x} p_X(x)[x\ a]_{X\bar{A}} \otimes \hat{\rho}_B^{a,x}. \tag{1}$$

Notation 1. In the above and what follows, we employ the shorthand $[x\ a]_{X\bar{A}}$ to denote $|x\rangle\langle x|_X \otimes |a\rangle\langle a|_{\bar{A}}$.

Assemblages are restricted by the no-signaling principle. That is, the reduced state of Bob's system should not depend on the input x to Alice's black box if the measurement output a is not available to him:

$$\sum_a \hat{\rho}_B^{a,x} = \sum_a \hat{\rho}_B^{a,x'} \quad \forall x, x' \in \mathcal{X}. \tag{2}$$

This is equivalent to $I(X;B)_\rho = 0$ for all input probability distributions $p_X(x)$, where $I(X;B)_\rho := H(X)_\rho + H(B)_\rho - H(XB)_\rho$ is the mutual information of the reduced state $\rho_{XB} = \mathrm{Tr}_{\bar{A}}(\rho_{X\bar{A}B})$.

An assemblage is referred to as local-hidden-state (LHS) if it arises from a classical shared hidden variable Λ in the following sense:

$$\hat{\rho}_B^{a,x} := \sum_\lambda p_\Lambda(\lambda) p_{\bar{A}|X\Lambda}(a|x,\lambda)\rho_B^\lambda. \tag{3}$$

We now recall a measure of steerability that was introduced in [KWW17]:

Definition 2 (**Restricted intrinsic steerability** [KWW17]). 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 no-signaling extension $\rho_{X\bar{A}BE}$ of $\rho_{X\bar{A}B}$ of the following form:

$$\rho_{X\bar{A}BE} := \sum_{a,x} p_X(x)[x\ a]_{X\bar{A}} \otimes \hat{\rho}_{BE}^{a,x}, \tag{4}$$

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:

$$\sum_a \hat{\rho}_{BE}^{a,x} = \sum_a \hat{\rho}_{BE}^{a,x'} \quad \forall x, x' \in \mathcal{X}. \tag{5}$$

The restricted intrinsic steerability of $\{\hat{\rho}_B^{a,x}\}_{a,x}$ is defined as follows:

$$S(\bar{A};B)_{\hat{\rho}} := \sup_{p_X} \inf_{\rho_{X\bar{A}BE}} I(X\bar{A};B|E)_\rho, \tag{6}$$

where the supremum is with respect to all probability distributions p_X and the infimum is with respect to all no-signaling extensions of $\rho_{X\bar{A}B}$ as specified above. Furthermore, the conditional mutual information of a tripartite state σ_{KLM} is defined as

$$I(K;L|M)_\sigma := H(KM)_\sigma + H(LM)_\sigma - H(M)_\sigma - H(KLM)_\sigma. \tag{7}$$

Using the no-signaling constraints, which imply that $I(X;B|E)_\rho = 0$, and the chain rule for conditional mutual information, it follows that

$$S(\bar{A};B)_{\hat{\rho}} := \sup_{p_X} \inf_{\rho_{X\bar{A}BE}} I(\bar{A};B|EX)_\rho. \tag{8}$$

3. Quantum non-locality

3.1. Correlations

Consider a two-component device that takes in two inputs and gives out two outputs. Let one component be with Alice and the other component be with Bob. Let us set some notation now. 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 $b \in \mathcal{B}$. We consider $\mathcal{X}$ and $\mathcal{Y}$ to be finite sets of quantum measurement choices and $\mathcal{A}$ and $\mathcal{B}$ to be finite sets of measurement outcomes. For simplicity, we consider $\mathcal{X} = \mathcal{Y} = [s]$ and $\mathcal{A} = \mathcal{B} = [r]$. The conditional probability distribution $\{p(a,b|x,y)\}_{a,b\in[r],x,y\in[s]}$ corresponding to the device is traditionally called a 'correlation.' Then the correlations can be divided as follows according to the constraints that they fulfill:

- *Local correlations*: A correlation is said to have a local hidden variable (LHV) description or be a local correlation if it can be written as

$$p(a,b|x,y) = \sum_\lambda p_\Lambda(\lambda) p(a|x,\lambda) p(b|y,\lambda), \tag{9}$$

  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 correlations that can be written as in (9). A device characterized by local correlations is known as a local box.

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

$$p(a,b|x,y) = \mathrm{Tr}([\Lambda_x^a \otimes \Lambda_y^b]\rho_{AB}), \tag{10}$$

  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 $\Lambda_x^a, \Lambda_y^b \geqslant 0$ 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 correlations*: The set *NS* corresponds to the set of correlations that fulfill the following no-signaling principle:

$$\sum_a p(a, b|x, y) = \sum_a p(a, b|x', y) = p(b|y), \quad \forall x, x' \in [s] \text{ and } \quad b \in [r], y \in [s]. \tag{11}$$

$$\sum_b p(a, b|x, y) = \sum_b p(a, b|x, y') = p(a|x), \quad \forall y, y' \in [s] \text{ and } \quad a \in [r], x \in [s]. \tag{12}$$

The no-signaling constraints (11) and (12) can be expressed equivalently in terms of conditional mutual informations, namely

$$\forall p(x, y) \;\; I(X; \bar{B}|Y)_p = 0 = I(Y; \bar{A}|X)_p, \tag{13}$$

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 contained in the set of quantum correlations, that is, $\mathbf{L} \subset \mathbf{Q}$. Since the correlations in $\mathbf{Q}$ fulfill the constraints in (11) and (12), we have that $\mathbf{Q} \subset \mathbf{NS}$. For more details on correlations, please refer to [BCP+14].

An example of a correlation that belongs to the set of no-signaling correlations, but not to the set of quantum correlations, is a Popescu–Rohrlich (PR) box [RP94] box, which is defined as follows:

Definition 3 (PR box). A PR box is a device corresponding to the following correlation $p(a, b|x, y)$:

$$\begin{aligned} p(0, 0|x, y) = p(1, 1|x, y) = \frac{1}{2} \quad \text{for} \quad (x, y) \neq (1, 1), \\ p(0, 1|x, y) = p(1, 0|x, y) = \frac{1}{2} \quad \text{for} \quad (x, y) = (1, 1), \end{aligned} \tag{14}$$

while $p(a, b|x, y) = 0$ for all other quadruples. This correlation is no-signaling between Alice and Bob, as defined in (11) and (12).

3.2. Local operations and shared randomness

Physically, local operations and shared randomness [FWW09, FW11] refers to an operation in which Alice and Bob share unlimited free randomness between their two components 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 [GA17]:

$$p_f(a_f, b_f|x_f, y_f) := \sum_{a,b,x,y} O^{(L)}(a_f, b_f|a, b, x, y, x_f, y_f)p_i(a, b|x, y)I^{(L)}(x, y|x_f, y_f). \tag{15}$$

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

$$I^{(L)}(x, y|x_f, y_f) = \sum_{\lambda_2} p_{\Lambda_2}(\lambda_2)I_A(x|x_f, \lambda_2)I_B(y|y_f, \lambda_2), \tag{16}$$

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

$$O^{(L)}(a_f, b_f|a, b, x, y, x_f, y_f) = \sum_{\lambda_1} p_{\Lambda_1}(\lambda_1)O_A(a_f|a, x, x_f, \lambda_1)O_B(b_f|b, y, y_f, \lambda_1). \tag{17}$$

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 1 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 [dV14, GA17], the resources are non-local correlations $p(a, b|x, y)$. Local operations and shared randomness are one possible set of free operations in this resource

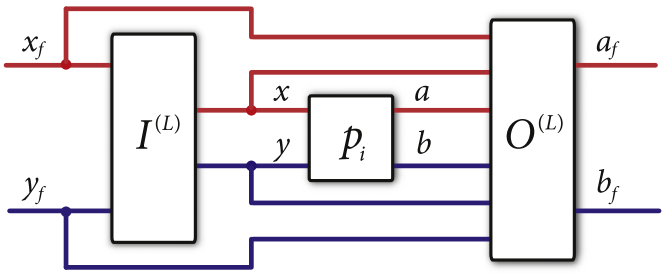

Figure 1. This figure depicts how local operations and shared randomness can act on an initial correlation $p_i(a, b|x, y)$ to produce a final correlation $p_f(a_f, b_f|x_f, y_f)$.

theory [dV14]. It can be shown from the definition of a local correlation, that the action of the local operations and share randomness transforms a local correlation to a correlation in **L**. Furthermore, a quantum correlation 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 (15) by separable states shared between Alice and Bob with the local states encoding the probability distributions required in (16) and (17) and the measurements as projective measurements.

In [GA17], a larger set of free operations known as wirings and prior-to-input classical communication (WPICC) was considered. It was also shown in lemma 6 of [GA17] that any quantifier that is a monotone under local operations and shared randomness is also a monotone under WPICCs.

3.3. Intrinsic non-locality

To calculate the amount of non-locality present in the correlation $p(a, b|x, y)$, we introduce a function N: $p(a, b|x, y) \to \mathbb{R}_{\geqslant 0}$, which we call *intrinsic non-locality*. Consider a correlation $p(a, b|x, y) \in$ **NS**. Now embed the correlation $p(a, b|x, y)$ into a classical–classical state as

$$\rho_{\bar{A}\bar{B}XY} := \sum_{a,b,x,y} p(x, y)p(a, b|x, y)[a\ b\ x\ y]_{\bar{A}\bar{B}XY}, \tag{18}$$

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 $\rho_{\bar{A}\bar{B}XY}$:

$$\rho_{\bar{A}\bar{B}XYE} := \sum_{a,b,x,y} p(x, y)[a\ b\ x\ y]_{\bar{A}\bar{B}XY} \otimes p(a, b|x, y)\rho_E^{a,b,x,y}, \tag{19}$$

such that $\mathrm{Tr}_E(\rho_{\bar{A}\bar{B}XYE}) = \rho_{\bar{A}\bar{B}XY}$, and the following no-signaling constraints hold:

$$\sum_a p(a, b|x, y)\rho_E^{a,b,x,y} = \sum_a p(a, b|x', y)\rho_E^{a,b,x',y} \quad \forall x, x' \in \mathcal{X}. \tag{20}$$

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:

$$I(\bar{B}E; X|Y)_\rho = 0 \quad \forall p(x, y). \tag{21}$$

Similarly, the following no-signaling constraints hold

$$\sum_b p(a, b|x, y)\rho_E^{a,b,x,y} = \sum_b p(a, b|x, y')\rho_E^{a,b,x,y'} \quad \forall y, y' \in \mathcal{Y}. \tag{22}$$

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

$$I(\bar{A}E; Y|X)_\rho = 0 \quad \forall p(x, y). \tag{23}$$

Finally, we have that

$$\sum_{a,b} p(a, b|x, y)\rho_E^{a,b,x,y} = \sum_{a,b} p(a, b|x', y)\rho_E^{a,b,x',y} \tag{24}$$

$$= \sum_{a,b} p(a, b|x', y')\rho_E^{a,b,x',y'} \quad \forall x, x' \in \mathcal{X}, y, y' \in \mathcal{Y}. \tag{25}$$

The first equality follows from (20), and the second equality follows from (22). 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 correlations in the correlation $p(a, b|x, y)$ 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 correlations of the two black boxes, we maximize over input probability distributions $p(x, y)$, leading us to the following definition:

Definition 4 (**Intrinsic non-locality**). The intrinsic non-locality of a correlation $p(a, b|x, y) \in \mathbf{NS}$ is defined as

$$N(\bar{A}; \bar{B})_p = \sup_{p(x,y)} \inf_{\rho_{\bar{A}\bar{B}XYE}} I(\bar{A}; \bar{B}|XYE)_\rho, \tag{26}$$

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 (20) and (22).

3.4. Quantum intrinsic non-locality

We now introduce a function N^Q: $p(a, b|x, y) \to \mathbb{R}_{\geqslant 0}$, which we call *quantum intrinsic non-locality*, with $p(a, b|x, y) \in \mathbf{Q}$. As stated above, the correlation 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[4]. 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 (18) is

$$\rho_{\bar{A}\bar{B}XYE} = \sum_{a,b,x,y} p(x, y)\mathrm{Tr}_{AB}[(\Lambda_x^a \otimes \Lambda_y^b \otimes \mathcal{I}_E)\rho_{ABE}][a\ b\ x\ y]_{\bar{A}\bar{B}XY}, \tag{27}$$

$$= \sum_{a,b,x,y} p(x, y)p(a, b|x, y)[a\ b\ x\ y]_{\bar{A}\bar{B}XY} \otimes \rho_E^{a,b,x,y}, \tag{28}$$

where $p(a, b|x, y)\rho_E^{a,b,x,y} := \mathrm{Tr}_{AB}[(\Lambda_x^a \otimes \Lambda_y^b \otimes \mathcal{I}_E)\rho_{ABE}]$. By definition, this extension is also a no-signaling extension and is subjected to the constraints in (20) and (22). We call the extensions of the form in (27) *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 correlation $p(a, b|x, y)$ 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. Whereas, the no-signaling extensions of such a correlation 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 LHV).

Therefore, to consider the regime in which there is an underlying quantum model, we define *quantum intrinsic non-locality* as follows:

Definition 5 (**Quantum intrinsic non-locality**). The quantum intrinsic non-locality of a correlation $p(a, b|x, y) \in \mathbf{Q}$ is defined as

$$N^Q(\bar{A}; \bar{B})_p = \sup_{p(x,y)} \inf_{\rho_{\bar{A}\bar{B}XYE}} I(\bar{A}; \bar{B}|XYE)_\rho, \tag{29}$$

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 (27).

Proposition 6. *If* $p(a, b|x, y) \in \mathbf{Q}$, *then*

$$N(\bar{A}; \bar{B})_p \leqslant N^Q(\bar{A}; \bar{B})_p. \tag{30}$$

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. ■

3.5. 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 correlations. These are the properties that are desirable for a measure of Bell non-locality to possess. We also prove that quantum intrinsic non-locality of a correlation is never larger than the intrinsic steerability of an associated assemblage.

Proposition 7. *Intrinsic non-locality and quantum intrinsic non-locality vanish for correlations 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

$$p(a, b|x, y) = \sum_\lambda p(\lambda)\, p(a|x, \lambda)\, p(b|y, \lambda). \tag{31}$$

[4] For certain quantum correlations, it is possible to pinpoint the underlying quantum state and POVMs up to certain isometries. See [YN13, MY04] in this context.

Embed this in a classical–classical state with $p(x, y)$ an arbitrary probability distribution over x, y:

$$\rho_{\bar{A}\bar{B}XY} = \sum_{a,b,x,y} p(x, y) \sum_{\lambda} p(\lambda)\, p(a|x, \lambda) p(b|y, \lambda) [a\ b\ x\ y]_{\bar{A}\bar{B}XY}. \tag{32}$$

Then, consider the following quantum extension

$$\rho_{\bar{A}\bar{B}XYE} := \sum_{a,b,x,y} p(x, y) [a\ b\ x\ y]_{\bar{A}\bar{B}XY} \otimes \sum_{\lambda} p(\lambda)\, p(a|x, \lambda) p(b|y, \lambda)\, [\lambda]_E. \tag{33}$$

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 (30) we conclude that $N(\bar{A}; \bar{B})_p = 0$. ■

We later prove in theorem 19 that $N(\bar{A}; \bar{B})_p = 0$ or $N^Q(\overline{A}; \overline{B})_p = 0$ implies that $p \in \mathbf{L}$.

We expect any quantifier of non-locality to be monotone under the free operations of 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 8 (**Monotonicity of intrinsic non-locality**). *Let $p_i(a, b|x, y)$ be a correlation, and let $p_f(a_f, b_f|x_f, y_f)$ be a correlation 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:*

$$p_f(a_f, b_f|x_f, y_f) := \sum_{a,b,x,y} O^{(L)}(a_f, b_f|a, b, x, y, x_f, y_f)\, p_i(a, b|x, y) I^{(L)}(x, y|x_f, y_f), \tag{34}$$

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 (16) and (17). Then,

$$N(\bar{A}; \bar{B})_{p_i} \geqslant N(\bar{A}_f; \bar{B}_f)_{p_f}. \tag{35}$$

Proof. First, we embed $p_f(a_f, b_f|x_f, y_f)$ in a quantum state:

$$\rho_{\bar{A}_f\bar{B}_f X_f Y_f} = \sum_{x_f,y_f,a_f,b_f} p(x_f, y_f) p_f(a_f, b_f|x_f, y_f) [x_f\ y_f\ a_f\ b_f]_{X_f Y_f \bar{A}_f \bar{B}_f}, \tag{36}$$

where $p(x_f, y_f)$ is an arbitrary probability distribution for x_f, y_f. Then invoking (15)–(17), we obtain

$$\begin{aligned}\rho_{\bar{A}_f\bar{B}_f X_f Y_f} = \sum_{x_f,y_f,a_f,b_f} p(x_f, y_f) \sum_{a,b,x,y} \sum_{\lambda_2} p_{\Lambda_2}(\lambda_2) O_A(a_f|a, x_f, x, \lambda_2) O_B(b_f|b, y, y_f, \lambda_2) \\ \times p_i(a, b|x, y) \sum_{\lambda_1} p_{\Lambda_1}(\lambda_1) I_A(x|x_f, \lambda_1) I_B(y|y_f, \lambda_1) [x_f\ y_f\ a_f\ b_f]_{X_f Y_f \bar{A}_f \bar{B}_f}.\end{aligned} \tag{37}$$

An arbitrary extension of the state in (36) is given by

$$\rho_{\bar{A}_f\bar{B}_f X_f Y_f E} = \sum_{x_f,y_f,a_f,b_f} p(x_f, y_f) p_f(a_f, b_f|x_f, y_f) [x_f\ y_f\ a_f\ b_f]_{X_f Y_f \bar{A}_f \bar{B}_f} \otimes \rho_E^{a_f,b_f,x_f,y_f}. \tag{38}$$

A particular extension of the state in (36) is given by

$$\begin{aligned}\zeta_{\bar{A}_f\bar{B}_f X_f Y_f E\Lambda_1\Lambda_2} = \sum_{x_f,y_f,a_f,b_f} p(x_f, y_f) \sum_{a,b,x,y} \sum_{\lambda_2} p_{\Lambda_2}(\lambda_2) O_A(a_f|a, x_f, x, \lambda_2) \\ \times O_B(b_f|b, y, y_f, \lambda_2) p_i(a, b|x, y) \\ \times \sum_{\lambda_1} p_{\Lambda_1}(\lambda_1) I_A(x|x_f, \lambda_1) I_B(y|y_f, \lambda_1) [x_f\ y_f\ a_f\ b_f]_{\bar{A}_f\bar{B}_f X_f Y_f} \otimes \tau_E^{a,b,x,y} \otimes [\lambda_1\lambda_2]_{\Lambda_1\Lambda_2}.\end{aligned} \tag{39}$$

This in turn is a marginal of the following state:

$$\begin{aligned}\zeta_{\bar{A}_f\bar{B}_f X_f Y_f E\Lambda_1\Lambda_2 XY\bar{A}\bar{B}} = \sum_{x_f,y_f,a_f,b_f} p(x_f, y_f) \sum_{a,b,x,y} \sum_{\lambda_2} p_{\Lambda_2}(\lambda_2) O_A(a_f|a, x_f, x, \lambda_2) O_B(b_f|b, y, y_f, \lambda_2) \\ \times p_i(a, b|x, y) \sum_{\lambda_1} p_{\Lambda_1}(\lambda_1) I_A(x|x_f, \lambda_1) I_B(y|y_f, \lambda_1) [x_f\ y_f\ a_f\ b_f]_{X_f Y_f \bar{A}_f \bar{B}_f} \\ \otimes \tau_E^{a,b,x,y} \otimes [\lambda_1\lambda_2]_{\Lambda_1\Lambda_2} \otimes [x\ y\ a\ b]_{XY\bar{A}\bar{B}}.\end{aligned} \tag{40}$$

Consider that

$$\inf_{\text{ext.in (38)}} I(\bar{A}_f; \bar{B}_f|X_f Y_f E)_\rho \leqslant I(\bar{A}_f; \bar{B}_f|X_f Y_f E\Lambda_1\Lambda_2)_\zeta \tag{41}$$

$$\leqslant I(\bar{A}X_f X\Lambda_2; \bar{B}Y_f Y\Lambda_2|X_f Y_f E\Lambda_1\Lambda_2)_\zeta \tag{42}$$

$$= I(\bar{A}X; \bar{B}Y|X_f Y_f E\Lambda_1\Lambda_2)_\zeta \tag{43}$$

$$= I(\bar{A}X;\ \bar{B}Y|X_f Y_f E\Lambda_1)_\zeta \tag{44}$$

$$= I(\bar{A};\ \bar{B}|XYX_f Y_f E\Lambda_1)_\zeta + I(X;\ \bar{B}|X_f Y_f E\Lambda_1 Y)_\zeta + I(Y;\ \bar{A}|X_f Y_f E\Lambda_1 X)_\zeta + I(X;\ Y|X_f Y_f \Lambda_1 E)_\zeta. \tag{45}$$

The first inequality follows from considering a particular extension in (39). The second inequality follows from data processing of conditional mutual information. The second equality follows because $\zeta_{\bar{A}\bar{B}XYX_f Y_f E\Lambda_1\Lambda_2} = \zeta_{\bar{A}\bar{B}XYX_f Y_f E\Lambda_1} \otimes \zeta_{\lambda_2}$. The last equality follows from the chain rule for conditional mutual information. Now, let us consider each term in (45). By inspection,

$$\zeta_{\bar{A}\bar{B}XYX_f Y_f E\Lambda_1} = \sum_{x_f,y_f} p(x_f,\ y_f) \sum_{a,b,x,y,\lambda} p(\lambda_1) p_i(a,\ b|x,\ y) p(x,\ y|x_f,\ y_f,\ \lambda_1)$$
$$[x_f\ y_f\ \lambda_1\ x\ y\ a\ b]_{X_f Y_f \Lambda_1 XY\bar{A}\bar{B}} \otimes \tau_E^{a,b,x,y}. \tag{46}$$

Upon re-arranging, we obtain

$$\zeta_{\bar{A}\bar{B}XYX_f Y_f E\Lambda_1} = \sum_{x,y} p(x,\ y) \sum_{x_f,y_f,\lambda_1} p(x_f,\ y_f,\ \lambda_1|x,\ y)[x\ y\ x_f\ y_f\ \lambda_1]_{XYX_f Y_f \Lambda_1}$$
$$\otimes \sum_{a,b} p_i(a,\ b|x,\ y)\tau_E^{a,b,x,y} \otimes [a\ b]_{\bar{A}\bar{B}}. \tag{47}$$

So, given $X,\ Y$, the states $\zeta^{x,y}_{\bar{A}\bar{B}E}$ and $\zeta^{x,y}_{X_f Y_f \Lambda_1}$ are in tensor product. Therefore $I(\bar{A};\ \bar{B}|XYX_f Y_f E\Lambda_1)_\zeta = I(\bar{A};\ \bar{B}|XYE)_\zeta$, where $\zeta_{\bar{A}\bar{B}XYE}$ is a no-signaling extension of $\rho_{\bar{A}\bar{B}XY}$. Now consider that

$$\zeta_{XX_f YY_f \bar{B}E\Lambda_1}$$
$$= \sum_{x,y,x_f,y_f,\lambda_1} p(x,\ y,\ x_f,\ y_f,\ \lambda_1)[x\ x_f\ y\ y_f\ \lambda_1]_{XX_f YY_f \Lambda_1} \otimes \sum_b p(b|y)\tau_E^{b,y} \otimes [b]_{\bar{B}}. \tag{48}$$

$$= \sum_y p(y)[y]_Y \otimes \sum_{x,x_f,y_f,\lambda_1} p(x_f,\ y_f,\ x,\ \lambda_1|y)[x\ x_f\ y\ y_f\ \lambda_1]_{XX_f YY_f \Lambda_1} \otimes \sum_b p(b|y)\tau_E^{b,y} \otimes [b]_{\bar{B}}. \tag{49}$$

Then, by inspection

$$I(X;\ \bar{B}|X_f Y_f E\Lambda_1 Y)_\zeta = 0. \tag{50}$$

Similarly, $I(Y;\ \bar{A}|Y_f X_f E\Lambda_1 X)_\zeta = 0$.

Now, consider the term $I(X;\ Y|X_f Y_f E\Lambda_1)_\zeta$, with

$$\zeta_{XYX_f Y_f E\Lambda_1} := \sum_{x_f,y_f} p(x_f,\ y_f) \sum_{x,y,\lambda_1} p(x|x_f,\ \lambda_1) p(y|y_f,\ \lambda_1)[x\ y\ x_f\ y_f\ \lambda_1]_{XYX_f Y_f \Lambda_1} \otimes \rho_E. \tag{51}$$

Here, X and Y are independent given X_f, Y_f, and Λ_1. Therefore, $I(X;\ Y|X_f Y_f E\Lambda_1)_\zeta = 0$. Combining the above equations, we obtain

$$\inf_{\text{ext.in (38)}} I(\bar{A}_f;\ \bar{B}_f|X_f Y_f E)_\rho \leqslant I(\bar{A};\ \bar{B}|XYE)_\zeta. \tag{52}$$

Since (52) 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

$$N(\bar{A}_f;\ \bar{B}_f)_{p_f} \leqslant N(\bar{A};\ \bar{B})_{p_i}. \tag{53}$$

This concludes the proof. ■

Proposition 9 (**Monotonicity of quantum intrinsic non-locality**). *Let $p_i(a,\ b|x,\ y) \in \mathbf{Q}$, and let $p_f(a_f,\ b_f|x_f,\ y_f)$ result 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:*

$$p_f(a_f,\ b_f|x_f,\ y_f) := \sum_{a,b,x,y} O^{(L)}(a_f,\ b_f|a,\ b,\ x,\ y,\ x_f,\ y_f)\ p_i(a,\ b|x,\ y) I^{(L)}(x,\ y|x_f,\ y_f), \tag{54}$$

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 (16) and (17). Then,

$$N^Q(\bar{A};\ \bar{B})_{p_i} \geqslant N^Q(\bar{A}_f;\ \bar{B}_f)_{p_f}. \tag{55}$$

Proof. First, we embed $p_f(a_f,\ b_f|x_f,\ y_f)$ in a quantum state:

$$\rho_{\bar{A}_f \bar{B}_f X_f Y_f} = \sum_{x_f,y_f,a_f,b_f} p(x_f,\ y_f) p_f(a_f,\ b_f|x_f,\ y_f)[x_f\ y_f\ a_f\ b_f]_{X_f Y_f \bar{A}_f \bar{B}_f}, \tag{56}$$

where $p(x_f,\ y_f)$ is an arbitrary probability distribution for $x_f,\ y_f$. The set of quantum correlations $\mathbf{Q}$ 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 correlation, we know that there exists an underlying state σ_{AB} and POVMs

$\{\Lambda_{x_f}^{a_f}\}_{a_f}$ and $\{\Lambda_{y_f}^{b_f}\}_{b_f}$, such that

$$p_f(a_f, b_f|x_f, y_f) = \mathrm{Tr}[(\Lambda_{x_f}^{a_f} \otimes \Lambda_{y_f}^{b_f})\sigma_{AB}]. \tag{57}$$

An arbitrary quantum extension of the state in (56) is given by

$$\sigma_{\bar{A}_f\bar{B}_f X_f Y_f E} = \sum_{x_f,y_f,a_f,b_f} p(x_f, y_f)p_f(a_f, b_f|x_f, y_f)[x_f\ y_f\ a_f\ b_f]_{X_f Y_f \bar{A}_f \bar{B}_f} \otimes \sigma_E^{a_f,b_f,x_f,y_f}, \tag{58}$$

where

$$\sigma_E^{a_f,b_f,x_f,y_f} = \frac{1}{p_f(a_f, b_f|x_f, y_f)} \mathrm{Tr}_{AB}[(\Lambda_{x_f}^{a_f} \otimes \Lambda_{y_f}^{b_f} \otimes I_E)\sigma_{ABE}], \tag{59}$$

and σ_{ABE} is an extension of σ_{AB}. Now, we know that

$$p_f(a_f, b_f|x_f, y_f) := \sum_{a,b,x,y} O^{(L)}(a_f, b_f|a, b, x, y, x_f, y_f)\, p_i(a, b|x, y) I^{(L)}(x, y|x_f, y_f), \tag{60}$$

and that the correlations $I^{(L)}(x, y|x_f y_f)$ and $O^{(L)}(a_f, b_f|a, b, x, y, x_f, y_f)$ are local correlations. Therefore, there exist separable states ρ_{XY} and $\rho_{A_F B_F}$, along with the POVMs which result in the correlations $I^{(L)}$ and $O^{(L)}$. That is,

$$I^{(L)}(x, y|x_f, y_f) = \mathrm{Tr}[(\Lambda_{x_f}^{x} \otimes \Lambda_{y_f}^{y})\rho_{XY}], \tag{61}$$

$$O^{(L)}(a_f, b_f|a, b, x, y, x_f, y_f) = \mathrm{Tr}[(\Lambda_{a,x_f,x}^{a_f} \otimes \Lambda_{b,b_f,y}^{b_f})\rho_{A_F B_F}] \tag{62}$$

Furthermore, we know that the correlation $p_i(a, b|x, y)$ is a quantum correlation. Therefore, it has an underlying state ρ_{AB} and POVMs characterized by $\{\Lambda_x^a\}_a$ and $\{\Lambda_y^b\}_b$. Then

$$p(a_f, b_f|x_f, y_f) = \sum_{a,b,x,y} \mathrm{Tr}[(\Lambda_{a,x_f,x}^{a_f} \otimes \Lambda_{b,b_f,y}^{b_f} \otimes \Lambda_x^a \otimes \Lambda_y^b \otimes \Lambda_{x_f}^x \otimes \Lambda_{y_f}^y)(\rho_{A_F B_F} \otimes \rho_{AB} \otimes \rho_{XY})]. \tag{63}$$

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 (56) is given by

$$\rho_{\bar{A}_f\bar{B}_f X_f Y_f E\Lambda_1\Lambda_2} = \sum_{x_f,y_f,a_f,b_f} p(x_f, y_f)p_f(a_f, b_f|x_f, y_f)[x_f, y_f, a_f, b_f]_{X_f Y_f A_f B_f} \rho_E^{a,b,x,y} \otimes [\lambda_1\lambda_2]_{\Lambda_1\Lambda_2}, \tag{64}$$

where

$$\rho_E^{a,b,x,y} = \frac{1}{p(a, b|x, y)} \mathrm{Tr}_{AB}[(\Lambda_x^a \otimes \Lambda_y^b \otimes I_E)\rho_{ABE}]. \tag{65}$$

Then it follows that

$$\begin{aligned}\rho_{\bar{A}_f\bar{B}_f X_f Y_f E\Lambda_1\Lambda_2} = \sum_{x_f,y_f,a_f,b_f} & p(x_f, y_f) \sum_{a,b,x,y} \sum_{\lambda_2} p_{\Lambda_2}(\lambda_2) O_A(a_f|a, x_f, x, \lambda_2) \\ & \times O_B(b_f|b, y, y_f, \lambda_2) p_i(a, b|x, y) \\ & \times \sum_{\lambda_1} p_{\Lambda_1}(\lambda_1) I_A(x|x_f, \lambda_1) I_B(y|y_f, \lambda_1)[x_f\ y_f\ a_f\ b_f]_{\bar{A}_f\bar{B}_f X_f Y_f} \otimes \rho_E^{a,b,x,y} \otimes [\lambda_1\lambda_2]_{\Lambda_1\Lambda_2}.\end{aligned} \tag{66}$$

This in turn is a marginal of the following state:

$$\begin{aligned}\rho_{\bar{A}_f\bar{B}_f X_f Y_f E\Lambda_1\Lambda_2 XY\bar{A}\bar{B}} = \sum_{x_f,y_f,a_f,b_f} & p(x_f, y_f) \sum_{a,b,x,y} \sum_{\lambda_2} p_{\Lambda_2}(\lambda_2) O_A(a_f|a, x_f, x, \lambda_2) O_B(b_f|b, y, y_f, \lambda_2) \\ & \times p_i(a, b|x, y) \sum_{\lambda_1} p_{\Lambda_1}(\lambda_1) I_A(x|x_f, \lambda_1) I_B(y|y_f, \lambda_1)[x_f\ y_f\ a_f\ b_f]_{X_f Y_f \bar{A}_f \bar{B}_f} \\ & \otimes \rho_E^{a,b,x,y} \otimes [\lambda_1\lambda_2]_{\Lambda_1\Lambda_2} \otimes [x\ y\ a\ b]_{XY\bar{A}\bar{B}}.\end{aligned} \tag{67}$$

Then, following arguments similar to that given in proposition 8, we obtain $N^Q(\bar{A}_f; \bar{B}_f)_{p_f} \leqslant N^Q(\bar{A}; \bar{B})_{p_i}$. ■

Proposition 10 (**Convexity of intrinsic non-locality**). *Let $p(a, b|x, y)$ and $q(a, b|x, y)$ be two correlations, and let $\lambda \in [0, 1]$. Let $t(a, b|x, y)$ be a mixture of the two correlations, defined as $t(a, b|x, y) = \lambda p(a, b|x, y) + (1 - \lambda)q(a, b|x, y)$. Then*

$$N(\bar{A}; \bar{B})_t \leqslant \lambda N(\bar{A}; \bar{B})_p + (1 - \lambda)N(\bar{A}; \bar{B})_q. \tag{68}$$

Proof. First, we embed the correlation $t(a, b|x, y)$ in the following classical–classical state $\tau_{\bar{A}\bar{B}XY}$:

$$\tau_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,\, y)t(a,\, b|x,\, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}}, \tag{69}$$

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 $\gamma_{\bar{A}\bar{B}XY}$:

$$\rho_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,\, y)p(a,\, b|x,\, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}}, \tag{70}$$

$$\gamma_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,\, y)q(a,\, b|x,\, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}}. \tag{71}$$

Next, consider an arbitrary no-signaling extension of $\tau_{\bar{A}\bar{B}XY}$:

$$\tau_{\bar{A}\bar{B}XYE} := \sum_{x,y,a,b} p(x,\, y)t(a,\, b|x,\, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}} \otimes \tau_E^{a,b,x,y}. \tag{72}$$

Similarly, consider an arbitrary no-signaling extension of $\rho_{\bar{A}\bar{B}XY}$ and $\gamma_{\bar{A}\bar{B}XY}$:

$$\rho_{\bar{A}\bar{B}XYE} = \sum_{x,y,a,b} p(x,\, y)p(a,\, b|x,\, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}} \otimes \rho_E^{a,b,x,y}, \tag{73}$$

$$\gamma_{\bar{A}\bar{B}XYE} = \sum_{x,y,a,b} p(x,\, y)q(a,\, b|x,\, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}} \otimes \gamma_E^{a,b,x,y}. \tag{74}$$

Now, consider the following particular no-signaling extension of $\tau_{\bar{A}\bar{B}XY}$:

$$\zeta_{\overline{A}\overline{B}XYEE'} := \sum_{x,y,a,b} p(x,\, y)[xyab]_{XY\overline{A}\overline{B}} \otimes (\lambda\, p(a,\, b|x,\, y)\rho_E^{a,b,x,y} \otimes [0]_{E'} + (1 - \lambda)q(a,\, b|x,\, y)\gamma_E^{a,b,x,y} \otimes [1]_{E'}). \tag{75}$$

Then,

$$\inf_{\text{ext.in (72)}} I(\bar{A};\, \bar{B}|XYE)_\tau \leqslant I(\bar{A};\, \bar{B}|XYEE')_\zeta \tag{76}$$

$$= \lambda I(\bar{A};\, \bar{B}|XYE)_\rho + (1 - \lambda)I(\bar{A};\, \bar{B}|XYE)_\gamma. \tag{77}$$

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 (73) and (74), we conclude that

$$\inf_{\text{ext.in (72)}} I(\bar{A};\, \bar{B}|XYE)_\zeta \leqslant \lambda \inf_{\text{ext.in (73)}} I(\bar{A};\, \bar{B}|XYE)_\rho + (1 - \lambda) \inf_{\text{ext.in (74)}} I(\bar{A};\, \bar{B}|XYE)_\gamma. \tag{78}$$

Taking the supremum over all measurement choices, we find that

$$\sup_{p(x,y)} \inf_{\text{ext.in (72)}} I(\bar{A};\, \bar{B}|XYE)_\zeta \leqslant \lambda \sup_{p(x,y)} \inf_{\text{ext.in (73)}} I(\bar{A};\, \bar{B}|XYE)_\rho + (1 - \lambda)\sup_{p(x,y)} \inf_{\text{ext.in (74)}} I(\bar{A};\, \bar{B}|XYE)_\gamma. \tag{79}$$

This completes the proof. ■

Proposition 11 (**Convexity of quantum intrinsic non-locality**). *Let $p(a,\, b|x,\, y)$ and $q(a,\, b|x,\, y)$ be correlations in* **Q**, *and let $\lambda \in [0,\, 1]$. Let $t(a,\, b|x,\, y)$ be a mixture of the correlations defined as $t(a,\, b|x,\, y) = \lambda p(a,\, b|x,\, y) + (1 - \lambda)q(a,\, b|x,\, y)$. Then*

$$N^{\mathrm{Q}}(\bar{A};\, \bar{B})_t \leqslant \lambda N^{\mathrm{Q}}(\bar{A};\, \bar{B})_p + (1 - \lambda)N^{\mathrm{Q}}(\bar{A};\, \bar{B})_q. \tag{80}$$

Proof. Since **Q** is a convex set [Pit86], we know that $t(a,\, b|x,\, y) \in \mathbf{Q}$. First, we embed the correlation $t(a,\, b|x,\, y)$ in the following quantum state $\tau_{\bar{A}\bar{B}XY}$:

$$\tau_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,\, y)t(a,\, b|x,\, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}}, \tag{81}$$

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 $\gamma_{\bar{A}\bar{B}XY}$:

$$\rho_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,\, y)p(a,\, b|x,\, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}}, \tag{82}$$

$$\gamma_{\bar{A}\bar{B}XY} := \sum_{x,y,a,b} p(x,\, y)q(a,\, b|x,\, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}}. \tag{83}$$

Next, consider an arbitrary quantum extension of $\tau_{\bar{A}\bar{B}XY}$:

$$\tau_{\bar{A}\bar{B}XYE} := \sum_{x,y,a,b} p(x, y)t(a, b|x, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}} \otimes \tau_E^{a,b,x,y}. \tag{84}$$

Similarly, consider an arbitrary quantum extension of $\rho_{\bar{A}\bar{B}XY}$ and $\gamma_{\bar{A}\bar{B}XY}$:

$$\rho_{\bar{A}\bar{B}XYE} = \sum_{x,y,a,b} p(x, y)p(a, b|x, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}} \otimes \rho_E^{a,b,x,y}, \tag{85}$$

$$\gamma_{\bar{A}\bar{B}XYE} = \sum_{x,y,a,b} p(x, y)q(a, b|x, y)[x\ y\ a\ b]_{XY\bar{A}\bar{B}} \otimes \gamma_E^{a,b,x,y}. \tag{86}$$

Let ρ_{AB} be a quantum state that, along with the POVMs characterized by Λ_x^a and Λ_y^b, yield the correlation $p(a, b|x, y)$. 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 correlation $q(a, b|x, y)$. Let γ_{ABE} be an extension of γ_{AB}. Then, a particular quantum state that realizes the correlation $t(a, b|x, y)$ is the following:

$$\tau_{ABA'B'} = \lambda\rho_{AB} \otimes |00\rangle\langle 00|_{A'B'} + (1 - \lambda)\gamma_{AB} \otimes |11\rangle\langle 11|_{A'B'}, \tag{87}$$

$$t(a, b|x, y) = \mathrm{Tr}\left[(\Lambda_x^a \otimes \Lambda_y^b \otimes (|00\rangle\langle 00|_{A'B'}) + M_x^a \otimes M_y^b \otimes (|11\rangle\langle 11|_{A'B'}))(\tau_{ABA'B'})\right], \tag{88}$$

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 $\tau_{ABA'B'}$:

$$\tau_{ABA'B'EE'} = \lambda\rho_{ABE} \otimes |000\rangle\langle 000|_{A'B'E'} + (1 - \lambda)\gamma_{ABE} \otimes |111\rangle\langle 111|_{A'B'E'}. \tag{89}$$

Furthermore, consider the following particular quantum extension of $\tau_{\bar{A}\bar{B}XY}$:

$$\begin{aligned}\zeta_{\bar{A}\bar{B}XYEE'} := & \sum_{x,y,a,b} p(x, y)[xyab]_{XY\bar{A}\bar{B}} \\ & \otimes(\lambda\, p(a, b|x, y)\rho_E^{a,b,x,y} \otimes [0]_{E'} + (1 - \lambda)q(a, b|x, y)\gamma_E^{a,b,x,y} \otimes [1]_{E'}).\end{aligned} \tag{90}$$

Then following similar arguments given in the proof of proposition 10, we obtain

$$N^Q(\bar{A}; \bar{B})_t \leqslant \lambda N^Q(\bar{A}; \bar{B})_p + (1 - \lambda)N^Q(\bar{A}; \bar{B})_q, \tag{91}$$

concluding the proof. ■

Proposition 12 (**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 correlation for which the following no-signaling constraints hold:*

$$\begin{aligned}&\textstyle\sum_{a_1} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) \\ &\quad = \textstyle\sum_{a_1} p(a_1, a_2, b_1, b_2|x_1', x_2, y_1, y_2) \quad \forall x_1', x_1, x_2, y_1, y_2 \in [s],\ a_2, b_1, b_2 \in [r], \\ &\textstyle\sum_{a_2} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) \\ &\quad = \textstyle\sum_{a_2} p(a_1, a_2, b_1, b_2|x_1, x_2', y_1, y_2) \quad \forall x_2', x_2, x_1, y_1, y_2 \in [s],\ a_1, b_1, b_2 \in [r], \\ &\textstyle\sum_{b_1} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) \\ &\quad = \textstyle\sum_{b_1} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1', y_2) \quad \forall y_1', y_1, x_1, x_2, y_2 \in [s],\ a_1, a_2, b_2 \in [r], \\ &\textstyle\sum_{b_2} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) \\ &\quad = \textstyle\sum_{b_2} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2') \quad \forall y_2', y_2, x_2, y_1, x_1 \in [s],\ a_1, a_2, b_1 \in [r].\end{aligned}$$

Let $t(a_1, b_1|x_1, y_1)$ and $r(a_2, b_2|x_2, y_2)$ be correlations corresponding to the marginals of the probability distribution $p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2)$. Then the intrinsic non-locality is super-additive, in the sense that

$$N(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2)_p \geqslant N(\bar{A}_1; \bar{B}_1)_t + N(\bar{A}_2; \bar{B}_2)_r. \tag{92}$$

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:

$$N(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2)_p = N(\bar{A}_1; \bar{B}_1)_t + N(\bar{A}_2; \bar{B}_2)_r. \tag{93}$$

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:

$$\begin{aligned}\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2E} = & \textstyle\sum_{x_1,x_2,y_1,y_2,a_1,a_2,b_1,b_2} p(x_1, y_1, x_2, y_2)p(a_1, b_1, a_2, b_2|x_1, x_2, y_1, y_2) \\ & [a_1\ b_1\ x_1\ y_1\ a_2\ b_2\ x_2\ y_2]_{\bar{A}_1\bar{B}_1X_1Y_1\bar{A}_2\bar{B}_2X_2Y_2} \otimes \rho_E^{a_1,b_1,x_1,y_1,a_2,b_2,x_2,y_2},\end{aligned} \tag{94}$$

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

$$I(\bar{A}_1\bar{A}_2;\ \bar{B}_1\bar{B}_2|X_1X_2Y_1Y_2E)_\rho = I(\bar{A}_1\bar{A}_2;\ \bar{B}_1|X_1Y_1X_2Y_2E) + I(\bar{A}_1\bar{A}_2;\ \bar{B}_2|EX_1Y_1X_2Y_2\bar{B}_1) \tag{95}$$

$$= I(\bar{A}_1;\ \bar{B}_1|X_1Y_1X_2Y_2E)_\rho + I(\bar{A}_2;\ \bar{B}_1|EX_1Y_1X_2Y_2\bar{A}_1)_\rho + I(\bar{A}_1;\ \bar{B}_2|X_1Y_1X_2Y_2E\bar{B}_1) + I(\bar{A}_2;\ \bar{B}_2|X_1Y_1X_2Y_2E\bar{A}_1\bar{B}_1) \tag{96}$$

$$\geqslant I(\bar{A}_1;\ \bar{B}_1|X_1Y_1X_2Y_2E)_\rho + I(\bar{A}_2;\ \bar{B}_2|X_1Y_1X_2Y_2E\bar{A}_1\bar{B}_1)_\rho. \tag{97}$$

From the no-signaling constraints in the statement of the proposition and (94), we obtain

$$\rho_{\bar{A}_1\bar{B}_1X_1X_2Y_1Y_2E} = \sum_{a_1,b_1,x_1,x_2,y_1,y_2} p(x_1,\ x_2,\ y_1,\ y_2)[a_1\ b_1\ x_1\ y_1\ x_2\ y_2]_{\bar{A}_1\bar{B}_2X_1Y_1X_2Y_2} \otimes p(a_1,\ b_1|x_1,\ y_1)\rho_E^{x_1,y_1,a_1,b_1}. \tag{98}$$

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:

$$\tau_{\bar{A}_1\bar{B}_1X_1Y_1E} := \sum_{x_1,y_1} p(x_1,\ y_1) \otimes \sum_{a_1,b_1} [x_1\ y_1\ a_1\ b_1]_{X_1Y_1\bar{A}_1\bar{B}_1} \otimes t(a_1,\ b_1|x_1,\ y_1)\tau_E^{a_1,b_1,x_1,y_1}, \tag{99}$$

$$\gamma_{\bar{A}_2\bar{B}_2X_2Y_2E} := \sum_{x_2,y_2} p(x_2,\ y_2) \otimes \sum_{a_2,b_2} [x_2\ y_2\ a_2\ b_2]_{X_2Y_2\bar{A}_2\bar{B}_2} \otimes r(a_2,\ b_2|x_2,\ y_2)\gamma_E^{a_2,b_2,x_2,y_2}. \tag{100}$$

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:

$$I(\bar{A}_1\bar{A}_2;\ \bar{B}_1\bar{B}_2|X_1X_2Y_1Y_2E)_\rho \geqslant I(\bar{A}_1;\ \bar{B}_1|X_1Y_1X_2Y_2E)_\rho + I(\bar{A}_2;\ \bar{B}_2|X_1Y_1X_2Y_2E\bar{A}_1\bar{B}_1)_\rho \tag{101}$$

$$\geqslant \inf_{\text{ext.in (99)}} I(\bar{A}_1;\ \bar{B}_1|X_1Y_1E)_\tau + \inf_{\text{ext.in (100)}} I(\bar{A}_2;\ \bar{B}_2|X_2Y_2E\bar{A}_1\bar{B}_1)_\gamma. \tag{102}$$

Since (102) holds for an arbitrary no-signaling extension of ρ, we obtain

$$\inf_{\text{ext.in (94)}} I(\bar{A}_1\bar{A}_2;\ \bar{B}_1\bar{B}_2|X_1X_2Y_1Y_2E)_\rho \geqslant \inf_{\text{ext.in (99)}} I(\bar{A}_1;\ \bar{B}_1|X_1Y_1E)_\tau + \inf_{\text{ext.in (100)}} I(\bar{A}_2;\ \bar{B}_2|X_2Y_2E\bar{A}_1\bar{B}_1)_\gamma \tag{103}$$

Since the above equation holds for arbitrary probability distributions, we can take a supremum over all probability distributions to obtain

$$\sup_{p(x_1,y_1)p(x_2,y_2)} \inf_{\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2E}} I(\bar{A}_1\bar{A}_2;\ \bar{B}_1\bar{B}_2|X_1X_2Y_1Y_2E)_\rho \geqslant \sup_{p(x_1,y_1)} \inf_{\tau_{\bar{A}_1\bar{B}_1X_1Y_1E}} I(\bar{A}_1;\ \bar{B}_1|X_1Y_1E)_\tau + \sup_{p(x_2,y_2)} \inf_{\gamma_{\bar{A}_2\bar{B}_2X_2Y_2E}} I(\bar{A}_2;\ \bar{B}_2|X_2Y_2E)_\gamma. \tag{104}$$

Since we have considered a supremum over product probability distributions for the measurement choices on the LHS, we can relax this to consider a supremum over all probability distributions $p(x_1,\ y_1,\ x_2,\ y_2)$ of the measurement choices. This concludes the proof of (92).

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:

$$N(\bar{A}_1\bar{A}_2;\ \bar{B}_1\bar{B}_2)_p \leqslant N(\bar{A}_1;\ \bar{B}_1)_t + N(\bar{A}_2;\ \bar{B}_2)_r. \tag{105}$$

Consider the following states

$$\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2} = \sum_{a_1,b_1,x_1,y_1,a_2,b_2,x_2,y_2} p(x_1,\ x_2,\ y_1,\ y_2)t(a_1,\ b_1|x_1,\ y_1)r(a_2,\ b_2|x_2,\ y_2) [a_1\ b_1\ a_2\ b_2\ x_1\ x_2\ y_1\ y_2]_{\bar{A}_1\bar{B}_1\bar{A}_2\bar{B}_2X_1Y_1X_2Y_2}. \tag{106}$$

Consider an arbitrary extension of the state $\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2}$

$$\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2E} := \sum_{a_1,b_1,x_1,y_1,a_2,b_2,x_2,y_2} p(x_1,\ x_2,\ y_1,\ y_2)t(a_1,\ b_1|x_1,\ y_1)r(a_2,\ b_2|x_2,\ y_2) [a_1\ b_1\ x_1\ y_1\ a_2\ b_2\ x_2\ y_2] \otimes \rho_E^{a_1,b_1,x_1,y_1,a_2,b_2,x_2,y_2}. \tag{107}$$

Now, consider a particular extension of the state $\rho_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2}$:

$$\zeta_{\bar{A}_1\bar{A}_2\bar{B}_1\bar{B}_2X_1X_2Y_1Y_2E_1E_2} := \sum_{a_1,b_1,x_1,y_1,a_2,b_2,x_2,y_2} p(x_1, x_2, y_1, y_2)\, t(a_1, b_1|x_1, y_1)\, r(a_2, b_2|x_2, y_2) \\ [a_1b_1a_2b_2x_1x_2y_1y_2]_{\bar{A}_1\bar{B}_1X_1Y_1\bar{A}_2\bar{B}_2X_2Y_2} \otimes \rho_{E_1}^{a_1,b_1,x_1,y_1} \otimes \rho_{E_2}^{a_2,b_2,x_2,y_2}. \tag{108}$$

Then, we have the following set of inequalities:

$$\inf_{\text{ext.in (107)}} I(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2|X_1Y_1X_2Y_2E)_\rho \leqslant I(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2|X_1Y_1X_2Y_2E_1E_2)_\zeta \tag{109}$$

$$= I(\bar{A}_1; \bar{B}_1|X_1Y_1X_2Y_2E_1E_2)_\zeta + I(\bar{A}_2; \bar{B}_1|E_1E_2X_1Y_1X_2Y_2\bar{A}_1)_\zeta \\ + I(\bar{A}_1; \bar{B}_2|X_1Y_1X_2Y_2E_1E_2\bar{B}_1)_\zeta + I(\bar{A}_2; \bar{B}_2|X_1Y_1X_2Y_2E_1E_2\bar{A}_1\bar{B}_1)_\zeta \tag{110}$$

$$= I(\bar{A}_1; \bar{B}_1|X_1Y_1X_2Y_2E_1E_2)_\zeta + I(\bar{A}_2; \bar{B}_2|X_1Y_1X_2Y_2E_1E_2\bar{A}_1\bar{B}_1)_\zeta. \tag{111}$$

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:

$$I(\bar{A}_2; \bar{B}_1|E_1E_2X_1Y_1X_2Y_2\bar{A}_1)_\zeta = H(\bar{A}_2|E_1E_2X_1Y_1X_2Y_2\bar{A}_1)_\zeta - H(\bar{A}_2|E_1E_2X_1Y_1X_2Y_2\bar{A}_1\bar{B}_1)_\zeta \tag{112}$$

$$= \sum_{x_1x_2y_1y_2} p(x_1, x_2, y_1, y_2)[H(\bar{A}_2|\bar{A}_1E_1E_2)_{\zeta^{x_1x_2y_1y_2}} - H(\bar{A}_2|\bar{A}_1E_1E_2\bar{B}_1)_{\zeta^{x_1x_2y_1y_2}}], \tag{113}$$

where

$$\zeta^{x_1,x_2,y_1,y_2}_{\bar{A}_1\bar{A}_2E_1E_2} = \sum_{a_1} t(a_1|x_1)[a_1]_{\bar{A}_1} \otimes \rho_{E_1}^{a_1,x_1} \otimes \sum_{a_2} r(a_2|x_2)[a_2]_{\bar{A}_2} \otimes \rho_{E_2}^{a_2,x_2}, \tag{114}$$

$$\zeta^{x_1,x_2,y_1,y_2}_{\bar{A}_1\bar{A}_2\bar{B}_1E_1E_2} = \sum_{a_1,b_1} t(a_1, b_1|x_1, y_1)[a_1b_1]_{\bar{A}_1\bar{B}_1} \otimes \rho_{E_1}^{a_1,x_1,b_1,y_1} \otimes \sum_{a_2} r(a_2|x_2)[a_2]_{\bar{A}_2} \otimes \rho_{E_2}^{a_2,x_2}. \tag{115}$$

Then, from (114) and (115), it follows that

$$H(\bar{A}_2|\bar{A}_1E_1E_2)_{\zeta^{x1x2y2y2}} = H(\bar{A}_2|E_2)_{\zeta^{x1x2y2y2}}, \tag{116}$$

$$H(\bar{A}_2|\bar{A}_1E_1E_2\bar{B}_1)_{\zeta^{x1x2y2y2}} = H(\bar{A}_2|E_2)_{\zeta^{x1x2y2y2}}. \tag{117}$$

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 (108), and from the no-signaling constraints, it follows that

$$\inf_{\text{ext.in (107)}} I(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2|X_1Y_1X_2Y_2E)_\rho \leqslant I(\bar{A}_1; \bar{B}_1|X_1Y_1E_1)_\zeta + I(\bar{A}_2; \bar{B}_2|X_2Y_2E_2)_\zeta. \tag{118}$$

Since the above statement holds for an arbitrary no-signaling extension of the form in (107), it follows that

$$\inf_{\text{ext.in (107)}} I(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2|X_1Y_1X_2Y_2E)_\rho \\ \leqslant \inf_{\text{ext.in (108)}} I(\bar{A}_1; \bar{B}_1|X_1Y_1E_1)_\zeta + \inf_{\text{ext.in (108)}} I(\bar{A}_2; \bar{B}_2|X_2Y_2E_2)_\zeta. \tag{119}$$

Since the above inequality holds for an arbitrary probability distribution $p(x_1, x_2, y_1, y_2)$, we find that

$$\sup_{p(x_1,x_2,y_1,y_2)} \inf_{\text{ext.in (107)}} I(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2|X_1Y_1X_2Y_2E)_\rho \\ \leqslant \sup_{p(x_1,y_1)} \inf_{\text{ext.in (108)}} I(\bar{A}_1; \bar{B}_1|X_1Y_1E_1)_\zeta + \sup_{p(x_2,y_2)} \inf_{\text{ext.in (108)}} I(\bar{A}_2; \bar{B}_2|X_2Y_2E_2)_\zeta. \tag{120}$$

This concludes the proof. ■

Proposition 13 (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 correlation that arises from a four-party state $\rho_{A_1A_2B_1B_2}$, and POVMs characterized by $\Lambda_{x_1}^{a_1}$, $\Lambda_{x_2}^{a_2}$, $\Lambda_{y_1}^{b_1}$ and $\Lambda_{y_2}^{b_2}$. Then the following no-signaling constraints hold:*

$$\sum_{a_1} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) = \sum_{a_1} p(a_1, a_2, b_1, b_2|x_1', x_2, y_1, y_2) \quad \forall\, x_1', x_1, x_2, y_1, y_2 \in [s],\ a_2, b_1, b_2 \in [r]$$

$$\sum_{a_2} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) = \sum_{a_2} p(a_1, a_2, b_1, b_2|x_1, x_2', y_1, y_2) \quad \forall\, x_2', x_2, x_1, y_1, y_2 \in [s],\ a_1, b_1, b_2 \in [r]$$

$$\sum_{b_1} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) = \sum_{b_1} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1', y_2) \quad \forall\, y_1', y_1, x_1, x_2, y_2 \in [s],\ a_1, a_2, b_2 \in [r]$$

$$\sum_{b_2} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2) = \sum_{b_2} p(a_1, a_2, b_1, b_2|x_1, x_2, y_1, y_2') \quad \forall\, y_2', y_2, x_2, y_1, x_1 \in [s],\ a_1, a_2, b_1 \in [r].$$

Let $t(a_1, b_1|x_1, y_1)$ and $r(a_2, b_2|x_2, y_2)$ be quantum correlations 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

$$N^Q(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2)_p \geqslant N^Q(\bar{A}_1; \bar{B}_1)_t + N^Q(\bar{A}_2; \bar{B}_2)_r. \tag{121}$$

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:*

$$N^{Q}(\bar{A}_1\bar{A}_2; \bar{B}_1\bar{B}_2)_p = N^{Q}(\bar{A}_1; \bar{B}_1)_t + N^{Q}(\bar{A}_2; \bar{B}_2)_r. \tag{122}$$

Proof. The proof follows by using similar techniques as proposition 12, and by taking appropriate quantum extensions. ■

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 $\{\Lambda_a^x\}$[5]. 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 correlations that can arise from this assemblage.

Proposition 14. *Let* $p(a, b|x, y)$ *be a quantum correlation that is obtained by performing a POVM* $\{\Lambda_y^b\}_b$ *on the assemblage* $\{p_{\bar{A}|X}(a|x)\rho_B^{a,x}\}_{a,x}$. *Then the quantum intrinsic non-locality of the correlation* p *does not exceed the intrinsic steerability of the assemblage* $\hat{\rho}$. *That is,*

$$N^{Q}(\bar{A}; \bar{B})_p \leqslant S(\bar{A}; B)_{\hat{\rho}}, \tag{123}$$

where we recall that $\hat{\rho}$ *is a shorthand to denote the assemblage.*

Proof. Let $p(a, b|x, y)$ be a quantum correlation that arises from the assemblage $p_{\bar{A}|X}(a|x)\rho_B^{a,x}$. That is,

$$p(a, b|x, y) = \mathrm{Tr}[\Lambda_y^b(p_{\bar{A}|X}(a|x)\rho_B^{a,x})]. \tag{124}$$

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

$$p(a, b|x, y)\rho_E^{a,x,b,y} = \mathrm{Tr}_B[\Lambda_y^b(p_{\bar{A}|X}(a|x)\rho_{BE}^{a,x})]. \tag{125}$$

From [SBC+15], 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 correlation embedded in a classical–classical state $\rho_{\bar{A}\bar{B}XY}$ with the following particular no-signaling extension:

$$\rho_{\bar{A}\bar{B}XYE} \coloneqq \sum_{a,b,x,y} p(x, y)p(a, b|x, y)[a\ b\ x\ y]_{\bar{A}\bar{B}XY} \otimes \rho_E^{a,b,x,y}, \tag{126}$$

and an arbitrary quantum extension:

$$\sigma_{\bar{A}\bar{B}XYE} \coloneqq \sum_{a,b,x,y} p(x, y)p(a, b|x, y)[a\ b\ x\ y]_{\bar{A}\bar{B}XY} \otimes \sigma_E^{a,b,x,y}. \tag{127}$$

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

$$\rho_{\bar{A}BXE} = \sum_{a,x} p(x)p_{\bar{A}|X}(a|x)[a\ x]_{\bar{A}X} \otimes \rho_{BE}^{a,x}. \tag{128}$$

Let

$$\rho_{\bar{A}BXYE} = \sum_{a,x} p(x, y)p_{\bar{A}|X}(a|x)[a\ x]_{\bar{A}X} \otimes \rho_{BE}^{a,x}. \tag{129}$$

Then,

$$I(\bar{A}; B|XE)_\rho = I(\bar{A}; BY|XE)_\rho. \tag{130}$$

This follows from chain rule of conditional mutual information and inspection of (129). Observe that Bob can perform a local operation and tranform the state $\rho_{\bar{A}BXYE}$ to $\rho_{\bar{A}\bar{B}XYE}$. Then, from the data-processing inequality, we find that

$$I(\bar{A}; B|XE)_\rho \geqslant I(\bar{A}; \bar{B}Y|XE)_\rho \tag{131}$$

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 correlation $p(a, b|x, y)$ derived from the assemblage $p_{\bar{A}|X}(a|x)\rho^{a,x}$, such that (131) is true. Therefore, we obtain the following:

$$\inf_{\text{ext in (128)}} I(\bar{A}; B|XE)_\rho \geqslant \inf_{\text{ext.in (126)}} I(\bar{A}; \bar{B}Y|XE)_\rho \tag{132}$$

[5] From [SBC+15], it can be seen that given a bipartite assemblage, we can always find an underlying quantum state and measurements.

$$\geqslant \inf_{\text{ext in (127)}} I(\bar{A};\, \bar{B}Y|XE)_{\sigma}. \tag{133}$$

This in turn implies that

$$S(\bar{A};\, B)_{\hat{\rho}} \geqslant N^{Q}(\bar{A};\, \bar{B})_{p}, \tag{134}$$

concluding the proof. ■

3.6. Intrinsic non-locality of a PR box

In this section, we calculate the intrinsic non-locality of a PR box.

Proposition 15. *The intrinsic non-locality of a PR box is equal to 1,* i.e., $N(\overline{A};\, \overline{B})_{p} = 1$, *where p is the correlation defined in (14).*

Proof. Consider the state

$$\rho_{\bar{A}\bar{B}XY} \coloneqq \sum_{a,b,x,y} p(x,\, y)p(a,\, b|x,\, y)[a\ b\ x\ y]_{\bar{A}\bar{B}XY}, \tag{135}$$

where $p(x,\, y)$ is an arbitrary probability distribution. Consider a no-signaling extension of the state

$$\rho_{\bar{A}\bar{B}XYE} \coloneqq \sum_{a,b,x,y} p(x,\, y)p(a,\, b|x,\, y)[a\ b\ x\ y]_{\bar{A}\bar{B}XY} \otimes \rho_{E}^{a,b,x,y}. \tag{136}$$

The no-signaling constraints are

$$\sum_{a,b,y} p(a,\, b|x,\, y)[b\ x\ y]_{\bar{B}XY} \otimes \rho_{E}^{x,y,a,b} = \sum_{a,b,y} p(a,\, b|x',\, y)[b\ x'\ y]_{\bar{B}XY} \otimes \rho_{E}^{x',y,a,b}, \tag{137}$$

$$\sum_{b,a,x} p(a,\, b|x,\, y)[a\ x\ y]_{\bar{A}XY} \otimes \rho_{E}^{x,y,a,b} = \sum_{b,a,x} p(a,\, b|x,\, y')[a\ x\ y']_{\bar{A}XY} \otimes \rho_{E}^{x,y',a,b}. \tag{138}$$

From (14), and the no-signaling constraint in (137), we arrive at the following constraints on the possible states of Eve's system:

$$\begin{bmatrix} \rho_{E}^{0000} & 0 & 0 & 0 \\ 0 & \rho_{E}^{0011} & 0 & 0 \\ 0 & 0 & \rho_{E}^{0100} & 0 \\ 0 & 0 & 0 & \rho_{E}^{0111} \end{bmatrix} = \begin{bmatrix} \rho_{E}^{1000} & 0 & 0 & 0 \\ 0 & \rho_{E}^{1011} & 0 & 0 \\ 0 & 0 & \rho_{E}^{1110} & 0 \\ 0 & 0 & 0 & \rho_{E}^{1101} \end{bmatrix}. \tag{139}$$

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 (139) can also be written as

$$\begin{aligned} &(1)\ \rho_{E}^{0000} = \rho_{E}^{1000}, \quad &&(2)\ \rho_{E}^{0011} = \rho_{E}^{1011}, \\ &(3)\ \rho_{E}^{0100} = \rho_{E}^{1110}, \quad &&(4)\ \rho_{E}^{0111} = \rho_{E}^{1101}. \end{aligned} \tag{140}$$

Similarly, from (14), and the no-signaling constraint in (138), we arrive at the following constraints on the possible states of Eve's system:

$$\begin{bmatrix} \rho_{E}^{0000} & 0 & 0 & 0 \\ 0 & \rho_{E}^{0011} & 0 & 0 \\ 0 & 0 & \rho_{E}^{1000} & 0 \\ 0 & 0 & 0 & \rho_{E}^{1011} \end{bmatrix} = \begin{bmatrix} \rho_{E}^{0100} & 0 & 0 & 0 \\ 0 & \rho_{E}^{0111} & 0 & 0 \\ 0 & 0 & \rho_{E}^{1101} & 0 \\ 0 & 0 & 0 & \rho_{E}^{1110} \end{bmatrix}. \tag{141}$$

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 (141) can also be written as

$$\begin{aligned} &(5)\ \rho_{E}^{0000} = \rho_{E}^{0100}, \quad &&(6)\ \rho_{E}^{0011} = \rho_{E}^{0111}, \\ &(7)\ \rho_{E}^{1000} = \rho_{E}^{1101}, \quad &&(8)\ \rho_{E}^{0111} = \rho_{E}^{1101}. \end{aligned} \tag{142}$$

By following $1 \to 7 \to 4 \to 6 \to 2 \to 8 \to 3 \to 5 \to 1$ 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,

$$I(\bar{A};\, \bar{B}|XYE)_{\rho} = I(\bar{A};\, \bar{B}|XY)_{\rho} = \sum_{x,y} p(x,\, y)I(\bar{A};\, \bar{B})_{\rho^{x,y}} \tag{143}$$

$$= \sum_{x,y} p(x,y)(H(\bar{A})_{\rho^{x,y}} - H(\bar{A}|\bar{B})_{\rho^{x,y}}) \tag{144}$$

$$= 1. \tag{145}$$

It is easy to check that given realizations of X, Y, the entropies $H(\bar{A}|\bar{B})_{\rho^{x,y}} = 0$ and $H(\bar{A})_{\rho^{x,y}} = 1$. ■

4. Faithfulness of restricted intrinsic steerability

In this section, we solve an open question from [KWW17], regarding the faithfulness of restricted intrinsic steerability.

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

$$\sup_{p_X(x)} \|\rho_{\bar{A}XB} - \sigma_{\bar{A}XB}\|_1 \leqslant |\mathcal{X}|\left(\varepsilon^{1/4} + \frac{\varepsilon^{1/16}|\mathcal{X}|^{1/2}}{1-\varepsilon^{1/16}|\mathcal{X}|^{1/2}} + 4|\mathcal{X}|e^{-\frac{\varepsilon^{-1/4}}{3}}\right). \tag{146}$$

Proof. The forward direction ('if') was established in [KWW17], Proposition 12. 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 [LW18].

Invoking theorem 5.1 of [FR15], we know that there exists a recovery channel $\mathcal{R}_{XE\to\bar{A}XE}$ such that

$$\|\rho_{\bar{A}XBE} - \mathcal{R}_{XE\to\bar{A}XE}(\rho_{BE}\otimes\rho_X)\|_1 \leqslant \sqrt{I(\bar{A};B|EX)_\rho \ln 2} =: t, \tag{147}$$

$$\|\rho_{\bar{A}XBE} - \mathcal{R}_{X_2E\to\bar{A}_2X_2E}\circ \mathrm{Tr}_{\bar{A}_1X_1}(\rho_{\bar{A}_1X_1BE}\otimes\rho_{X_2})\|_1 \leqslant t, \tag{148}$$

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

$$\|\mathcal{R}_{X_3E\to\bar{A}_3X_3E}\circ \mathrm{Tr}_{X_2\bar{A}_2}(\rho_{\bar{A}_2X_2BE}\otimes\rho_{X_3}) \quad - \bigcirc_{i=2}^3 \mathcal{R}_{X_iE\to\bar{A}_iX_iE}\circ \mathrm{Tr}_{A_{i-1}X_{i-1}}(\rho_{\bar{A}_1X_1BE}\otimes\rho_{X_2}\otimes\rho_{X_3})\|_1 \leqslant t. \tag{149}$$

which follows from the monotonicity of trace distance with respect to $\mathcal{R}_{X_3E\to\bar{A}_3X_3E}\circ \mathrm{Tr}_{X_2\bar{A}_2}$. Then, combining the above equation with (147) via the triangle inequality, we obtain

$$\|\rho_{\bar{A}XBE} - \bigcirc_{i=2}^3 \mathcal{R}_{X_iE\to\bar{A}_iX_iE}\circ \mathrm{Tr}_{A_{i-1}X_{i-1}}(\rho_{\bar{A}_1X_1BE}\otimes\rho_{X_2}\otimes\rho_{X_3})\|_1 \leqslant 2t. \tag{150}$$

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

$$\|\rho_{\bar{A}XB} - \mathrm{Tr}_E\{\bigcirc_{i=2}^j \mathcal{R}_{X_iE\to\bar{A}_iX_iE}\circ \mathrm{Tr}_{A_{i-1}X_{i-1}}(\rho_{\bar{A}_1X_1BE}\otimes\rho_X^{\otimes j})\}\|_1 \leqslant nt. \tag{151}$$

The recovery channel $\mathcal{R}_{X_iE\to\bar{A}_iX_iE}$ can be taken as [Wil15]

$$\mathcal{R}_{XE\to\bar{A}XE}(\cdot) = \rho_{\bar{A}XE}^{\frac{1}{2}+i\omega}\rho_{X\bar{E}}^{-\frac{1}{2}-i\omega}(\cdot)\rho_{X\bar{E}}^{-\frac{1}{2}+i\omega}\rho_{\bar{A}XE}^{\frac{1}{2}-i\omega}, \tag{152}$$

$$= \sum_x |x\rangle\langle x|_X \otimes (\rho_{\bar{A}E}^x)^{\frac{1}{2}+i\omega}\rho_E^{-\frac{1}{2}+i\omega}(\cdot)\rho_E^{-\frac{1}{2}+i\omega}(\rho_{\bar{A}E}^x)^{\frac{1}{2}-i\omega}, \tag{153}$$

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

$$\sigma_{\bar{A}^nX^nBE} = (\mathcal{R}_{X_nE\to\bar{A}_nX_nE}\circ\cdots\circ\mathcal{R}_{X_1E\to\bar{A}_1X_1E})(\sigma_{BE}\otimes\sigma_X^{\otimes n}) \tag{154}$$

$$= \sum_{a^n,x^n} p_{X^n}(x^n)q_{\bar{A}^n|X^n}(a^n|x^n)|x^n\rangle\langle x^n|_{X^n}\otimes|a^n\rangle\langle a^n|_{A^n}\otimes\sigma_{BE}^{a^n,x^n}. \tag{155}$$

$$\sigma_{\bar{A}^nX^nB} = \mathrm{Tr}_E(\sigma_{\bar{A}^nX^nBE}) \tag{156}$$

$$= \sum_{a^n,x^n} p_{X^n}(x^n)q_{\bar{A}^n|X^n}(a^n|x^n)|x^n\rangle\langle x^n|_{X^n}\otimes|a^n\rangle\langle a^n|_{A^n}\otimes\sigma_B^{a^n,x^n}. \tag{157}$$

$$\sigma_{\bar{A}_iX_iB} = \mathrm{Tr}_{A^{[n]\setminus\{i\}}X^{[n]\setminus\{i\}}}(\sigma_{\bar{A}^nX^nB}) \tag{158}$$

$$= \sum_{a^n,x^n} p_{X^n}(x^n)q_{\bar{A}^n|X^n}(a^n|x^n)|x^n\rangle\langle x^n|_{X^n}\otimes|a^n\rangle\langle a^n|_{A^n}\otimes\sigma_B^{a^n,x^n}, \tag{159}$$

where $A^{[n]\setminus\{i\}} = A_1A_2\cdots A_{i-1}A_{i+1}\cdots A_n$ and similarly $X^{[n]\setminus\{i\}} = X_1X_2\cdots X_{i-1}X_{i+1}\cdots X_n$. 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 (151), we obtain for all $i \in \{1, 2, \ldots, n\}$ that

$$\|\rho_{\bar{A}XB} - \sigma_{\bar{A}_iX_iB}\|_1 \leqslant nt. \tag{160}$$

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 LHV model in [TDS03].

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

$$p_{I|\tilde{X}X^n}(i|\tilde{x}x^n) = \frac{1}{N(\tilde{x}|x^n)}\delta_{x_i,\tilde{x}}, \tag{161}$$

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

$$p_{\tilde{A}|A^nI}(\tilde{a}|a^ni) = \delta_{\tilde{a},a_i}. \tag{162}$$

Therefore,

$$p_{\tilde{A}|\tilde{X}X^nA^n}(\tilde{a}|\tilde{x}x^na^n) = \sum_{i=1}^{n} p_{\tilde{A}|A^nIX^n\tilde{X}}(\tilde{a}|a^nix^n\tilde{x})p_{I|\tilde{X}X^nA^n}(i|\tilde{x}x^na^n) \tag{163}$$

$$= \sum_{i=1}^{n} p_{\tilde{A}|A^nI}(\tilde{a}|a^ni)p_{I|\tilde{X}X^n}(i|\tilde{x}x^n). \tag{164}$$

$$= \sum_{i=1}^{n} \frac{1}{N(\tilde{x}|x^n)}\delta_{\tilde{x}x_i}\delta_{\tilde{a}a_i}. \tag{165}$$

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:

$$p_{\tilde{A}|\tilde{X}X^nA^n}(\tilde{a}|\tilde{x}, x^n, a^n) := \begin{cases} \dfrac{1}{|A|}, & \text{if } N(\tilde{x}|x^n) = 0 \\ \sum_{i=1}^{n} \dfrac{1}{N(\tilde{x}|x^n)}\delta_{\tilde{x},x_i}\delta_{a_i,\tilde{a}} & \text{else.} \end{cases} \tag{166}$$

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 [OR01] for the analysis.

Definition 17 (Robust typicality [OR01]). 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

$$T_\delta^{X^n} := \left\{x^n: \forall\, \mathcal{X}, \left|\frac{1}{n}N(x|x^n) - p_X(x)\right| \leqslant \delta p(x)\right\}. \tag{167}$$

The following result holds for $0 < \delta < 1$:

Property 18. The probability of a sequence x^n to be in the robustly typical set is bounded from below as

$$\Pr\{X^n \in T_\delta^{X^n}\} \geqslant 1 - 2|\mathcal{X}|\exp^{-\frac{n\delta^2\mu_X}{3}}, \tag{168}$$

where

$$\mu_X := \min_{x\in\mathcal{X}, p_X(x)>0} p_X(x). \tag{169}$$

The state generated after the application of the algorithm in (166) is as follows:

$$\sigma_{\tilde{A}\tilde{X}B} = \sum_{\tilde{x},\tilde{a}} p_{\tilde{X}}(\tilde{x})|\tilde{x}\rangle\langle\tilde{x}|_{\tilde{X}} \otimes \sum_{x^n,a^n} p_{\tilde{A}|\tilde{X}X^nA^n}(\tilde{a}|\tilde{x}, x^n, a^n)p_{X^n}(x^n)q_{\bar{A}^n|X^n}(a^n|x^n)|\tilde{a}\rangle\langle\tilde{a}|_{\tilde{A}} \otimes \sigma_B^{a^n,x^n}. \tag{170}$$

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

$$\sigma_{\tilde{A}\tilde{X}B} = \sum_{\tilde{x},\tilde{a}} p_{\tilde{X}}(\tilde{x})|\tilde{x}\rangle\langle\tilde{x}|_{\tilde{X}} \otimes \Bigg(\sum_{x^n\in S_1(\tilde{x}),a^n} p(\tilde{a}|\tilde{x},\, x^n,\, a^n)|\tilde{a}\rangle\langle\tilde{a}| \otimes q(a^n,\, x^n)\sigma_B^{a^n,x^n}$$

$$+ \sum_{x^n\in S_2(\tilde{x}),a^n} p(\tilde{a}|\tilde{x},\, x^n,\, a^n)|\tilde{a}\rangle\langle\tilde{a}| \otimes q(a^n,\, x^n)\sigma_B^{a^n,x^n} + \sum_{x^n\in S_3,a^n} p(\tilde{a}|\tilde{x},\, x^n,\, a^n)|\tilde{a}\rangle\langle\tilde{a}| \otimes q(a^n,\, x^n)\sigma_B^{a^n,x^n} \Bigg), \quad (171)$$

$$\sigma_{\tilde{A}\tilde{X}B} = \sigma^{(1)}_{\tilde{A}\tilde{X}B} + \sigma^{(2)}_{\tilde{A}\tilde{X}B} + \sigma^{(3)}_{\tilde{A}\tilde{X}B}. \quad (172)$$

From the triangle inequality, we obtain the following:

$$\|\rho_{\tilde{A}\tilde{X}B} - \sigma_{\tilde{A}\tilde{X}B}\|_1 \leqslant \|\rho_{\tilde{A}\tilde{X}B} - \sigma^{(1)}_{\tilde{A}\tilde{X}B}\|_1 + \|\sigma^{(2)}_{\tilde{A}\tilde{X}B}\|_1 + \|\sigma^{(3)}_{\tilde{A}\tilde{X}B}\|_1, \quad (173)$$

where

$$\|\rho_{\tilde{A}XB} - \sigma^{(1)}_{\tilde{A}\tilde{X}B}\|_1 \leqslant \left\| \rho_{\tilde{A}\tilde{X}B} - \frac{1}{n}\sum_{i=1}^n \sigma_{\bar{A}_i X_i B} \right\|_1 + \left\| \frac{1}{n}\sum_{i=1}^n \sigma_{\bar{A}_i X_i B} - \sigma^{(1)}_{\tilde{A}\tilde{X}B} \right\|_1 \quad (174)$$

$$\leqslant nt + \left\| \frac{1}{n}\sum_{i=1}^n \sigma_{\bar{A}_i X_i B} - \sigma^{(1)}_{\tilde{A}\tilde{X}B} \right\|_1. \quad (175)$$

Let us analyze each term individually, beginning with

$$\|\sigma^{(3)}_{\tilde{A}\tilde{X}B}\|_1 = \left\| \sum_{\tilde{x},\tilde{a}} p_{\tilde{X}}(\tilde{x})|\tilde{x}\rangle\langle\tilde{x}|_{\tilde{X}} \otimes \sum_{x^n\in S_3,a^n} p(\tilde{a}|\tilde{x},\, x^n,\, a^n)|\tilde{a}\rangle\langle\tilde{a}| \otimes q(a^n,\, x^n)\sigma_B^{a^n,x^n} \right\|_1 \quad (176)$$

$$\leqslant \sum_{\tilde{x},\tilde{a}} p(\tilde{x}) \sum_{x^n\in S_3,a^n} p(x^n)q(a^n|x^n)p(\tilde{a}|\tilde{x},\, x^n,\, a^n)\||\tilde{x}\rangle\langle\tilde{x}| \otimes |\tilde{a}\rangle\langle\tilde{a}| \otimes \sigma_B^{a^n,x^n}\|_1 \quad (177)$$

$$= \sum_{\tilde{x}} p(\tilde{x}) \sum_{x^n\in S_3} p(x^n) \sum_{a^n} q(a^n|x^n) \sum_{\tilde{a}} p(\tilde{a}|\tilde{x},\, x^n,\, a^n) \leqslant \varepsilon_1, \quad (178)$$

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 (168).

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 17, we know that for the robustly-typical set, the following condition holds

$$x^n\colon \forall\, x \in \mathcal{X}, \quad \left| \frac{1}{n}N(x|x^n) - p_X(x) \right| \leqslant \delta p_X(x). \quad (179)$$

For a robustly-typical sequence to have an empirical distribution $N(x|x^n) = 0$, it is required that $\delta \geqslant 1$. 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,

$$\|\sigma^{(2)}_{\tilde{A}\tilde{X}B}\|_1 = 0. \quad (180)$$

Consider that

$$\sigma^{(1)}_{\tilde{X}\tilde{A}B} = \sum_{\tilde{x}} p(\tilde{x})[\tilde{x}]_{\tilde{X}} \otimes \sum_{a^n,x^n\in S_1(\tilde{x}),\tilde{a}} \times \sum_{i=1}^n \frac{1}{N(\tilde{x}|x^n)}\delta_{a_i,\tilde{a}}\delta_{\tilde{x},x_i}[\tilde{a}]_{\tilde{A}} \otimes p_{X^n}(x^n)q_{A^n|X^n}(a^n|x^n)\sigma_B^{a^n,x^n}, \quad (181)$$

$$= \sum_{\tilde{x}} p(\tilde{x})[\tilde{x}]_{\tilde{X}} \otimes \sum_{\tilde{a}}[\tilde{a}]_{\tilde{A}} \otimes \frac{1}{n}\sum_{i=1}^n \sum_{x^{[n]\setminus\{i\}},\tilde{x}\in S_1(\tilde{x}),a^{[n]\setminus\{i\}}} \frac{p_{\tilde{X}}(\tilde{x})}{N(\tilde{x}|x^n)/n} p_{X^{[n]\setminus\{i\}}}(x^{[n]\setminus\{i\}}|\tilde{x})q(\tilde{a}|x^{[n]\setminus\{i\},\tilde{x}})$$
$$q(a^{[n]\setminus\{i\}}|x^{[n]\setminus\{i\},\tilde{x}}\tilde{a})\sigma_B^{a^{[n]\setminus\{i\}},x^{[n]\setminus\{i\}},\tilde{x},\tilde{a}}, \quad (182)$$

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 (175):

$$\left\| \frac{1}{n}\sum_{i=1}^n \sigma_{\bar{A}_i X_i B} - \sigma^{(1)}_{\tilde{A}XB} \right\|_1, \quad (183)$$

where

$$\sigma_{\bar{A}_i X_i B} = \sum_{a^n, x^n} p_{X^n}(x^n) q_{\bar{A}^n|X^n}(a^n|x^n) |x_i\rangle\langle x_i|_{X_i} \otimes |a_i\rangle\langle a_i|_{\bar{A}_i} \otimes \sigma_B^{a^n, x^n}. \tag{184}$$

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,

$$\begin{aligned}\sigma_{\bar{A}_i X_i B} = &\sum_{a^n, x^n \in S_1(x_i)} p_{X^n}(x^n) q_{\bar{A}^n|X^n}(a^n|x^n) |x_i\rangle\langle x_i|_{X_i} \otimes |a_i\rangle\langle a_i|_{\bar{A}_i} \otimes \sigma_B^{a^n, x^n} \\ &+ \sum_{a^n, x^n \in S_2(x_i)} p_{X^n}(x^n) q_{\bar{A}^n|X^n}(a^n|x^n) |x_i\rangle\langle x_i|_{X_i} \otimes |a_i\rangle\langle a_i|_{\bar{A}_i} \otimes \sigma_B^{a^n, x^n} \\ &+ \sum_{a^n, x^n \in S_3} p_{X^n}(x^n) q_{\bar{A}^n|X^n}(a^n|x^n) |x_i\rangle\langle x_i|_{X_i} \otimes |a_i\rangle\langle a_i|_{\bar{A}_i} \otimes \sigma_B^{a^n, x^n} \qquad (185)\\ = &\sigma^{(1)}_{\bar{A}_i X_i B} + \sigma^{(2)}_{\bar{A}_i X_i B} + \sigma^{(3)}_{\bar{A}_i X_i B}. \qquad (186)\end{aligned}$$

Then, using the convexity of trace distance with (183) and typicality arguments similar to (178) and (180), we find that

$$\left\| \frac{1}{n}\sum_{i=1}^n \sigma_{\bar{A}_i X_i B} - \sigma^{(1)}_{\bar{A}XB} \right\|_1 \leqslant \frac{1}{n}\sum_{i=1}^n \|\sigma_{\bar{A}_i X_i B} - \sigma^{(1),i}_{\bar{A}XB}\|_1 \tag{187}$$

$$\leqslant \frac{1}{n}\sum_{i=1}^n \|\sigma^{(1)}_{\bar{A}_i X_i B} - \sigma^{(1),i}_{\bar{A}XB}\|_1 + \varepsilon_1, \tag{188}$$

where

$$\begin{aligned}\sigma^{(1)}_{\bar{A}_i X_i B} = &\sum_{x_i} p_{X_i}(x_i) [x_i]_{X_i} \otimes \sum_{a_i} [a_i]_{\bar{A}_i} \\ &\otimes \sum_{x^{[n]\setminus\{i\}}, x_i \in S_1(x_i), a^{[n]\setminus\{i\}}} p(x^{[n]\setminus\{i\}}|x_i) q(\tilde{a}|x^{[n]\setminus\{i\}}, x_i) q(a^{[n]\setminus\{i\}}|x^{[n]\setminus\{i\}}, x_i, \tilde{a}) \sigma_B^{a^{[n]\setminus\{i\}} x^{[n]\setminus\{i\}}, x_i, a_i}. \qquad (189)\end{aligned}$$

and

$$\begin{aligned}\sigma^{(1),i}_{\bar{A}XB} = &\sum_{\tilde{x}} p(\tilde{x}) [\tilde{x}]_{\tilde{X}} \otimes \sum_{\tilde{a}} [\tilde{a}]_{\tilde{A}} \otimes \sum_{x^{[n]\setminus\{i\}}, \tilde{x} \in S_1(\tilde{x}), a^{[n]\setminus\{i\}}} \frac{p_{\tilde{X}}(\tilde{x})}{N(\tilde{x}|x^n)/n} p_{X^{[n]\setminus\{i\}}}(x^{[n]\setminus\{i\}}|\tilde{x}) q(\tilde{a}|x^{[n]\setminus\{i\}, \tilde{x}}) \\ &q(a^{[n]\setminus\{i\}}|x^{[n]\setminus\{i\}, \tilde{x}}\tilde{a}) \sigma_B^{a^{[n]\setminus\{i\}}, x^{[n]\setminus\{i\}}, \tilde{x}, \tilde{a}}. \qquad (190)\end{aligned}$$

Invoking (179), we find that

$$\frac{1}{n}\sum_{i=1}^n \|\sigma^{(1)}_{\bar{A}_i X_i B} - \sigma^{(1),i}_{\bar{A}XB}\|_1 \leqslant \frac{\delta}{1-\delta}, \tag{191}$$

where $\delta \in (0, 1)$. After combining (178), (180), (188), and (191), we obtain

$$\|\rho_{\bar{A}XB} - \sigma_{\bar{A}\tilde{X}B}\|_1 \leqslant nt + \frac{\delta}{1-\delta} + 2\varepsilon_1. \tag{192}$$

Minimizing over all possible no-signaling extensions, as required by the definition, we find that

$$\|\rho_{\bar{A}XB} - \sigma_{\bar{A}\tilde{X}B}\|_1 \leqslant n \inf_{\rho_{\bar{A}XBE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1. \tag{193}$$

Since $\rho_{\bar{A}XB}$ and $\sigma_{\bar{A}XB}$ are classical quantum states with $p_X(x) = \frac{1}{|\mathcal{X}|}$, we obtain

$$\sum_x \|\rho^x_{\bar{A}B} - \sigma^x_{\bar{A}B}\|_1 \leqslant |\mathcal{X}| \left(n \inf_{\rho_{\bar{A}XBE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1 \right). \tag{194}$$

This implies that the following inequality holds for all $x \in \mathcal{X}$:

$$\|\rho^x_{\bar{A}B} - \sigma^x_{\bar{A}B}\|_1 \leqslant |\mathcal{X}| \left(n \inf_{\rho_{\bar{A}XBE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1 \right). \tag{195}$$

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:

$$\sup_{p_X(x)} \|\rho_{\bar{A}BX} - \sigma_{\bar{A}BX}\|_1 \leqslant |\mathcal{X}|\left(n \sup_{p_X(x)} \inf_{\rho_{\bar{A}XBE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1\right), \tag{196}$$

which implies that

$$\sup_{p_X(x)} \|\rho_{\bar{A}BX} - \sigma_{\bar{A}BX}\|_1 \leqslant |\mathcal{X}|\left(n\sqrt{S(\bar{A};\, B)_{\hat{\rho}} \ln 2} + \frac{\delta}{1-\delta} + 2\varepsilon_1\right). \tag{197}$$

Given $S(\bar{A};\, B)_{\hat{\rho}} \leqslant \varepsilon$ (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 [OR01] that $\varepsilon_1 = 2|\mathcal{X}|e^{-\frac{1}{3|\mathcal{X}|}\delta^2 n}$. Substituting these values, we find that

$$\|\rho_{\bar{A}BX} - \sigma_{\bar{A}BX}\|_1 \leqslant |\mathcal{X}|\left(\varepsilon^{1/4} + \frac{\varepsilon^{1/16}|\mathcal{X}|^{1/2}}{1-\varepsilon^{1/16}|\mathcal{X}|^{1/2}} + 4|\mathcal{X}|e^{-\frac{\varepsilon^{-1/4}}{3}}\right). \tag{198}$$

This concludes the proof. ■

5. Faithfulness of intrinsic non-locality

The following theorem, combined with proposition 7, establishes that intrinsic non-locality is faithful.

Theorem 19 (**Faithfulness of intrinsic non-locality**). *For every no-signaling or quantum correlation $p(a,\, b|x,\, y)$, the intrinsic non-locality $N(\bar{A};\, \bar{B})_p = 0$, if and only if it has a LHV description. Quantitatively, if $N(\bar{A};\, \bar{B})_p \leqslant \varepsilon$, 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 LHV description, such that*

$$\sup_{p_{XY}(x,y)} \|\rho_{\bar{A}X\bar{B}Y} - \gamma_{\bar{A}X\bar{B}Y}\|_1 \leqslant d\left(\varepsilon^{1/4} + \frac{\varepsilon^{1/16}d^{1/2}}{1-\varepsilon^{1/16}d^{1/2}} + 4de^{-\frac{\varepsilon^{-1/4}}{3}}\right), \tag{199}$$

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 intrinsic steerability. 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 [FR15], we know that there exists a recovery channel $\mathcal{R}_{XE\to\bar{A}XE}$ such that

$$\|\rho_{\bar{A}X\bar{B}YE} - \mathcal{R}_{XE\to\bar{A}XE}(\rho_{\bar{B}YE} \otimes \rho_X)\|_1 \leqslant \sqrt{I(\bar{A};\, \bar{B}Y|XE)_\rho \ln 2} = t. \tag{200}$$

Since $I(\bar{B}E;\, X|Y)_\rho = 0$ from (21), and $p_{XY}(x,\, y) = \frac{1}{x}.\frac{1}{y}$, we can write $\rho_{\bar{B}XYE} = \rho_{\bar{B}YE} \otimes \rho_X$. Following an argument similar to (148)–(151), we obtain the following inequality:

$$\|\rho_{\bar{A}\bar{B}XY} - \omega_{A_iX_iBY}\| \leqslant nt, \tag{201}$$

where

$$\omega_{\bar{A}^nX^n\bar{B}YE} = \bigcirc_{i=1}^n \mathcal{R}_{X_iE\to\bar{A}_iX_iE}(\rho_{\bar{B}YE} \otimes \rho_X^{\otimes n}), \tag{202}$$

$$\omega_{\bar{A}_iX_iBY} = \mathrm{Tr}_{E\bar{A}^{n/i}X^{n/i}}(\omega_{\bar{A}^nX^n\bar{B}YE}). \tag{203}$$

Since the distributions $p_X(x)$ and $p_Y(y)$ are independent, we have

$$I(X^n;\, Y)_p = 0. \tag{204}$$

From the no-signaling constraints, we have

$$I(X^nY;\, E)_\rho = 0. \tag{205}$$

This implies that

$$I(X^nE;\, Y)_\rho = I(X^n;\, Y)_\rho + I(E;\, Y|X^n)_\rho = 0. \tag{206}$$

Since the systems $\bar{A}^nX^nE$ of $\omega_{\bar{A}^nX^n\bar{B}YE}$ are obtained from the application of the recovery channel on systems X_nE of the state $\rho_{X_nYE\bar{B}}$, we can use quantum data processing for mutual information to obtain the following inequality:

$$I(A^nX^n;\, Y)_\omega = 0. \tag{207}$$

This implies that

$$\omega_{\bar{A}^n X^n \bar{B} Y} = \sum_{x^n, a^n, y, b} p(x^n) q(a^n|x^n)\, p(y) q(b|a^n x^n y) [x^n\ a^n\ b\ y]_{X^n \bar{A}^n \bar{B} Y}. \tag{208}$$

Alice's strategy is exactly the same as before, and the following state is obtained after the application of the algorithm in (166):

$$\gamma_{\tilde{A}\tilde{X}\bar{B}Y} := \sum_{\tilde{x}, \tilde{a}, b, y} p_X(\tilde{x}) |\tilde{x}\rangle\langle\tilde{x}|_{\tilde{X}} \otimes \sum_{x^n, a^n,} p_{\tilde{A}|\tilde{X} X^n A^n}(\tilde{a}|\tilde{x}, x^n, a^n) p_{X^n}(x^n) q_{A^n|X^n}(a^n|x^n) p(y) q(b|a^n x^n y) [\tilde{a}\ b\ y]_{\tilde{A}\bar{B}Y}. \tag{209}$$

Note that this state is a local-hidden-variable state. This construction of the local-hidden-variable state shares some similarities with [TDS03]. By following the arguments given for the proof of faithfulness of intrinsic steerability, we obtain

$$\|\rho_{\tilde{A}X\bar{B}Y} - \gamma_{\tilde{A}\tilde{X}\bar{B}Y}\|_1 \leqslant nt + \frac{\delta}{1-\delta} + 2\varepsilon_1. \tag{210}$$

This implies

$$\|\rho_{\tilde{A}X\bar{B}Y} - \gamma_{\tilde{A}\tilde{X}\bar{B}Y}\|_1 \leqslant n \inf_{\rho_{\tilde{A}X\bar{B}YE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1. \tag{211}$$

This implies

$$\sum_{a,b} |p(a, b|x, y) - l(a, b|x, y)| \leqslant |\mathcal{X}||\mathcal{Y}| \left(\inf_{\rho_{\tilde{A}X\bar{B}YE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1 \right) \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}. \tag{212}$$

Now, using triangle inequality, we obtain the following for any arbitrary distribution $p(x, y)$:

$$\|\rho_{\tilde{A}X\bar{B}Y} - \gamma_{\tilde{A}\tilde{X}\bar{B}Y}\|_1 \leqslant |\mathcal{X}||\mathcal{Y}| \left(\inf_{\rho_{\tilde{A}X\bar{B}YE}} t + \frac{\delta}{1-\delta} + 2\varepsilon_1 \right). \tag{213}$$

This implies

$$\sup_{p_{XY}(x,y)} \|\rho_{\tilde{A}X\bar{B}Y} - \gamma_{\tilde{A}\tilde{X}\bar{B}Y}\|_1 \leqslant |\mathcal{X}||\mathcal{Y}| \left(\sqrt{N(\bar{A}; \bar{B})_p \ln 2} + \frac{\delta}{1-\delta} + 2\varepsilon_1 \right). \tag{214}$$

Given $N(\bar{A}; \bar{B})_p \leqslant \varepsilon$ (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 [OR01] that $\varepsilon_1 = 2|\mathcal{X}||\mathcal{Y}| e^{-\frac{1}{3|\mathcal{X}||\mathcal{Y}|}\delta^2 n}$. Substituting these values, we obtain

$$\|\rho_{\tilde{A}X\bar{B}Y} - \gamma_{\tilde{A}\tilde{X}\bar{B}Y}\|_1 \leqslant |\mathcal{X}| \cdot |\mathcal{Y}| \left(\varepsilon^{1/4} + \frac{\varepsilon^{1/16} |\mathcal{X}|^{1/2} \cdot |\mathcal{Y}|^{1/2}}{1 + \varepsilon^{1/16} |\mathcal{X}|^{1/2} \cdot |\mathcal{Y}|^{1/2}} + 4|\mathcal{X}| \cdot |\mathcal{Y}| e^{-\frac{\varepsilon^{-1/4}}{3}} \right). \tag{215}$$

This concludes the proof. ■

Corollary 20 (Faithfulness of quantum intrinsic non-locality). *For every quantum correlation $p(a, b|x, y)$, the quantum intrinsic non-locality $N^Q(\bar{A}; \bar{B})_p = 0$, if and only if it has a LHV description. Quantitatively, if $N^Q(\bar{A}; \bar{B})_p \leqslant \varepsilon$, 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*

$$\sup_{p_{XY}(x,y)} \|\rho_{\tilde{A}X\bar{B}Y} - \gamma_{\tilde{A}X\bar{B}Y}\|_1 \leqslant d \left(\varepsilon^{1/4} + \frac{\varepsilon^{1/16} d^{1/2}}{1 - \varepsilon^{1/16} d^{1/2}} + 4d e^{-\frac{\varepsilon^{-1/4}}{3}} \right), \tag{216}$$

where $\rho_{\tilde{A}X\bar{B}Y}$ correponds to the classical–classical state $p_{XY}(x, y) p(a, b|x, y)$ and $\gamma_{\tilde{A}X\bar{B}Y}$ 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 7. The only-if part follows from proposition 6 and theorem 19. ■

6. Upper bounds on secret key rates in device-independent QKD

We now consider the task of device-independent QKD. We consider two honest parties, Alice and Bob, who share a two-component device and want to extract a shared secret key from this device.

In general, in the device-independent literature, many prior works have devised lower bounds on the key rates for particular protocols, as done in [ABG+07, AFDF+18]. By a protocol, we mean a sequence of steps in which Alice and Bob interact with their devices and communicate publicly with each other.

Here, we are interested in a different question. We fix the black-box device that is shared by Alice and Bob. We assume that the correlations generated from this device are characterized by a correlation $p(a, b|x, y)$. We then pose the following question:

> Given a device characterized by $p(a, b|x, y)$, what is a non-trivial upper bound on the secret-key rate that can be extracted from this device with any possible protocol?

We answer this question for an i.i.d. device, which means that in each round of the protocol, the device considered is characterized by the correlation $p(a, b|x, y)$. 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 the correlation $p(a, b|x, y)$ is not a drawback since we are interested in determining upper bounds on secret-key rates here. In what follows, we prove that the quantifiers introduced above are upper bounds on the secret-key rates that can be generated from the device.

In device-independent 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 (10). Let the correlation shared between Alice, Bob and Eve be defined by $p(a, b|x, y)\rho_E^{a,b,x,y}$. If $p(a, b|x, y)\rho_E^{a,b,x,y}$ has an underlying quantum strategy as described in (27), 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 (20) and (22), then we call the eavesdropper a no-signaling Eve.

6.1. Device-independent protocols

We now state the general form of a device-independent protocol with no-signaling eavesdropper for which our upper bounds hold. Such protocols have previously been considered in [BHK05, Mas09, MRC+14]. Let $n \in \mathbb{Z}^+$, $R \geqslant 0$, and $\varepsilon \in [0, 1]$. Let $p(a, b|x, y)$ be the correlation of the device shared between Alice and Bob. We define an (n, R, ε) device-independent secret-key agreement protocol as follows:

- Alice and Bob give 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
$$\sigma_{\bar{A}^n\bar{B}^nX^nY^n} := \sum_{x^n,y^n,a^n,b^n} p_{X^nY^n}(x^n, y^n)p^n(a^n, b^n|x^n, y^n)[a^nb^nx^ny^n]_{\bar{A}^n\bar{B}^nX^nY^n}, \tag{217}$$
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, C_B the classical register 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 $\omega_{K_AK_BE\bar{C}_A\bar{C}_BX^nY^n}$ that satisfies
$$\|\omega_{K_AK_BEX^nY^n\bar{C}_A\bar{C}_B} - \overline{\Phi}_{K_AK_B} \otimes \omega_{EX^nY^n\bar{C}_A\bar{C}_B}\|_1 \leqslant \varepsilon, \tag{218}$$
for all no-signaling extensions, where
$$\overline{\Phi}_{K_AK_B} = \frac{1}{2^{nR}}\sum_{k=1}^{2^{nR}} |kk\rangle\langle kk|_{K_AK_B}. \tag{219}$$

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.

Theorem 21. *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:*
$$DI(p) \leqslant N(\bar{A}; \bar{B})_p. \tag{220}$$

Proof. For an arbitrary (n, R, ε) protocol, consider that
$$nR = I(K_A; K_B|EX^nY^n\bar{C}_A\bar{C}_B)_{\tilde{\Phi}\otimes\omega} \tag{221}$$

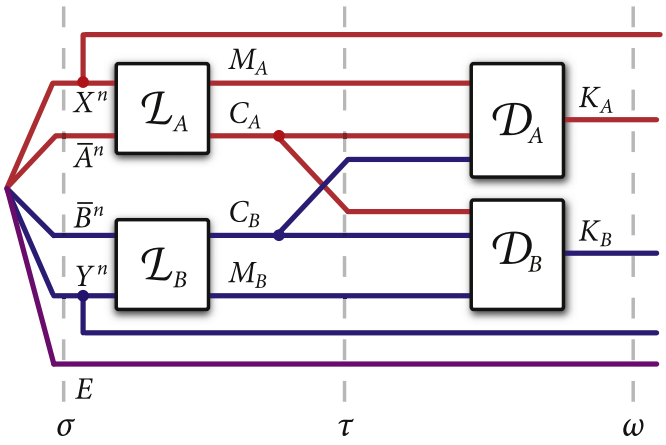

Figure 2. This figure depicts a device-independent secret-key agreement protocol, along with the various states involved at each time step. The classical systems $\bar{A}^n X^n \bar{B}^n Y^n$ are processed by local operations and classical communication to produce classical key systems K_A and K_B that are approximately independent of the no-signaling extension system E, the input systems X^n and Y^n, and copies $\bar{C}_A$ and $\bar{C}_B$ of the classical systems C_A and C_B, respectively.

$$\leqslant I(K_A;\, K_B|EX^nY^n\bar{C}_A\bar{C}_B)_\omega + \varepsilon' \tag{222}$$

$$\leqslant I(M_AC_BC_A;\, M_BC_BC_A|EX^nY^n\bar{C}_A\bar{C}_B)_\tau + \varepsilon' \tag{223}$$

$$= I(M_AC_A;\, M_BC_B|EX^nY^n\bar{C}_A\bar{C}_B)_\tau + \varepsilon' \tag{224}$$

$$\leqslant I(M_AC_A;\, M_BC_B|EX^nY^n)_\tau + \varepsilon' \tag{225}$$

$$\leqslant I(\bar{A}^n;\, \bar{B}^n|EX^nY^n)_\sigma + \varepsilon', \tag{226}$$

where

$$\varepsilon' = nR\varepsilon + 2[(1+\varepsilon)\log(1+\varepsilon)) - \varepsilon\log\varepsilon]. \tag{227}$$

In the above equations, $\sigma_{X^n\bar{A}^n\bar{B}^nY^n}$ 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 $\sigma_{X^n\bar{A}^n\bar{B}^nY^nE}$. 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 2.

The first inequality follows from the uniform continuity of conditional mutual information [Shi17, Proposition 1]. The second inequality follows from data processing. The second equality and third inequality follow from 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 $\sigma_{\bar{A}^n\bar{B}^nX^nY^n}$, we find that

$$nR \leqslant \inf_{\sigma_{\bar{A}^n\bar{B}^nX^nY^nE}} I(\bar{A}^n;\, \bar{B}^n|X^nY^nE)_\sigma + \varepsilon'. \tag{228}$$

This implies that

$$nR \leqslant N(\bar{A}^n;\, \bar{B}^n)_p + \varepsilon'. \tag{229}$$

By the assumption that the device is i.i.d, we can invoke the additivity of intrinsic non-locality from proposition 12 to obtain

$$(1-\varepsilon)R \leqslant N(\bar{A};\, \bar{B})_p + 2[(1+\varepsilon)\log(1+\varepsilon)) - \varepsilon\log\varepsilon]/n. \tag{230}$$

Taking the limit as $n \to \infty$ and $\varepsilon \to 0$ then leads to $DI(p) \leqslant N(\bar{A};\, \bar{B})_p$. ■

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 [ABG+07, AFDF+18]. The general form of a device-independent protocol with a quantum eavesdropper remains the same except that we now consider a quantum extension (27) of the state in (217). We then arrive at the following theorem:

Theorem 22. *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:*

$$DI(p) \leqslant N^{Q}(\bar{A};\bar{B})_{p}. \tag{231}$$

Proof. The proof of the theorem is similar to that of theorem 21. ■

We should explicitly point out that the general form for protocols that we consider allow 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 the one-way error-correcting protocol [BA07, WMUK07, KL17]. This question has only recently been explored in DI-QKD in [TLR19]. Therefore, it is possible that the upper bound via the intrinsic non-locality will not be tight for the existing DI-QKD protocols [ABG+07, AFDF+18] 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 [RPMP15] (note that [SARG04] introduced this concept in the device-dependent QKD literature). For such DI-QKD protocols, our upper bounds do not hold.

6.1.1. Other considerations

Bounds on device-independent QKD protocols based on certain states were also previously discussed in [HM15].

There is yet another way to model a no-signaling adversary in the device-independent secret agreement protocols which has been considered in [BHK05]. This model is set in 'box world', where 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 [WDH19], the authors introduced squashed non-locality to provide 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 number of measurements.

6.2. One-SDI protocol

Let $n \in \mathbb{Z}^{+}$, $R \geqslant 0$, and $\varepsilon \in [0,\ 1]$. We define an $(n,\ R,\ \varepsilon)$ one-SDI 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

$$\rho_{\bar{A}^{n}X^{n}B^{n}} := \sum_{x^{n},a^{n}} p_{X^{n}}(x^{n})p_{A^{n}|X^{n}}(a^{n}|x^{n})[x^{n},\ a^{n}]_{X^{n}A^{n}} \otimes \rho_{B^{n}}^{a^{n},x^{n}}, \tag{232}$$

  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:

$$\rho_{\bar{A}^{n}X^{n}B^{n}E} := \sum_{x^{n},a^{n}} p_{X^{n}}(x^{n})p_{A^{n}|X^{n}}(a^{n}|x^{n})[x^{n},\ a^{n}]_{X^{n}A^{n}} \otimes \rho_{B^{n}E}^{a^{n},x^{n}}. \tag{233}$$

- 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 $\sigma_{\bar{A}^{n}X^{n}\bar{B}^{n}Y^{n}E}$.

$$\begin{aligned}\sigma_{\bar{A}^{n}X^{n}Y^{n}\bar{B}^{n}E} := \sum_{x^{n},a^{n}} p_{X^{n}}(x^{n})p_{\bar{A}^{n}|X^{n}}(a^{n}|x^{n})[x^{n},\ a^{n}]_{X^{n}\bar{A}^{n}} \otimes \sum_{y^{n}b^{n}} p_{Y^{n}}(y^{n})[y^{n}]_{Y_{n}} \\ \otimes (Y_{b^{n}}^{n}\rho_{B^{n}E}^{a^{n},x^{n}}(Y_{b^{n}}^{n})^{\dagger}).\end{aligned} \tag{234}$$

- 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}, C_{B} the classical register 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

$$\|\omega_{K_{A}K_{B}EX^{n}Y^{n}\bar{C}_{A}\bar{C}_{B}} - \overline{\Phi}_{K_{A}K_{B}} \otimes \omega_{EX^{n}Y^{n}\bar{C}_{A}\bar{C}_{B}}\|_{1} \leqslant \varepsilon, \tag{235}$$

  for all no-signaling extensions, where

$$\overline{\Phi}_{K_A K_B} = \frac{1}{2^{nR}} \sum_{k=1}^{2^{nR}} |kk\rangle \langle kk|_{K_A K_B}. \tag{236}$$

A rate R is achievable for a device characterized by $\hat{\rho}$ if there exists an $(n, R - \delta, \varepsilon)$ one-SDI protocol for all $\varepsilon \in (0, 1)$, $\delta > 0$, and sufficiently large n. The one-SDI capacity $\mathrm{SDI}(\hat{\rho})$ of the device characterized by $\hat{\rho}$ is defined as the supremum of all achievable rates for $\hat{\rho}$.

Theorem 23. *The restricted intrinsic steerability $S(\bar{A}; \bar{B})_{\hat{\rho}}$ is an upper bound on the one-SDI secret-key-agreement capacity $\mathrm{SDI}(\hat{\rho})$ of a device characterized by $\hat{\rho}$:*

$$\mathrm{SDI}(\hat{\rho}) \leqslant S(\bar{A}; B)_{\hat{\rho}}. \tag{237}$$

Proof. For obtaining the upper bound in the one-SDI setting, we continue from (226) as follows:

$$nR \leqslant I(\bar{A}^n; \bar{B}^n Y^n | E X^n)_\sigma - I(\bar{A}^n; Y^n | E X^n)_\sigma + \varepsilon' \tag{238}$$

$$\leqslant I(\bar{A}^n; \bar{B}^n Y^n | E X^n)_\sigma + \varepsilon' \tag{239}$$

$$\leqslant I(\bar{A}^n; B^n | E X^n)_\rho + \varepsilon'. \tag{240}$$

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

$$nR \leqslant \inf_{\rho_{\bar{A}^n X^n B^n E}} I(\bar{A}^n; B^n | X^n E)_\rho + \varepsilon'. \tag{241}$$

This implies that

$$nR \leqslant S(\bar{A}^n; B^n)_{\hat{\rho}} + \varepsilon'. \tag{242}$$

Since we assume an i.i.d device, we find by applying the additivity of restricted intrinsic steerability [KWW17] that

$$(1 - \varepsilon) R \leqslant S(\bar{A}; B)_{\hat{\rho}} + 2[(1 + \varepsilon)\log(1 + \varepsilon)) - \varepsilon \log \varepsilon]/n. \tag{243}$$

Taking the limit as $n \to \infty$ and $\varepsilon \to 0$ then leads to the desired inequality $\mathrm{SDI}(\hat{\rho}) \leqslant S(\bar{A}; B)_{\hat{\rho}}$. ■

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 [HHHO09].

Proposition 24. *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 $p(a, b|x, y)$ be 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-SDI secret-key-agreement capacity of $\hat{\rho}$, which in turn does not exceed the distillable key of the state ρ_{AB}:*

$$\mathrm{DI}(p) \leqslant \mathrm{SDI}(\hat{\rho}) \leqslant K(\rho_{AB}). \tag{244}$$

Proof. The proof is a consequence of the following observation: the DI secret-key-agreement protocol is a special case of the SDI secret-key-agreement 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. Examples

7.1. 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 $\omega_{AB}^p = (1 - p)\Phi_{AB} + p\pi_A \otimes \pi_B$, 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 [ABG+07], which was then used to obtain a lower bound on the key rate from the above specified correlation.

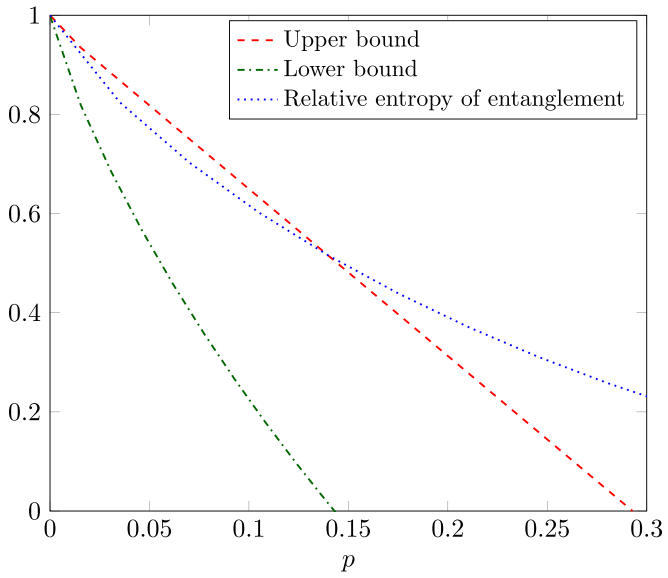

Figure 3. In this figure, we plot the upper bound in (253) and the lower bound from [ABG+07] for the device-independent protocol described in section 7.1. Relative entropy of entanglement of qubit-qubit isotropic state is given in [VPRK97]. For further explanation of this plot, see the next section.

The secret-key rate in a device-independent protocol is bounded from above as follows (Theorem 22):

$$R \leqslant \sup_{p(x,y)} \inf_{\rho_{\bar{A}\bar{B}XYE}} \sum_{x,y} p_{XY}(x, y) I(\bar{A}; \bar{B}|E)_{x,y}. \tag{245}$$

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 [GEW16] to obtain upper bounds on the squashed entanglement of a depolarizing channel. An isotropic state is Bell local if $p \geqslant 1 - \frac{1}{\sqrt{2}}$ [HHH95]. This implies that the quantum intrinsic non-locality of a correlation derived from ω_{AB}^{p} is equal to zero for $p \geqslant 1 - \frac{1}{\sqrt{2}}$ (Proposition 7). For $\epsilon \leqslant p \leqslant 1 - \frac{1}{\sqrt{2}}$, we can write the probability distribution $q_{\omega^p}(a, b|x, y)$ obtained from ω_{AB}^{p} as a convex combination of probability distributions obtained from ω^{ϵ} and $\omega^{1-1/\sqrt{2}}$. That is, for some $0 \leqslant \alpha \leqslant 1$, we have

$$q_{\omega^p}(a, b|x, y) = (1 - \alpha(\epsilon)) q_{\omega^{\epsilon}}(a, b|x, y) + \alpha(\epsilon) q_{1-\omega^{1/\sqrt{2}}}(a, b|x, y). \tag{246}$$

By simple algebra, we obtain

$$\alpha(\epsilon) = \frac{p - \epsilon}{1 - \frac{1}{\sqrt{2}} - \epsilon}. \tag{247}$$

Equation (246) can be written as

$$q_{\omega^p}(a, b|x, y) = (1 - \alpha(\epsilon)) q_{\omega^{\epsilon}}(a, b|x, y) + \alpha(\epsilon) \sum_{\lambda} p(\lambda) q_{\omega^{1/\sqrt{2}}}(a, |x, \lambda) q_{\omega^{1/\sqrt{2}}}(b, |y, \lambda). \tag{248}$$

Then, from convexity of quantum intrinsic non-locality (Proposition 11), we obtain

$$N^{Q}(\bar{A}; \bar{B})_{q_{\omega^p}} \leqslant (1 - \alpha(\epsilon)) N^{Q}(\bar{A}; \bar{B})_{q_{\omega^{\epsilon}}}. \tag{249}$$

Since the above equation is true for all α, we find that

$$N^{Q}(\bar{A}; \bar{B})_{q_{\omega^p}} \leqslant \min_{0 \leqslant \epsilon \leqslant p} (1 - \alpha(\epsilon)) N^{Q}(\bar{A}; \bar{B})_{q_{\omega^{\epsilon}}}. \tag{250}$$

This implies that

$$N^{Q}(\bar{A}; \bar{B})_{q_{\omega^p}} \leqslant \min_{0 \leqslant \epsilon \leqslant p} (1 - \alpha(\epsilon)) \sup_{p(x,y)} \inf_{\rho_{\bar{A}\bar{B}XYE}(\epsilon)} \sum_{x,y} p(x, y) I(\bar{A}; \bar{B}|E)_{\rho^{x,y}_{\bar{A}\bar{B}E}(\epsilon)}, \tag{251}$$

where q_{ω}^{ϵ} is encoded in $\rho_{\bar{A}\bar{B}XY}(\epsilon)$ with $\rho_{\bar{A}\bar{B}XYE}(\epsilon)$ as the quantum extension. Let us choose a trivial extension of the state $\rho^{x,y}_{\bar{A}\bar{B}}(\epsilon)$. It is easy to see that

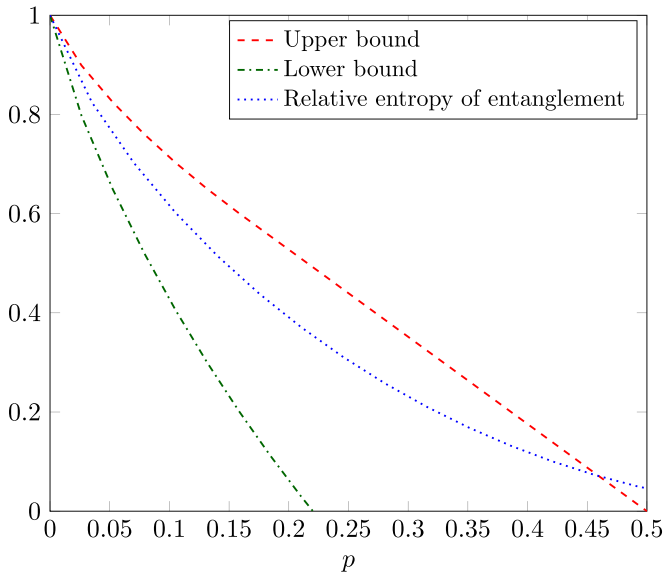

Figure 4. In this figure, we plot the upper bound in (261) and the lower bound from [BCW+12] for the one-sided-device-independent protocol described in section 7.2. The relative entropy of entanglement of a qubit-qubit isotropic state is given in [VPRK97].

$$I(\bar{A};\bar{B})_{\rho^{0,1}_{\bar{A}\bar{B}}(\epsilon)} \geqslant I(\bar{A};\bar{B})_{\rho^{x,y}_{\bar{A}\bar{B}}(\epsilon)} \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}. \tag{252}$$

Therefore,

$$R \leqslant \min_{0\leqslant\epsilon\leqslant p}(1-\alpha(\epsilon))I(\bar{A};\bar{B})_{\rho^{0,1}_{\bar{A}\bar{B}}(\epsilon)} = \min_{0\leqslant\epsilon\leqslant p}(1-\alpha(\epsilon))\left(\frac{2-\epsilon}{2}\log_2(2-\epsilon)+\frac{\epsilon}{2}\log_2\epsilon\right). \tag{253}$$

We plot this upper bound in figure 3, and we interpret it and explain the relative entropy of entanglement bound in the next subsection.

7.2. 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

$$\begin{aligned}\rho_{X\bar{A}B}(p) = &\frac{1}{4}(|0\rangle\langle 0|_X \otimes [|0\rangle\langle 0|_{\bar{A}} \otimes ((1-p)|0\rangle\langle 0|_B + p\pi_B)])\\ &+\frac{1}{4}(|0\rangle\langle 0|_X \otimes [|1\rangle\langle 1|_{\bar{A}} \otimes ((1-p)|1\rangle\langle 1|_B + p\pi_B)])\\ &+\frac{1}{4}(|1\rangle\langle 1|_X \otimes [|0\rangle\langle 0|_{\bar{A}} \otimes ((1-p)|+\rangle\langle +|_B + p\pi_B)])\\ &+\frac{1}{4}(|1\rangle\langle 1|_X \otimes [|1\rangle\langle 1|_{\bar{A}} \otimes ((1-p)|-\rangle\langle -|_B + p\pi_B)]).\end{aligned} \tag{254}$$

If $p \geqslant 1/2$, it is known that $\rho_{X\bar{A}B}$ is unsteerable [WJD07], and therefore intrinsic steerability is zero for $p \geqslant \frac{1}{2}$ ([KWW17], Proposition 7). For $\epsilon \leqslant p \leqslant \frac{1}{2}$, we can write the $\rho_{X\bar{A}B}(p)$ 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 $0 \leqslant \alpha \leqslant 1$

$$\rho_{X\bar{A}B}(p) = (1-\alpha)\rho_{X\bar{A}B}(\epsilon) + \alpha\rho_{X\bar{A}B}\left(\tfrac{1}{2}\right). \tag{255}$$

Then, by simple algebra we obtain

$$\alpha(\epsilon) = \frac{p-\epsilon}{\frac{1}{2}-\epsilon}. \tag{256}$$

From convexity of intrinsic steerability (Proposition 10 [KWW17]), we obtain

$$S(\bar{A};B)_{\hat{\rho}(p)} \leqslant S(\bar{A};B)_{\hat{\rho}(\epsilon)}. \tag{257}$$

Following the same argument as before, we obtain

$$S(\bar{A};B)_{\hat{\rho}(p)} \leqslant \min_{0\leqslant\epsilon\leqslant p}(1-\alpha(\epsilon))\sup_{p_X(x)}\inf_{\rho_{\bar{A}BXE(\epsilon)}}\sum_{p_X(x)} p_X(x)I(\bar{A};B|E)_{\rho_{\bar{A}BE}(\epsilon)}. \tag{258}$$

Let us now choose a trivial extension of the assemblage. It is easy to see that

$$I(\bar{A};B)_{\rho^0(\epsilon)} = I(\bar{A};B)_{\rho^1(\epsilon)} \tag{259}$$

$$= 1 + \left(\tfrac{\epsilon}{2}\right)\log\left(\tfrac{\epsilon}{2}\right) + \left(1-\tfrac{\epsilon}{2}\right)\log\left(1-\tfrac{\epsilon}{2}\right). \tag{260}$$

We therefore obtain

$$S(\bar{A};B)_\rho = \min_{0\leqslant\epsilon\leqslant p}(1-\alpha(\epsilon))\left(1 + \left(\tfrac{\epsilon}{2}\right)\log\left(\tfrac{\epsilon}{2}\right) + \left(1-\tfrac{\epsilon}{2}\right)\log\left(1-\tfrac{\epsilon}{2}\right)\right). \tag{261}$$

We plot this bound in figure 4.

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 [CEH+07, Wil16], 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 the secret key can be distilled in one-SDI 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 [GEW16] is greater than the bound obtained on intrinsic steerability of the assemblage considered above. Therefore, we do not plot the squashed-entanglement bounds in figure 3 or 4.

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-SDI and device-independent protocols [HHHO09]. The relative entropy of entanglement of qubit-qubit isotropic states has been calculated in [VPRK97], 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 [GA15, KW17] and relative entropy of non-locality [vDGG05] would be useful as upper bounds for one-SDI and device-independent QKD, respectively.

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 SDI-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, pertaining to two-way error correction that is allowed in the protocols considered above.

8. Conclusion and outlook

In the present work, we have introduced information-theoretic measures of non-locality called *intrinsic non-locality* and *quantum intrinsic non-locality*. They are inspired by the intrinsic information [MW99] and have a form similar to squashed entanglement [CW04] and intrinsic steerability [KWW17]. We have proven that intrinsic non-locality and quantum intrinsic non-locality are upper bounds on secret-key rates in device-independent secret-key-agreement protocols. Similarly, we have proven that restricted intrinsic steerability is an upper bound on secret-key rates in one-SDI secret-key-agreement protocols. To our knowledge, this is the first time that monotones of Bell non-locality and steering have been used to obtain upper bounds on device-independent and one-sided-device-independent secret-key rates, respectively. The faithfulness properties for intrinsic steerability and intrinsic non-locality that we have proven here are of independent interest.

We now give an overview of the remaining open problems not addressed by the present work. It is not known if either intrinsic non-locality or intrinsic steerability are asymptotically continuous. A naive approach for establishing these properties is to follow the proof for asymptotic continuity of squashed entanglement [AF04]; however, this approach does not straightforwardly apply due to the no-signaling constraints on the extension system. From a foundational perspective, it would be interesting to provide an example of a probability distribution for which the intrinsic non-locality with a classical no-signaling extension is different from intrinsic non-locality with a quantum no-signaling extension.

We also suspect that the squashed entanglement of a bipartite state ρ_{AB} is greater than or equal to the intrinsic steerability of an assemblage that results from measuring ρ_{AB}. The approach in proposition 14 does not apply because it does not account for the factor of $1/2$ present in the definition of squashed entanglement.

Another promising direction to pursue is to improve the upper bounds on secret-key rates for device-independent and one-sided-device independent protocols. Several works in the classical information theory literature have introduced modifications of classical intrinsic information [RW03, GA10] in order to obtain better bounds on secret-key rates than intrinsic information. In [RW03], a modified measure of intrinsic information, called reduced intrinsic information, was introduced and proved to be a better upper bound on secret-key rate than intrinsic information [MW99]. This bound was also subsequently improved further in

[GA10]. It would be interesting to check if these techniques lead to improvements on the upper bounds presented by intrinsic non-locality and intrinsic steerability.

One of the most important open questions is to determine if the relative entropy of steering [GA15, KW17] and relative entropy of non-locality [vDGG05] would be useful as upper bounds for one-SDI and device-independent secret-key-agreement protocols, respectively. It is possible that this might be the case; if true, it could lead to tighter upper bounds for certain device-independent and one-SDI protocols.

Acknowledgments

We are grateful to Rotem Arnon-Friedman for discussions on device-independent QKD. Eneet Kaur and Mark M Wilde acknowledge support from the US Office of Naval Research and the National Science Foundation under Grant No. 1350397. Andreas Winter acknowledges support from the ERC Advanced Grant IRQUAT, the Spanish MINECO (project FIS2016-86681-P), with the support of FEDER funds, and the Generalitat de Catalunya, CIRIT project 2014-SGR-966.

ORCID iDs

Mark M Wilde https://orcid.org/0000-0002-3916-4462
Andreas Winter https://orcid.org/0000-0001-6344-4870

References

[ABG+07] Acín A, Brunner N, Gisin N, Massar S, Pironio S and Scarani V 2007 Device-independent security of quantum cryptography against collective attacks *Phys. Rev. Lett.* **98** 230501
[AF04] Alicki R and Fannes M 2004 Continuity of quantum conditional information *J. Phys. A: Math. Gen.* **37** L55
[AFDF+18] Arnon-Friedman R, Dupuis F, Fawzi O, Renner R and Vidick T 2018 Practical device-independent quantum cryptography via entropy accumulation *Nat. Commun.* **9** 459
[BA07] Bae J and Acín A 2007 Key distillation from quantum channels using two-way communication protocols *Phys. Rev.* A **75** 012334
[BB84] Bennett C H and Brassard G 1984 Quantum cryptography: public-key distribution and coin tossing *Proc. IEEE Int. Conf. on Computers, Systems and Signal Processing (Bangalore, India)* pp 175–9
[BCP+14] Brunner N, Cavalcanti D, Pironio S, Scarani V and Wehner S 2014 Bell nonlocality *Rev. Mod. Phys.* **86** 419–78
[BCW+12] Branciard C, Cavalcanti E G, Walborn S P, Scarani V and Wiseman H M 2012 One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering *Phys. Rev.* A **85** 010301
[BHK05] Barrett J, Hardy L and Kent A 2005 No signaling and quantum key distribution *Phys. Rev. Lett.* **95** 010503
[CEH+07] Christandl M, Ekert A, Horodecki M, Horodecki P, Oppenheim J and Renner R 2007 Unifying classical and quantum key distillation *Theory of Cryptography* (Berlin: Springer) pp 456–78
[CHSH69] Clauser J F, Horne M A, Shimony A and Holt R A 1969 Proposed experiment to test local hidden-variable theories *Phys. Rev. Lett.* **23** 880–4
[CS17] Cavalcanti D and Skrzypczyk P 2017 Quantum steering: a review with focus on semidefinite programming *Rep. Prog. Phys.* **80** 024001
[CW04] Christandl M and Winter A 2004 Squashed entanglement: an additive entanglement measure *J. Math. Phys.* **45** 829–40
[dV14] de Vicente J I 2014 On nonlocality as a resource theory and nonlocality measures *J. Phys. A: Math. Theor.* **47** 424017
[Eke91] Ekert A K 1991 Quantum cryptography based on Bell's theorem *Phys. Rev. Lett.* **67** 661–3
[FR15] Fawzi O and Renner R 2015 Quantum conditional mutual information and approximate Markov chains *Commun. Math. Phys.* **340** 575–611
[FW11] Forster M and Wolf S 2011 Bipartite units of nonlocality *Phys. Rev.* A **84** 042112
[FWW09] Forster M, Winkler S and Wolf S 2009 Distilling nonlocality *Phys. Rev. Lett.* **102** 120401
[GA10] Gohari A A and Anantharam V 2010 Information-theoretic key agreement of multiple terminals. I *IEEE Trans. Inf. Theory* **56** 3973–96
[GA15] Gallego R and Aolita L 2015 Resource theory of steering *Phys. Rev.* X **5** 041008
[GA17] Gallego R and Aolita L 2017 Nonlocality free wirings and the distinguishability between Bell boxes *Phys. Rev.* A **95** 032118
[GEW16] Goodenough K, Elkouss D and Wehner S 2016 Assessing the performance of quantum repeaters for all phase-insensitive Gaussian bosonic channels *New J. Phys.* **18** 063005
[HHH95] Horodecki R, Horodecki P and Horodecki M 1995 Violating Bell inequality by mixed spin-1/2 states: necessary and sufficient condition *Phys. Lett.* A **200** 340–4
[HHHH09] Horodecki R, Horodecki P, Horodecki M and Horodecki K 2009 Quantum entanglement *Rev. Mod. Phys.* **81** 865–942
[HHHO09] Horodecki K, Horodecki M, Horodecki P and Oppenheim J 2009 General paradigm for distilling classical key from quantum states *IEEE Trans. Inf. Theory* **55** 1898–929
[HM15] Horodecki K and Murta G 2015 Bounds on quantum nonlocality via partial transposition *Phys. Rev.* A **92** 010301
[KL17] Khatri S and Lütkenhaus N 2017 Numerical evidence for bound secrecy from two-way postprocessing in quantum key distribution *Phys. Rev.* A **95** 042320
[KW17] Kaur E and Wilde M M 2017 Relative entropy of steering: on its definition and properties *J. Phys. A: Math. Theor.* **50** 465301
[KWW17] Kaur E, Wang X and Wilde M M 2017 Conditional mutual information and quantum steering *Phys. Rev.* A **96** 022332
[LCT14] Lo H-K, Curty M and Tamaki K 2014 Secure quantum key distribution *Nat. Photon.* **8** 604
[LW18] Li K and Winter A 2018 Squashed entanglement, k-extendibility, quantum Markov chains, and recovery maps *Found. Phys.* **48** 910–24
[Mas09] Masanes L 2009 Universally composable privacy amplification from causality constraints *Phys. Rev. Lett.* **102** 140501

[May01] Mayers D 2001 Unconditional security in quantum cryptography *J. ACM* **48** 351–406
[MRC+14] Masanes L, Renner R, Christandl M, Winter A and Barrett J 2014 Full security of quantum key distribution from no-signaling constraints *IEEE Trans. Inf. Theory* **60** 4973–86
[MW99] Maurer U M and Wolf S 1999 Unconditionally secure key agreement and the intrinsic conditional information *IEEE Trans. Inf. Theory* **45** 499–514
[MY98] Mayers D and Yao A 1998 Quantum cryptography with imperfect apparatus *Proc. 39th Annual Symp. on Foundations of Computer Science, Foundations of Computer Science '98*
[MY04] Mayers D and Yao A 2004 Self testing quantum apparatus *Quantum Inf. Comput.* **4** 273–86
[OR01] Orlitsky A and Roche J R 2001 Coding for computing *IEEE Trans. Inf. Theory* **47** 903–17
[Pit86] Pitowsky I 1986 The range of quantum probability *J. Math. Phys.* **27** 1556–65
[Pus13] Pusey M F 2013 Negativity and steering: a stronger Peres conjecture *Phys. Rev.* A **88** 032313
[RP94] Rorhlich D and Popescu S 1994 Quantum nonlocality as an axiom *Found. Phys.* **24** 379–85
[RPMP15] Rahaman R, Parker M G, Mironowicz P and Pawłowski M 2015 Device-independent quantum key distribution based on measurement inputs *Phys. Rev.* A **92** 062304
[RW03] Renner R and Wolf S 2003 New bounds in secret-key agreement: the gap between formation and secrecy extraction *Advances in Cryptology–EUROCRYPT 2003* (Berlin: Springer) pp 562–77
[SARG04] Scarani V, Acín A, Ribordy G and Gisin N 2004 Quantum cryptography protocols robust against photon number splitting attacks for weak laser pulse implementations *Phys. Rev. Lett.* **92** 057901
[SBC+15] Sainz A B, Brunner N, Cavalcanti D, Skrzypczyk P and Vértesi T 2015 Postquantum steering *Phys. Rev. Lett.* **115** 190403
[SBPC+09] Scarani V, Bechmann-Pasquinucci H, Cerf N J, Dušek M, Lütkenhaus N and Peev M 2009 The security of practical quantum key distribution *Rev. Mod. Phys.* **81** 1301–50
[Shi17] Shirokov M E 2017 Tight continuity bounds for the quantum conditional mutual information, for the Holevo quantity and for capacities of quantum channels *J. Math. Phys.* **58** 102202
[SP00] Shor P W and Preskill J 2000 Simple proof of security of the BB84 quantum key distribution protocol *Phys. Rev. Lett.* **85** 441–4
[TDS03] Terhal B M, Doherty A C and Schwab D 2003 Symmetric extensions of quantum states and local hidden variable theories *Phys. Rev. Lett.* **90** 157903
[TGW14] Takeoka M, Guha S and Wilde M M 2014 The squashed entanglement of a quantum channel *IEEE Trans. Inf. Theory* **60** 4987–98
[TLR19] Tan E Y-Z, Lim C C-W and Renner R 2019 Advantage distillation for device-independent quantum key distribution arXiv:1903.10535
[TR11] Tomamichel M and Renner R 2011 Uncertainty relation for smooth entropies *Phys. Rev. Lett.* **106** 110506
[Tuc02] Tucci R R 2002 Entanglement of distillation and conditional mutual information arXiv:quant-ph/0202144
[vDGG05] van Dam W, Gill R D and Grunwald P D 2005 The statistical strength of nonlocality proofs *IEEE Trans. Inf. Theory* **51** 2812–35
[VPRK97] Vedral V, Plenio M B, Rippin M A and Knight P L 1997 Quantifying entanglement *Phys. Rev. Lett.* **78** 2275–9
[VV14] Vazirani U and Vidick T 2014 Fully device-independent quantum key distribution *Phys. Rev. Lett.* **113** 140501
[WDH19] Winczewski M, Das T and Horodecki K 2019 Upper bounds on secure key against non-signaling adversary via non-signaling squashed secrecy monotones arXiv:1903.12154
[Wil15] Wilde M M 2015 Recoverability in quantum information theory *Proc. R. Soc.* A **471** 20150338
[Wil16] Wilde M M 2016 Squashed entanglement and approximate private states *Quantum Inf. Process.* **15** 4563–80
[WJD07] Wiseman H M, Jones S J and Doherty A C 2007 Steering, entanglement, nonlocality, and the Einstein–Podolsky–Rosen paradox *Phys. Rev. Lett.* **98** 140402
[WMUK07] Watanabe S, Matsumoto R, Uyematsu T and Kawano Y 2007 Key rate of quantum key distribution with hashed two-way classical communication *Phys. Rev.* A **76** 032312
[YN13] Yang T H and Navascués M 2013 Robust self-testing of unknown quantum systems into any entangled two-qubit states *Phys. Rev.* A **87** 050102