Non-Locality Distillation as Cryptographic Game

Gilles Brassard*†, Benno Salwey*‡, Stefan Wolf‡
*Laboratoire d'informatique théorique et quantique, Université de Montréal
†Canadian Institute for Advanced Research
‡Informatics Faculty, University of Lugano

***Abstract*—Besides being one of the most puzzling aspects of quantum information theory, *non-locality* has been recognised as a valuable resource for various cryptographic protocols. We study the phenomenon of *distillation of non-locality*, which is the ability to generate a stronger instance of non-locality from weaker ones. We construct an eavesdropping third party who gains knowledge about the outputs of distillation protocols. This knowledge directly implies an upper bound on the degree of non-locality of the output of the protocol.**

I. Introduction

Since the introduction of the first quantum key distribution protocol [5], quantum information theory has been a field of intense study. The promises of quantum theory are far reaching and address the very foundations of information theory: the possibility of *randomness expansion* and *amplification*, and of *information-theoretic security* for cryptographic cyphers with no need for a long secret shared key [9], [10], [23]. The essence of these protocols lies in the fundamental inability to predict the outcomes of measurements of quantum systems. In his seminal work [4], Bell showed that the conditional probabilities $\mathsf{P}(ab|xy)$ produced by quantum mechanics when several observables are measured on potentially spatially separated systems — $x \in X$ and $y \in Y$ denote the choices of local observables, $a \in A$ and $b \in B$ the respective measurement outcomes — cannot be described by a so-called *local hidden variable model* (LHV). If a distribution $\mathsf{P}(ab|xy)$ arises from a LHV, its probabilities must satisfy the CHSH [8] inequality

$$\textit{CHSH}(\mathsf{P}) := \frac{1}{4}\sum_{xy} \mathsf{P}(a \oplus b = x \cdot y \mid xy) \leq \frac{3}{4} \ . \qquad (1)$$

Distributions violating (1) have consequently been coined to display *non-locality*, a property that is directly linked to the secrecy of the outputs a and b, which is at least proportional to the violation of (1).

Besides being a valuable resource for cryptographic tasks, non-locality also plays a role in better understanding the foundations of quantum information theory. In the spirit of asking "why is nature as it is?", one may wonder about what is so special concerning the distributions $\mathsf{P}(ab|xy)$ entailed by quantum mechanics. The question has been examined from a *purely information-theoretic* perspective, in particular by Popescu and Rohrlich [19]. They introduced distributions, now called *non-local boxes* and denoted PR_ε in their honour, where ε resembles a noise parameter (which we omit if $\varepsilon = 0$),

$$\mathsf{PR}_\varepsilon(ab|xy) = \begin{cases} \frac{1-\varepsilon}{2} & \text{if } a \oplus b = x \cdot y \\ \varepsilon/2 & \text{otherwise} \ , \end{cases} \qquad (2)$$

which do not enable instantaneous communication, i.e., are *no-signalling*. They exceed the quantum bound whenever $\varepsilon < \varepsilon_q = (2-\sqrt{2})/4 \sim 0.15$ and violate (1) even up to the algebraic maximum of 1 for $\varepsilon = 0$.

What consequences would the existence of post-quantum distributions $\mathsf{P}(ab|xy)$ imply, even when they are restricted to being no-signalling? A first answer was given in [22], where it was shown that if Alice and Bob are given PR distributions as a resource, they can compute any distributed boolean function using a single bit of classical communication, and thus render communication complexity trivial. In [6], this result was extended to the probabilistic setting for any $\varepsilon \lesssim 0.09$. In [15], it was shown that non-locality can be *distilled*; using many identical no-signalling boxes P, Alice and Bob can generate a box $\hat{\mathsf{P}}$ with $\textit{CHSH}(\hat{\mathsf{P}}) > \textit{CHSH}(\mathsf{P})$ without communicating. Then, in [7], a protocol was presented that enables Alice and Bob to use *correlated boxes* P, with $\textit{CHSH}(\mathsf{P}) = 3/4 + \delta$ and δ arbitrarily small, to generate a PR box with arbitrary precision. Correlated boxes are mixtures between a PR box and a box that always outputs perfectly random but correlated bits a and b for any input. Therefore, boxes arbitrarily close to the set of quantum boxes render communication complexity trivial as well.

This stimulated further research on the distillation of non-locality. Other distillation protocols were found [1], [20], which implied that more post-quantum boxes $\mathsf{P}(ab|xy)$ render communication complexity trivial, but indications in [1] and common intuition led to the conjecture that distributions PR_ε *cannot* be used to distill non-locality. In [21], it has been shown that distillation is impossible by protocols using only two PR_ε. By numerical analysis, this no-go result could be extended to protocols using up to nine PR_ε as resource [14]. Using an unbounded number of resources, the impossibility of distillation was shown for slightly restricted *non-adaptive* distillation protocols [17], where the inputs to the resources are chosen independently of other outputs. For general distillation protocols, it was shown that as long as the resources PR_ε are quantum, i.e., $\varepsilon \geq \varepsilon_q$, distillation is strongly limited [11]. For general distillation protocols of *super-quantum* correlations, using n PR_ε as resource, a bound $\textit{CHSH}(\hat{\mathsf{P}}) \leq 1 - \theta(\varepsilon^{-n/2})$ can be derived by considering the so-called Elitzur-Popescu-Rohrlich decomposition [12] of the resource. The idea is to probabilistically decompose the resource into a non-local part and a local part, i.e., a distribution satisfying (1), the weight of the latter being maximal. Consequently, with the same probability weight, $\hat{\mathsf{P}}$ must be local and satisfy (1).

The drawback of this approach is that it cannot yield stronger bounds; the weight of the local part of n PR_ε is exactly of the order of $\theta(\varepsilon^{-n/2})$ [13]. Only recently, a breakthrough was achieved in [3]. Through a rather involved argument, these authors prove the *complete* impossibility of non-locality distillation by general protocols when the resources PR_ε are super-quantum, i.e., $\varepsilon < \varepsilon_q$. A complementary impossibility theorem for the quantum region is still due.

We present a completely new approach to derive general no-go theorems for non-locality distillation: If a box P violates the CHSH inequality (1), its outputs contain some secrecy with respect to any non-signalling adversary, where the secrecy is proportional to the violation. Conversely, a lower bound on the guessing probability of the adversary provides an upper bound on *CHSH*(P). We construct no-signalling adversaries that guess Alice's output a of the box $\hat{\mathsf{P}}(ab|xy)$ generated by the distillation protocol and thereby derive limitations on *CHSH*($\hat{\mathsf{P}}$). The key contribution we provide is to establish a relation between the type of the distillation protocol and the constrains of a no-signalling adversary who attacks the resources used in the protocol. Intuitively, the adversary has to obey no-signalling conditions consistent with the order in which the resources are used in the distillation protocol. So far, our method has not yielded optimal bounds, which are achieved in [3] for super-quantum resources, but the argument is simpler: our sole formal ingredient consists in extending the resource distribution with an additional party. We provide a generalised version of a so-called time-ordered no-signalling adversary considered in [2] and show that this version can be applied to *any* non-locality distillation protocol. This allows us to conclude:

Theorem I.1. *For any non-locality distillation protocol using n PR_ε as resource, the upper bound on the CHSH value is*

$$\mathit{CHSH}(\hat{\mathsf{P}}) \leq 1 - \frac{\varepsilon}{3n} \ . \tag{3}$$

This provides an improvement over the bound implied for general distillation protocols by the Elitzur-Popescu-Rohrlich decomposition [12]. Furthermore, we show that the less restricted no-signalling adversary considered in [16] can be applied to any non-adaptive protocol, which yields:

Theorem I.2. *For any non-adaptive non-locality distillation protocol using an arbitrary number of PR_ε as resource, the upper bound on the resulting CHSH value is*

$$\mathit{CHSH}(\hat{\mathsf{P}}) \leq 1 - \frac{\varepsilon}{4} \ . \tag{4}$$

Theorem I.2 further implies that non-adaptive distillation is virtually impossible for infinitely many values of ε: for any $\delta > 0$ and any $\varepsilon \leq 1/4$, there is a subset of $[\varepsilon/4, \varepsilon]$ of non-zero measure such that *CHSH*($\hat{\mathsf{P}}$) $\leq 1-\varepsilon+\delta$ for non-adaptive protocols using any number of PR_ε as resource.

II. Preliminaries

A. Notation

We refer to a *system* as a black box with an interface consisting of an input and an output; the latter is obtained instantaneously once the former has been inserted. We consider boxes that are shared between two or three parties, which we identify with Alice (A), Bob (B) and Eve (E). If a party holds several subsystems, we shall denote each of them with a subscript, e.g., A_i for the i-th subsystem of Alice. We also use contracted indices and define the shorthand notation $A_{\leq i} := A_1 A_2 \ldots A_i$ for the union of i subsystems. Boxes or systems are identified with conditional probability distributions. For two systems A and B with inputs $x, y \in \mathcal{X} \times \mathcal{Y}$ and outputs $a, b \in \mathcal{A} \times \mathcal{B}$, $\mathsf{P}(ab|xy)$ is the probability of obtaining the outputs (a, b) if the inputs are (x, y). Thus, the whole table of probabilities $\mathsf{P}(ab|xy)$ specifies completely the joint input-output behaviour of the subsystems A and B. When considering a more complicated event, e.g., $f(a) = e$ on the outputs of a system AE, we define $\mathsf{P}(f(a) = e) = \sum_{a,e:f(a)=e} \mathsf{P}(ae)$.

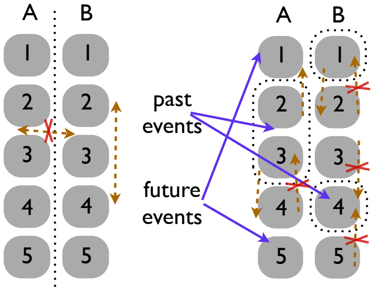

Fig. 1: *Schematic representation of Alice-Bob no-signalling conditions (ABNS) on the left and dynamic time-ordered no-signalling conditions (TONS) on the right. The parties hold each $n = 5$ subsystems and the dotted lines indicate partitions between which no-signalling conditions hold. The crossed arrows indicate no-signalling directions and the uncrossed arrows indicate allowed directions of signalling. On the right hand side, we choose $(i_A, i_B) = (2, 2)$ for an explicit dynamic TONS condition with orders $(j_1, j_2) = (2, 3)$ on Alice's and $(k_1, k_2) = (4, 1)$ on Bob's side.*

B. Several no-signalling conditions

Definition II.1. A system $\mathsf{P}(ab|xy)$ is *no-signalling* if

$$\sum_a \mathsf{P}(ab|xy) = \sum_a \mathsf{P}(ab|x'y) \quad \forall\ b, x, x', y$$
$$\text{and} \quad \sum_b \mathsf{P}(ab|xy) = \sum_b \mathsf{P}(ab|xy') \quad \forall\ a, x, y, y'. \tag{5}$$

Note that (only) when such conditions hold, it is possible to define valid *marginal* boxes so that $\mathsf{P}(a|x)$ and $\mathsf{P}(b|y)$ that are independent of y and x, respectively.

Definition II.2. A system $\mathsf{P}(a_{\leq n} b_{\leq n} | x_{\leq n} y_{\leq n})$ is *Alice-Bob no-signalling (ABNS)* if (5) holds for $A := A_{\leq n}$ and $B := B_{\leq n}$.

In an ABNS system $\mathsf{P}(a_{\leq n} b_{\leq n} | x_{\leq n} y_{\leq n})$, a bit a_i may not only depend on the input x_i but also on other inputs x_j, $j \neq i$. If in a protocol Alice is allowed to access her subsystems *consecutively*, she can choose a later input x_j as a function of the previously obtained output a_i. Then, in order to avoid circular definitions of the variables, a_i should not depend on x_j.

This motivates us to consider a stricter set of conditions than the ABNS conditions, which we call *dynamic time-ordered no-signalling* conditions (dynamic TONS). We allow Alice and Bob to use their n subsystems in *any* order $j_1, j_2, ..., j_n$ and $k_1, k_2, ..., k_n$, respectively. As this order may depend also on outputs obtained during the protocol, we speak of a *dynamic* order. Using again contracted indices $j_{\leq i} := (j_1, ..., j_i)$ and $a_{j_{\leq i}} := (a_{j_1}, a_{j_2}, ..., a_{j_i})$, we uniquely define a dynamic order by the set of functions $\{j_1, j_2(a_{j_1}), ..., j_n(a_{j_{<n}})\} := \{j_i\}$ for Alice and similarly for Bob.

Definition II.3. A system $\mathsf{P}(a_{\leq n}b_{\leq n}|x_{\leq n}y_{\leq n})$ is *dynamic time-ordered no-signalling* with respect to the dynamic orders $\{j_i\}$ and $\{k_i\}$ [1] if

$$\sum_{\substack{a_{j_{>i_A}}\\ b_{k_{>i_B}}}} \mathsf{P}(a_{j_{\leq i_A}} a_{j_{>i_A}} b_{k_{\leq i_B}} b_{k_{>i_B}} | x_{j_{\leq i_A}} x_{j_{>i_A}} y_{k_{\leq i_B}} y_{k_{>i_B}})$$

$$= \sum_{\substack{a_{j_{>i_A}}\\ b_{k_{>i_B}}}} \mathsf{P}(a_{j_{\leq i_A}} a_{j_{>i_A}} b_{k_{\leq i_B}} b_{k_{>i_B}} | x_{j_{\leq i_A}} x'_{j_{>i_A}} y_{k_{\leq i_B}} y'_{k_{>i_B}})$$

$$\forall (a_{j_{\leq i_A}}, b_{k_{\leq i_B}}, x_{j_{\leq i_A}}, y_{k_{\leq i_A}}), (x_{j_{>i_A}}, y_{k_{>i_B}}), (x'_{j_{>i_A}}, y'_{k_{>i_B}})$$
$$\text{for} \quad 0 \leq i_A, i_B \leq n \ . \tag{6}$$

III. Non-locality distillation protocols

We consider bipartite non-locality distillation protocols: Two players Alice and Bob share n resource boxes denoted R. Without communication, their goal is to use these boxes to generate a single box $\hat{\mathsf{P}}(ab|xy)$ such that *CHSH*$(\hat{\mathsf{P}}) >$ *CHSH*(R).[2] To encompass the most general case, we allow the players to use their boxes in any given (dynamic) order, see Figure 2, which also depends on x and y.

Definition III.1. A *bipartite non-locality distillation protocol* using n resource boxes R is defined by the tuple of functions $(\{j_i^x\}, \{k_i^y\}, \{x_{j_i}^x\}, \{y_{k_i}^y\}, f^x, g^y)$ for $1 \leq i \leq n$ in the following way: Given (x, y), the outputs (a, b) of the box $\hat{\mathsf{P}}(ab|xy)$ are functions of the outputs $a_{\leq n}, b_{\leq n}$ of the n resource boxes, i.e., $a = f^x(a_{\leq n})$, $b = g^x(b_{\leq n})$. At the i-th step of the protocol, the function $j_i^x = j_i^x(a_{j_{<i}})$ of the previously obtained outputs $a_{j_{<i}}$ determines which is the next box Alice uses and function $x_{j_i}^x(a_{j_{<i}})$ determines what to input in this box (similarly for Bob, k_i^y and $y_{k_i}^y$).[3]

The functions $\{j_i^x\}$, $\{k_i^y\}$, $\{x_{j_i}^x\}$ and $\{y_{k_i}^y\}$ fix the order of usage of and the inputs to the resources at any given step in the protocol and induce a first mapping; the output functions $f^x(a_{\leq n}), g^y(b_{\leq n})$ induce a second mapping:

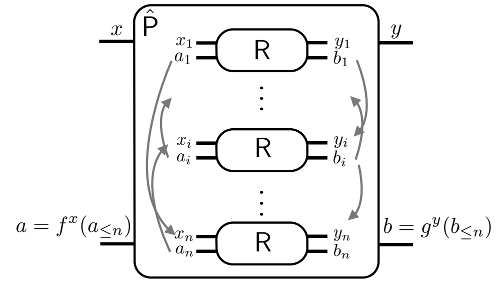

Fig. 2: *Schematic representation of a general distillation protocol. The arrows indicate the orders in which Alice and Bob use their resource boxes. The i-th system of Bob has two outgoing arrows pointing to distinct systems in use, which indicate a dynamic order that depends non-trivially on b_i.*

$$\begin{aligned} &\mathsf{R}^{\otimes n}(a_{\leq n}b_{\leq n}|x_{\leq n}y_{\leq n}) \\ &\downarrow \{j_i^x\}, \{k_i^y\}, \{x_{j_i}^x\}, \{y_{k_i}^y\} \\ \mathsf{P}(a_{\leq n}b_{\leq n}|xy) = &\mathsf{R}^{\otimes n}(a_{\leq n}b_{\leq n}|x_{\leq n}^x(a_{\leq n})y_{\leq n}^y(b_{\leq n})) \\ &\downarrow \{f^x, g^y\} \\ \hat{\mathsf{P}}(ab|xy) = &\sum_{\substack{a_{\leq n}: f^x(a_{\leq n})=a \\ b_{\leq n}: g^y(b_{\leq n})=b}} \mathsf{P}(a_{\leq n}b_{\leq n}|xy) \ , \end{aligned} \tag{7}$$

where we write $x_{\leq n}^x(a_{\leq n})$ for the vector of functions $x_{j_i}^x(a_{j_{<i}})$. A commonly addressed class of distillation protocols are so-called *non-adaptive* distillation protocols, one example being the first distillation protocol presented in [15].

Definition III.2. A distillation protocol is *non-adaptive* if the inputs to the resource are restricted to being independent from any outputs $(a_{\leq n},\ b_{\leq n})$ of the resources.

Note that for non-adaptive protocols there is also no need to specify order functions $\{j_i^x\}$, $\{k_i^y\}$; all inputs can be inserted *simultaneously*.

IV. No-signalling attacks on non-locality distillation protocols

A. The no-signalling adversary Eve

Assume that Alice and Bob hold a box $\mathsf{P}(a_{\leq n}b_{\leq n}|x_{\leq n}y_{\leq n})$ and Alice computes a Boolean function $f(a_{\leq n})$. In order to analyse the privacy of $f(a_{\leq n})$ against a no-signalling adversary, one considers, in analogy to the quantum case, an adversary Eve holding a "no-signalling purifying system" E with input Z. Here we restrict Eve to a *single* input z (which we neglect in the following) and a binary output $e \in \{0, 1\}$, which is independent of Alice's and Bob's inputs.

Definition IV.1. $\mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n})$ is a *no-signalling attack* on $\mathsf{P}'(a_{\leq n}b_{\leq n}|x_{\leq n}y_{\leq n})$ if $\mathsf{P}(e|x_{\leq n}y_{\leq n}) = \mathsf{P}(e)$,

$$\sum_e \mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n}) = \mathsf{P}'(a_{\leq n}b_{\leq n}|x_{\leq n}y_{\leq n})$$

$$\forall a_{\leq n}, b_{\leq n}, x_{\leq n}, y_{\leq n} \ , \tag{8}$$

[1] In [2] only the set on TONS conditions where Alice and Bob use their systems in the same standard order $j_i = i = k_i$ is considered.

[2] Conditional probability distributions that are outputs of non-locality distillation protocols are marked with an accent $\hat{\mathsf{P}}$. The inputs and outputs a_i, b_i, x_i, y_i of the resource boxes R are indexed, in contrast to the inputs and outputs a, b, x, y of the resulting composed protocol.

[3] Note that in the present discussion Alice and Bob do not use shared randomness in addition to their non-local resources. However, the linearity of the CHSH value in the output probabilities of the distillation protocol allows us to extend the results in section V straightforwardly to distillation protocols with shared randomness.

and $\mathsf{P}(a_{\leq n}b_{\leq n}|ex_{\leq n}y_{\leq n})$ satisfies the no-signalling conditions under interest. Here, these are either according to Definitions II.1, II.2, or II.3, defining a standard no-signalling attack, an ABNS-attack or a dynamic TONS-attack, respectively.

B. Non-locality and Eve's guessing probability

Lemma IV.2. *[16] For any no-signalling attack $\mathsf{P}(abe|xy)$ on a box $\mathsf{P}(ab|xy)$ with CHSH$(\mathsf{P}) = 1 - \varepsilon$,*

$$\sum_e \mathsf{P}(e)\max_{a'}[\mathsf{P}(a = a'|ex)] \leq \tfrac{1}{2} + 2\varepsilon \quad \forall x \ . \tag{9}$$

Corollary IV.3. *Let $\hat{\mathsf{P}}(abe|xy)$ be a no-signalling attack on $\hat{\mathsf{P}}(ab|xy)$. If there exists an x such that $\hat{\mathsf{P}}(a = e|x) \geq \frac{1}{2} + 2\varepsilon$, then CHSH$(\hat{\mathsf{P}}) \leq 1 - \varepsilon$.*

The idea is to create an adversary $\mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n})$ to the resource $\mathsf{R}^{\otimes n}(a_{\leq n}b_{\leq n}|x_{\leq n}y_{\leq n})$ who gains knowledge on Alice's output of the distillation protocol $f^{x=0}(a_{\leq n})$. The mappings (7) induce a box $\hat{\mathsf{P}}(abe|xy)$ with $\hat{\mathsf{P}}(e|xy) = \mathsf{P}(e)$ and $\sum_e \hat{\mathsf{P}}(abe|xy) = \hat{\mathsf{P}}(ab|xy)$ for the output $\hat{\mathsf{P}}(ab|xy)$ of the distillation protocol. The crucial part is to see under which conditions $\hat{\mathsf{P}}(ab|exy)$ is no-signalling. Once this is guaranteed, we are able to apply Corollary IV.3 in order to derive limitations on the value of *CHSH*$(\hat{\mathsf{P}})$.

C. Sufficient conditions for a no-signalling attack on $\hat{\mathsf{P}}(ab|xy)$

Theorem IV.4. *Let $\hat{\mathsf{P}}(ab|xy)$ denote a box generated by a general distillation protocol using n R boxes as resource and dynamic orders $\{j_i^x\}$,$\{k_i^y\}$. Let $\mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n})$ be a dynamic TONS attack on $\mathsf{R}^{\otimes n}(a_{\leq n}b_{\leq n}|x_{\leq n}y_{\leq n})$ that fulfils* each *of the four dynamic TONS conditions specified by the orders $(\{j_i^x\},\{k_i^y\})$ for $x \in \{0,1\}$ and $y \in \{0,1\}$. Then $\mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n})$ induces a no-signalling attack $\hat{\mathsf{P}}(abe|xy)$ on distribution $\hat{\mathsf{P}}(ab|xy)$.*

Theorem IV.5. *Let $\hat{\mathsf{P}}(ab|xy)$ denote a box generated by a non-adaptive distillation protocol using n R boxes as resource. Let $\mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n})$ be an ABNS attack on $\mathsf{R}^{\otimes n}(a_{\leq n}b_{\leq n}|x_{\leq n}y_{\leq n})$. Then $\mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n})$ induces a no-signalling attack $\hat{\mathsf{P}}(abe|xy)$ on $\hat{\mathsf{P}}(ab|xy)$.*

The intuition behind the above statements is that as long as the attack $\mathsf{P}(a_{\leq n}b_{\leq n}|ex_{\leq n}y_{\leq n})$ on the resources respects the orders of use of the distillation protocol, it also induces a no-signalling attack on the output of the protocol. In general protocols, Alice can choose the inputs $x_{j_{>i}}$ as a function of $a_{j_{\leq i}}$, which is why we require the outputs $a_{j_{\leq i}}$ to be independent of the inputs $x_{j_{>i}}$ by enforcing the dynamic TONS conditions on $\mathsf{P}(a_{\leq n}b_{\leq n}|ex_{\leq n}y_{\leq n})$; similarly for Bob. For non-adaptive protocols, the inputs $x_{\leq n}$ are independent of the outputs $a_{\leq n}$ and therefore it is sufficient to enforce the ABNS conditions. To prove Theorems IV.4 and IV.5, one needs to show that $\hat{\mathsf{P}}(ab|exy)$ induced by $\mathsf{P}(a_{\leq n}b_{\leq n}|ex_{\leq n}y_{\leq n})$ is no-signalling. This is done using the Definitions II.2 and II.3 in the respective case.

V. No-signalling attacks and their consequences

A. A dynamic TONS-attack on $\mathsf{PR}_\varepsilon^{\otimes n}$

We construct a dynamic TONS attack on $\mathsf{PR}_\varepsilon^{\otimes n}$, which is a generalisation of the attack constructed in [2]. Fix a function $f(a_{\leq n})$. Let $C = \{c_1, c_2, ..., c_k\}$ be a binary prefix code with $|c_m| \leq n-1$ where we assign to each codeword c_m also a bit $a^*(m)$. The code C and the bits $a^*(1), ..., a^*(k)$ are chosen according to the function $f(a_{\leq n})$. We refer the reader to [2] for details. Here is the construction of $\mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n}) = \mathsf{P}(e)\mathsf{P}(a_{\leq n}b_{\leq n}|ex_{\leq n}y_{\leq n})$ for a given dynamic order $\{j_i^x\}$, where we omit the superscript x for better readability and set $\mathsf{P}(e) = 1/2$: [4]

$$\mathsf{P}(a_{\leq n}b_{\leq n}|ex_{\leq n}y_{\leq n}) = \prod_{i=1}^{n} \mathsf{P}(a_{j_i}b_{j_i}|a_{j_{<i}}b_{j_{<i}}ex_{j_{\leq i}}y_{j_{\leq i}}) \tag{10}$$

where $\mathsf{P}(a_{j_i}b_{j_i}|a_{j_{<i}}b_{j_{<i}}ex_{j_{\leq i}}y_{j_{\leq i}})$ is defined as

$$\begin{cases} \mathsf{PR}_\varepsilon(a_{j_i}b_{j_i}|x_{j_i}y_{j_i}) \ \text{ if } a_{j_{<i}} \notin C, \text{ or} & \text{(11a)} \\ (1-2\varepsilon)\cdot\mathsf{PR}(a_{j_i}b_{j_i}|x_{j_i}y_{j_i}) + 2\varepsilon\cdot\delta(a^*(m)\oplus e, a_{j_i}) & \\ \qquad\qquad\qquad \text{otherwise, if } a_{j_{<i}} = c_m \in C & \text{(11b)} \end{cases}$$

and $\delta(a, a') = 1$ if $a = a'$ and $\delta(a, a') = 0$ otherwise. Note that through (11b), the construction of $\mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n})$ implicitly depends on x if the sets of functions $\{j_i^{x=0}\}$ and $\{j_i^{x=1}\}$ differ. However, one can define the box $\mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n})$ *independently* of (x, y) with *exactly* the same conditional distribution for any protocol where the players use their systems consecutively. At any time in the protocol one can also regard the information j_i^x, namely which box to use in the next step, as an *additional input* of Alice to the system $\mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n})$; instead of Alice inserting input $x_{j_i}^x(a_{j_{<i}})$ into the slot $j_i^x(a_{j_{<i}})$, one obtains the same distribution if Alice inserts the pair $(j_i^x(a_{j_{<i}}), x_{j_i}^x(a_{j_{<i}}))$ into a box $\mathsf{P}(a_{\leq n}b_{\leq n}e|j_{\leq n}x_{\leq n}y_{\leq n})$ with a *single* input slot on Alice's side, which is *repeatedly* used.

Theorem V.1. *The construction* (10) - (11b) *defines a dynamic TONS attack on $\mathsf{PR}_\varepsilon^{\otimes n}(a_{\leq n}b_{\leq n}|x_{\leq n}y_{\leq n})$ for the dynamic orders $\{j_i^x\}, \{k_i^y\}$. Furthermore,*

$$\mathsf{P}(f(a_{\leq n}) = e \mid x_{\leq n}) \geq \frac{1}{2} + \frac{2\varepsilon}{3n} \quad \forall x_{\leq n} \ . \tag{12}$$

To prove Theorem V.1, one needs to show that $\mathsf{P}(a_{\leq n}b_{\leq n}e|x_{\leq n}y_{\leq n})$ satisfies three properties:

1) It is an extension of $\mathsf{PR}_\varepsilon^{\otimes n}$: This follows from the fact that the j_i-th box is either directly a PR_ε box by (11a), or by (11b) in the average over e.
2) It satisfies (6) for all four orderings $\{j_i^x\}, \{k_i^y\}$: using (10) - (11b) it is straightforward to show that the marginal of $A_{j^x_{\leq i_A}} B_{k^y_{\leq i_B}}$ is independent of $(x^x_{j_{>i_A}}, y^y_{k_{>i_B}})$.
3) (12) holds: Intuitively, for any balanced function $f(a_{\leq n})$, there must exist at least one bit a_i that influences $f(a_{\leq n})$

[4] We present here the construction for functions $f(a_{\leq n})$ that are (almost) unbiased. For biased functions, one can apply Corollary IV.3 directly.

by at least a value of roughly $1/n$. This bit is then biased towards the preferred direction by (11b). For more details we refer the reader to [2].

A direct consequence of Corollary IV.3, Theorem IV.4 and Theorem V.1 is Corollary V.2.

Corollary V.2. *Let $\hat{\mathsf{P}}(ab|xy)$ be generated by a* general *distillation protocol using n PR_ε boxes as resource. Then the CHSH value of the generated box must satisfy*

$$\mathit{CHSH}(\hat{\mathsf{P}}) \leq 1 - \frac{\varepsilon}{3n} \ . \tag{13}$$

B. An ABNS-attack on $\mathsf{PR}_\varepsilon^{\otimes n}$

Theorem V.3. *[16] Let $\mathsf{R} = \mathsf{PR}_\varepsilon$. For any $0 \leq \varepsilon \leq 1/4$ and any function $f : \{0,1\}^n \to \{0,1\}$, there exists an ABNS attack $\mathsf{P}(a_{\leq n} b_{\leq n} e | x_{\leq n} y_{\leq n})$ on $\mathsf{R}^{\otimes n}$ such that*

$$\mathsf{P}(f(a_{\leq n}) = e \mid x_{\leq n}) \geq \frac{1}{2} + \frac{\varepsilon}{2} \quad \forall x_{\leq n} \ . \tag{14}$$

Corollary V.4 (below) is a direct consequence of Corollary IV.3, Theorem IV.5 and Theorem V.3.

Corollary V.4. *Let $\hat{\mathsf{P}}(ab|xy)$ be generated by a* non-adaptive *distillation protocol using n PR_ε boxes as resource, where n is arbitrary. Then the CHSH value of the generated box must satisfy*

$$\mathit{CHSH}(\hat{\mathsf{P}}) \leq 1 - \frac{\varepsilon}{4} \ . \tag{15}$$

Note that the proof of impossibility of distillation for non-adaptive protocols in [17] is restricted to protocols with $f^{x=0} = f^{x=1}$ and $g^{y=0} = g^{y=1}$. As (15) holds for any number n of resources, one can derive bounds even stronger than (15) for most values of ε by an argument combining sub-protocols. For a given ε, define $1 - \varepsilon' \leq 1 - \varepsilon/4$ as the supremum for $\mathit{CHSH}(\hat{\mathsf{P}})$ for all non-adaptive distillation protocols using PR_ε resources. If $\varepsilon' < \varepsilon$, then via use of a depolarisation protocol [18], and potentially some noise, it is possible to generate any $\mathsf{PR}_{\varepsilon''}$ box with $\varepsilon' < \varepsilon'' < \varepsilon$ with PR_ε boxes. Thus, for all non-adaptive distillation protocols using $\mathsf{PR}_{\varepsilon''}$ as resource, $1 - \varepsilon'$ is also the supremum of $\mathit{CHSH}(\hat{\mathsf{P}})$.

VI. Conclusion and Outlook

We presented a novel method for deriving bounds on distillation protocols. So far, our method does not yield optimal bounds on distillation, as is achieved in [3] for super-quantum correlations. However, the argument presented here does not require elaborate mathematical tools such as those deployed in [3]. The method consists in the construction of no-signalling attacks on output bits of distillation protocols. We established sufficient conditions for no-signalling attacks on non-adaptive and on general distillation protocols. A suitable generalisation of the attack in [2] to a broader set of no-signalling conditions allowed us to argue that the reduction of noise in a PR_ε box can at best be proportional to the number of resources used in a general distillation protocol. We also found that the noise level of a PR_ε box can at best be reduced by a factor of 4 through non-adaptive distillation in general and virtually not at all for some values of ε.

Acknowledgements. The authors are grateful to Rotem Arnon-Friedman, Ämin Baumeler, Omar Fawzi and Jibran Rashid for fruitful discussions. The work of GB is supported in part by the Canadian Institute for Advanced Research, the Canada Research Chair program, Canada's NSERC, Québec's INTRIQ, and the Institute for Theoretical Studies at ETH Zürich. The work of SW is supported by the Swiss National Science Foundation, the NCCR QSIT, and the COST action on "Fundamental Problems in Quantum Physics". Throughout this work, BS has been supported by several of the above.

References

[1] Jonathan Allcock, Nicolas Brunner, Noah Linden, Sandu Popescu, Paul Skrzypczyk, and Tamás Vértesi. Closed sets of nonlocal correlations. *Phys. Rev. A*, 80:062107, Dec 2009.

[2] Rotem Arnon-Friedman and Amnon Ta-Shma. Limits of privacy amplification against nonsignaling memory attacks. *Phys. Rev. A*, 86:062333, Dec 2012.

[3] Salman Beigi and Amin Gohari. A monotone measure for non-local correlations. *http://arxiv.org/abs/1409.3665v3*, Nov 2014.

[4] John S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, Nov 1964.

[5] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of the International Conference on Computers, Systems and Signal Processing*, Bangalore, pages 175–179, Dec 1984.

[6] Gilles Brassard, Harry Buhrman, Noah Linden, André Allan Méthot, Alain Tapp, and Falk Unger. Limit on nonlocality in any world in which communication complexity is not trivial. *Phys. Rev. Lett.*, 96:250401, Jun 2006.

[7] Nicolas Brunner and Paul Skrzypczyk. Nonlocality distillation and postquantum theories with trivial communication complexity. *Phys. Rev. Lett.*, 102:160403, Apr 2009.

[8] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, Oct 1969.

[9] Roger Colbeck. *Quantum and Relativistic Protocols for Secure Multi-Party Computation*. PhD thesis, University of Cambridge, 2006.

[10] Roger Colbeck and Renato Renner. Free randomness can be amplified. *Nat. Phys.*, 8(6):450–453, Jun 2012.

[11] Dejan Dukaric and Stefan Wolf. A limit on non-locality distillation. *http://arxiv.org/abs/0808.3317*, Aug 2008.

[12] Avshalom C. Elitzur, Sandu Popescu, and Daniel Rohrlich. Quantum nonlocality for each pair in an ensemble. *Phys. Lett. A*, 162(1):25–28, Jan 1992.

[13] Matthias Fitzi, Esther Hänggi, Valerio Scarani, and Stefan Wolf. The non-locality of n noisy Popescu-Rohrlich boxes. *J. Phys. A: Math. Theor.*, 43(46):465305, Oct 2010.

[14] Manuel Forster. Bounds for nonlocality distillation protocols. *Phys. Rev. A*, 83:062114, Jun 2011.

[15] Manuel Forster, Severin Winkler, and Stefan Wolf. Distilling nonlocality. *Phys. Rev. Lett.*, 102:120401, Mar 2009.

[16] Esther Hänggi, Renato Renner, and Stefan Wolf. The impossibility of non-signaling privacy amplification. *Theor. Comput. Sci.*, 486:27–42, May 2013.

[17] Peter Høyer and Jibran Rashid. Optimal protocols for nonlocality distillation. *Phys. Rev. A*, 82:042118, Oct 2010.

[18] Lluis Masanes, Antonio Acín, and Nicolas Gisin. General properties of nonsignaling theories. *Phys. Rev. A*, 73:012112, Jan 2006.

[19] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Found. Phys.*, 24:379–385, Mar 1994.

[20] Jibran Rashid. *Limits and Consequences of Nonlocality Distillation*. PhD thesis, University of Calgary, Apr 2012.

[21] Anthony J. Short. No deterministic purification for two copies of a noisy entangled state. *Phys. Rev. Lett.*, 102:180502, May 2009.

[22] Wim van Dam. Implausible consequences of superstrong nonlocality. *arXiv:quant-ph/0501159*, Jan 2005.

[23] Umesh Vazirani and Thomas Vidick. Fully device-independent quantum key distribution. *Phys. Rev. Lett.*, 113:140501, Sep 2014.