Robust Device-Independent Randomness Amplification with Few Devices

Fernando G.S.L. Brandão,1 Ravishankar Ramanathan,2,3 Andrzej Grudka,4 Karol Horodecki,2,5 Michał Horodecki,2,3 and Paweł Horodecki2,6
1Department of Computer Science, University College London
2National Quantum Information Center of Gdańsk, 81-824 Sopot, Poland
3Institute of Theoretical Physics and Astrophysics, University of Gdańsk, 80-952 Gdańsk, Poland
4Faculty of Physics, Adam Mickiewicz University, 61-614 Poznań, Poland
5Institute of Informatics, University of Gdańsk, 80-952 Gdańsk, Poland
6Faculty of Applied Physics and Mathematics, Technical University of Gdańsk, 80-233 Gdańsk, Poland
(Dated: October 18, 2013)

Randomness amplification is the task of transforming a source of somewhat random bits into a source of fully random bits. Although it is impossible to amplify randomness from a single source by classical means, the situation is different considering non-local correlations allowed by quantum mechanics. Here we give the first device-independent protocol for randomness amplification using a constant number of devices. The protocol involves eight devices, can amplify any non-deterministic source into a fully random source, tolerates a constant rate of error, and has its correctness based solely on the assumption of non-signalling between the devices. In contrast all previous protocols either required an infinite number of devices, or could only amplify sources sufficiently close to fully random.

I. INTRODUCTION

Randomness is a useful resource in a variety of applications, ranging from numerical simulations to cryptography. However almost always one needs ideal, or close to ideal, randomness, while typical random sources are far from ideal. Randomness amplification, defined as a process that maps a source of randomness into another closer to an ideal source, is a potential solution to this problem. But is randomness amplification possible at all?

We model a source of randomness as an $\varepsilon$-Santha-Vazirani source ($\varepsilon$-SV source), given by a probability distribution $p(x_1, \ldots, x_n)$ over bit strings such that for every $i \leq n$,

$$\frac{1}{2} - \varepsilon \leq p(x_i|x_1, \ldots, x_{i-1}) \leq \frac{1}{2} + \varepsilon.$$  \hspace{1cm} (1)

For example, when $\varepsilon = 0$ the source is fully random, while when $\varepsilon = 1/2$ it may be fully deterministic. The previous equation is the only assumption on the source, which otherwise is completely unknown. Given an SV source, is it possible to process the bits so that the quality of the randomness is improved? In particular, can we obtain a fully random bit by processing an arbitrary large number of bits from a SV source? In [1], Santha and Vazirani answered the question in the negative: It is impossible to improve the quality of the randomness of SV sources \footnote{Similar results are also known for other kinds of sources of randomness [2].}. However their argument only applies to classical protocols and it leaves open the possibility that the situation might be different once quantum resources are considered.

Indeed it is trivial to generate randomness in quantum mechanics, e.g. by preparing a state and measuring it in a complementary basis. However this assumes that one has full control of the state preparation and the measurement. A more demanding task is to generate randomness in a device-independent manner, treating the quantum system as a black-box and obtaining randomness
only as a consequence of the correlations in measurement outcomes and fundamental physical principles such as the no-signaling principle. The existence of non-local quantum correlations violating Bell inequalities already suggests that device-independent randomness amplification could be achieved. However, the violation of Bell inequalities requires that the measurements be performed in a random manner, independent of the system upon which they are performed [3, 4].

In a seminal work [5], Colbeck and Renner showed that despite this difficulty non-local quantum correlations can be used to amplify the randomness of Santha-Vazirani sources sufficiently close to fully random. This result was later improved by Gallego et al. [7], who gave a protocol using quantum non-local correlations to amplify general SV sources, as long as they are not deterministic. Neither of the two protocols tolerate noise however. In [8], we gave a different protocol that is robust to noise and can transform any non-deterministic SV source into a fully random one. A major drawback of the protocols in [7, 8] is that they require an infinite number of space-like separated devices. Therefore a natural open question is the existence of a randomness amplification protocol using a fixed number of devices, on the one hand, and allowing for the amplification of arbitrary non-deterministic sources, on the other hand. See also [10–15] for more recent work in the area.

A related, but distinct, task to randomness amplification is (device-independent) randomness expansion, where one assumes that an input seed of perfect random bits is available and the goal is to expand it into a larger random bit string. Quantum non-locality has found application also in this latter task [16, 19–23, 27, 29] (as well as in device-independent quantum key-distribution (see e.g. [24–26, 28]).

### A. Results

In this paper we overcome the shortcomings of previous protocols and obtain:

**Theorem 1.** For every $\varepsilon > 0$, there is a protocol using an $\varepsilon$-SV source and eight non-signalling devices with the following properties:

- Using the devices $\text{poly}(n, 1/\delta)$ times, the protocol either aborts or produces $n$ bits which are $\delta$-close to uniform and independent of any side information (e.g. held by an adversary).

- Local measurements on many copies of a four-partite entangled state, with $\text{poly}(1 - 2\varepsilon)$ error rate, give rise to devices that do not abort the protocol with probability larger than $1 - 2^{-\Omega(n)}$.

See Proposition 7 in Section III for a precise formulation of the theorem.

### B. Protocol and Outline of its Correctness Proof

The protocol we use for randomness amplification is in fact simpler than previous protocols [5, 7, 8]. It is given precisely in Fig. 2 and illustrated in Fig. 1 but its rough structure is the following: First one uses several bits from the Santha-Vazirani source in order to choose inputs for eight non-signalling boxes (each of which is reused many times). Then one collects the outputs of the boxes and using the empirical data on inputs and outputs decides whether to abort. Then if the protocol is not aborted, one applies a randomness extractor (see Section II A) to all output bits of the eight devices. The output bits of the protocol, which by Theorem 1 are close to fully random, are just the output bits of the extractor.

The proof of correctness of the protocol consists of three main steps.
(i) We show that we can effectively work with two independent devices, each consisting of four boxes; by independent we mean devices that given any fixed inputs produce uncorrelated outputs.

(ii) We show that conditioned on passing a certain test on the inputs and outputs of the boxes, for a single device and a given input, with high probability, the outputs are a source of linear min-entropy.

(iii) We use existing results on extracting randomness from uncorrelated min-entropy sources.

Step (i) is presented in Section IV D. The idea is to adapt recent de Finetti theorems for non-signalling devices [30, 31] (based on information-theoretical methods) to the situation in which subsystems are selected from a Santha-Vazirani source, instead of being selected uniformly at random. In Lemma 15 we show that given two non-signalling boxes, if we select at random (using a SV source) a block of boxes from the second device among a sufficiently large number of blocks, this block of boxes will be approximately uncorrelated with the first device.

Step (ii) is established by a sequence of implications (which have a similar flavour to the estimation of [19–21] for the related task of randomness expansion) as follows:

• In Lemma 8 we present an estimation procedure which ensures that with high probability the value of the Bell expression with settings chosen from an SV source is small \(^2\) for a linear fraction of boxes conditioned on previous inputs and outputs. This will follow from an application of Azuma’s inequality.

• In Lemma 11 we show that by the very definition of SV sources, the value of the Bell expression with uniformly random settings is also small.

• In Lemma 10 we show that a small uniform Bell value implies that for any setting appearing in the Bell expression, the probability of any output is bounded away from one. This is achieved by linear-programming, analogously to the approach of [7, 8].

• In Lemma 14, in turn, we show that if a linear fraction of conditional boxes have probability of outputs bounded away from one (which is ensured by Lemmas 8, 11 and 10), then given the input, the output has linear min-entropy with high probability.

Finally, Step (iii) is an application of known results regarding extracting randomness from two independent min-entropy source (see Section II A).

It is instructive to point out already here where the non-local nature of our protocol plays a role. Clearly the first step is entirely classical, and it does not use any feature of non-local correlations. It only exploits the fact that the devices do not signal to each other and are time-ordered (the remarkable fact here is that the de Finetti bound of Lemma 15 does apply also to general non-signaling correlations). This step ensures that given a particular input, the outputs of the selected boxes in each device are independent of each other. Note this is the case both for non-local as well as for classical boxes. It is also clear that the third step is independent of any non-local feature of the devices.

It is in the second step that the non-local nature of correlations are exploited. There looking at the inputs and outputs of the devices obtained, one can verify in a device-independent manner that the outputs must have been somewhat random. This is impossible classically without making further assumptions.

\(^2\) The Bell inequality we consider is such that the maximum possible violation corresponds to the zero value. Thus the smaller the Bell value, the larger the violation.
C. Extensions

Theorem 1 has the drawback that the extractor used in Protocol 2 is non-explicit (we only know it exists by the probabilistic method). We can improve on this aspect if we are willing to increase the number of non-signalling devices.

**Theorem 2.** For every $\varepsilon > 0$, there is a protocol using an $\varepsilon$-SV source and twelve non-signalling devices with the following properties:

- Using the devices $\text{poly}(n, 1/\delta)$ times, the protocol either aborts or produces $n$ bits which are $\delta$-close to uniform and independent of any side information (e.g. held by an adversary).
- Local measurements on many copies of a four-partite entangled state, with $\text{poly}(1 - 2\varepsilon)$ error rate, give rise to devices that do not abort the protocol with probability larger than $1 - 2^{-\Omega(n)}$.

Moreover the protocol is fully explicit.

The protocol is given in Fig. 3 (for the theorem above we use it with $k = 3$ and the four-partite Bell inequality given in Section II.C). Its proof of correctness is completely analogous to the proof of Theorem 1, the only differences being that we use a version of the de Finetti bound for more devices (given in Lemma 16) and the explicit extractor from part (ii) of Lemma 5.

II. PRELIMINARIES

A. Randomness Extractors

In the protocol, we use randomness extractors to obtain nearly uniform bits from a finite number of randomness sources. We use the min-entropy to measure the amount of randomness in a source $S$.

**Definition 3.** The min-entropy of a random variable $S$ is given by

$$H_{\min}(S) = \min_{s \in \text{supp}(S)} \log \frac{1}{\Pr(S = s)}. \quad (2)$$

For $S \in \{0, 1\}^n$, $S$ is called an $(n, H_{\min}(S))$ source with entropy rate $\frac{1}{n} H_{\min}(S)$.

From multiple weak sources of randomness, we use independent source extractors [6] to extract nearly uniform random bits.

**Definition 4 (Independent-source extractor).** An independent source extractor is a function $\text{Ext} : (\{0, 1\}^n)^k \to \{0, 1\}^m$ that acting on $k$ independent $(n, H_{\min}(S))$ sources, outputs $m$ bits with error $\xi$, i.e. for $k$ independent $(n, H_{\min}(S))$ sources $S_1, \ldots, S_k$ we have

$$|\text{Ext}(S_1, \ldots, S_k) - U_m| \leq \xi, \quad (3)$$

where $|.|$ is the distance between the two distributions and $U_m$ denotes the uniform distribution on the $m$ bits.

The results about extractors that we will use are summarized in the following lemma

**Lemma 5.** [Extractors Constructions]
(i) [2] There exists a (non-explicit) deterministic extractor that given two independent sources of min-entropy larger than $h$, outputs $\Omega(h)$ bits $2^{-\Omega(h)}$-close to uniform.

(ii) [18] There exists an explicit extractor that given three independent sources, one having min-entropy larger than $\tau n$ (for any $\tau > 0$) and the other two larger than $h \geq \text{polylog}(n)$, outputs $\Omega(h)$ bits $2^{-h\Omega(1)}$-close to uniform.

Theorem 1 uses the non-explicit extractor for two sources from part (i), while Theorem 2 uses Rao’s extractor [18] from part (ii).

B. Assumptions on the Devices

The protocol presented in this paper can be applied to any Bell inequality involving $t$ parties with the following properties:

- The maximal no-signaling value of the Bell expression is attainable by quantum correlations.
- For every measurement setting $u = (u^1, \ldots, u^t)$, the probability of any output $x = (x^1, \ldots, x^t)$ is bounded away from one for all no-signaling boxes that achieve maximum violation of the Bell inequality.

As an example of a Bell inequality with the above properties, we present a four-partite inequality in the next section. The assumptions needed for the protocol are now explained in this four partite setting, but can be straightforwardly generalized to a Bell expression involving any other number of parties as well.

(ii) Non-signaling Structure: In the protocol, we consider 4$k$ parties performing $k$ Bell experiments in parallel, for some finite $k$ (the simplest case for which the protocol works is $k = 2$). The parties act on $k$ devices each of which consists of four no-signaling apparatuses. Moreover, the $k$ devices are also no-signaling between each other. Each device may be thought of as a channel that transforms inputs (settings) $u$ into outputs $x$. An external party, for example an eavesdropper (Eve), acts on a system that is no-signaling with respect to the $k$ devices. On the $j$-th device for $j \in [k]$, the parties conduct $M_j$ runs of the Bell experiment sequentially with a time-ordered no-signaling structure, i.e. the outputs of the $l$-th run may depend on the inputs and outputs of the previous runs.

We describe the correlations in the $k + 1$ devices by the following joint probability distribution:

$$P(x^1_{\leq M_1}, \ldots, x^k_{\leq M_k}, Z| \mathbf{u}^1_{\leq M_1}, \ldots, \mathbf{u}^k_{\leq M_k}, \mathbb{W}),$$

where e.g. $x^j_{\leq M_1}$ denotes the vector $(x^j_1, \ldots, x^j_{M_1})$. For device $j$ with $1 \leq j \leq k$, the inputs and outputs of the $l_j$-th box (associated to the $l_j$-th use of the device) are given by $u^j_{l_j}$ and $x^j_{l_j}$, respectively. Likewise, the input and output of the last device (held by the eavesdropper) are given by $\mathbb{W}$ and $Z$. One may think of the adversary holding the set of random variables $Z, \mathbb{W}$ and supplying $k$ devices which produce the conditional probability distribution $P(\cdot | Z, \mathbb{W})$ whose behaviour depends on $Z$ and $\mathbb{W}$.

We say such $P$ is a $(k; M_1, \ldots, M_k)$ time-ordered non-signaling box, meaning that it is non-signaling between each of the $k$ devices and time-ordered non-signaling within each device.

(iii) Conditional Boxes: The settings are drawn from SV-source described by a distribution $\nu$. For each device $j \in [k]$, this defines a probability space given by sequences of inputs and outputs $(u^j_1, x^j_1, \ldots, u^j_{M_j}, x^j_{M_j})$, with measure $P(u^j_1, x^j_1, \ldots, u^j_{M_j}, x^j_{M_j}) =$
random variables through the no-signaling adversary Eve. In other words, we assume that the SV source, Eve’s inputs. Classically this cannot be achieved, since there is no way of certifying that a channel is not deterministic. However the users of the device do not know this a priori. In the protocol we propose, they make the assumption that the \( \epsilon \)-SV source and the boxes can be correlated with each other only through the no-signaling adversary Eve. In other words, we assume that the SV source, Eve’s random variables \((Z, W)\) and the box \(P\) constitute a Markov chain, so that given \((Z = z, W = w)\), the box \(P(x_{i}^{1}, \ldots, x_{M_{k}}^{k} | u_{j_{1}}^{1}, \ldots, u_{j_{M_{k}}}^{k}) = \) independent of the bits produced by the source.

The device is formally a channel applied to SV source. Its analogue in the Santha-Vazirani theorem is just a deterministic hash function. In our situation, the channel is not deterministic. However the users of the device do not know this a priori. In the protocol we propose, they are able to verify that the device has randomness by looking at the statistics of inputs and outputs. Classically this cannot be achieved, since there is no way of certifying that a channel is not deterministic without extra assumptions.

C. The Bell inequality

The inequality we consider involves four spatially separated parties with measurement settings \( \mathbf{u} = \{u^{1}, u^{2}, u^{3}, u^{4}\} \) and respective outcomes \( \mathbf{x} = \{x^{1}, x^{2}, x^{3}, x^{4}\} \). Each party chooses one of two measurement settings with two outcomes each so that \( u^{i} \in \{0, 1\} \) and \( x^{i} \in \{0, 1\} \) for \( i \in \{1, \ldots, 4\} \). The measurement settings for which non-trivial constraints are imposed by the inequality can be divided into two sets

\[
U_{0} = \{\{0001\}, \{0010\}, \{0100\}, \{1000\}\} \quad \text{and} \quad U_{1} = \{\{0111\}, \{1011\}, \{1101\}, \{1110\}\}. \tag{5}
\]

The inequality is then \[ \sum_{x, u} (I_{ UT_{-1} x^{i}=0} I_{ u \in U_{0}} + I_{ UT_{-1} x^{i}=1} I_{ u \in U_{1}}) P(x | u) \geq 2, \tag{6} \]

where the indicator function \( I_{L} = 1 \) if \( L \) is true and 0 otherwise. The local hidden variable bound is 2 and there exist no-signaling distributions that reach the algebraic limit of 0. For any no-signaling box represented by a vector of probabilities \( \{P(x | u)\} \), the Bell inequality may be written as

\[
\mathbf{B} \cdot \{P(x | u)\} = \sum_{x, u} B(x, u) P(x | u) \geq 2, \tag{7}
\]

where \( \mathbf{B} \) is an indicator vector for the Bell inequality with \( 2^{4} \times 2^{4} \) entries

\[
B(x, u) = I_{ UT_{-1} x^{i}=0} I_{ u \in U_{0}} + I_{ UT_{-1} x^{i}=1} I_{ u \in U_{1}}. \tag{8}
\]

Consider the quantum state

\[
| \Psi \rangle = \frac{1}{\sqrt{2}} (| \phi_{-} \rangle | \phi_{+} \rangle + | \psi_{+} \rangle | \psi_{-} \rangle), \tag{9}
\]

where \( | \phi_{-} \rangle = \frac{1}{\sqrt{2}} (|0\rangle |0\rangle - |1\rangle |1\rangle), \quad | \psi_{+} \rangle = \frac{1}{\sqrt{2}} (|0\rangle |1\rangle + |1\rangle |0\rangle), \quad | \phi_{+} \rangle = \frac{1}{\sqrt{2}} (|0\rangle |+\rangle + |1\rangle |-\rangle), \) and \( | \psi_{-} \rangle = \frac{1}{\sqrt{2}} (|0\rangle |-\rangle - |1\rangle |+\rangle). \)

Measurements in the \( X \) basis

\[
\{|+\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle), \quad |-\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle)\}. \tag{10}
\]
correspond to $u^i = 0$ and measurements in the $Z$ basis
\[ \{0, 1\} \] (11)
correspond to $u^i = 1$ for each of the four parties $i \in \{1, \ldots, 4\}$. These measurements on $|\Psi\rangle$ lead to the algebraic violation of the inequality, i.e., the sum of the probabilities appearing in the inequality is zero.

The reason for the choice of this Bell inequality is twofold. Firstly, as we have seen, there exist quantum correlations achieving the maximal no-signaling violation of the inequality, which implies that free randomness amplification starting from any initial $\epsilon$ of the SV source may be possible. Secondly, we will show (in Lemma 10) that for any measurement setting $u$ out of the $2^4$ possible settings in the inequality, the probability of any of the $2^4$ output bit strings $x$ is bounded away from one (for any no-signaling box) by a linear function of the uniform value of the Bell expression.

D. Randomness Criterion

To quantify the quality of the output we will use of the distance to uniform of a random variable $S \in \Sigma$, conditioned on Eve’s input and output:
\[ d(S|Z, W) := \max_{s, w, z} |P(S = s|Z = z, W = w) - 1/|\Sigma||. \] (12)

Although this function is convenient to work, it is not universally composable. A better definition of randomness relative to an eavesdropper is the following (see e.g. [26]):
\[ d_c(S|Z, W) := \sum_{s=0}^{||\Sigma||-1} \max_w \sum_z P(Z = z|W = w)P(S = s|Z = z, W = w) - 1/|\Sigma||, \] (13)
with $|\Sigma|$ the size of $\Sigma$.

However the following relation between them holds:

**Lemma 6.** For a random variable $S \in \Sigma$,
\[ d_c(S|Z, W) \leq |\Sigma|d(S|Z, W). \] (14)

**Proof.**
\[
\begin{align*}
\sum_{s=0}^{||\Sigma||-1} \max_w \sum_z P(Z = z|W = w)P(S = s|Z = z, W = w) - 1/|\Sigma||
\leq |\Sigma| \max_{s,w} \sum_z P(Z = z|W = w)P(S = s|Z = z, W = w) - 1/|\Sigma||
\leq |\Sigma| \max_{s,w,z} P(S = s|Z = z, W = w) - 1/|\Sigma||
= |\Sigma|d(S|Z, W).
\end{align*}
\] (15)
FIG. 1: Illustration of the protocol for randomness amplification from two no-signaling devices with $N_1 = 1$ block of $n$ runs from the first device and $N_2$ block of $n$ runs from the second.

III. PROOF OF CORRECTNESS OF THE PROTOCOL

Let us consider the protocol in Fig. 2 performed on $k = 2$ devices with the following choice of number of blocks:

$$N_1 = 1, \quad N_2^{1-\log(1+2\varepsilon)} = 8n \ln (2) t^4. \quad (16)$$

for a parameter $t > 0$. Then our main result is the following

**Proposition 7.** Let $(N_1, N_2)$ be given by Eq. (16). Then conditioned on not aborting, the protocol (when applied to the four-partite Bell inequality given in Section II C) outputs $\Omega(n)$ bits $S_n$ such that

$$d_c(S_n|Z, W) \leq 4 \exp \left( -\Omega \left( n\delta^2(1 - \mu)^2(\frac{1}{2} - \varepsilon)^8 \right) \right) + \frac{6}{\sqrt{t}}, \quad (17)$$

Moreover if the boxes are realized by performing measurements which are $O(\delta(1 - \mu)(1 - 2\varepsilon)^4)$-close to either one of the measurements of Eqs. (10, 11) on states which are $O(\delta(1 - \mu)(1 - 2\varepsilon)^4)$-close to the state of Eq. (9), then the protocol accepts with high probability.
Protocol I

1. The $\varepsilon$-SV source is used to choose the measurement settings $u_{1 \leq M_1}, u_{2 \leq M_2}$ for the 2 devices. The devices produce output bits $x_{1 \leq M_1}, x_{2 \leq M_2}$.
2. The measurements in device $j$ (for $j = 1, 2$) are partitioned into $N_j$ blocks of boxes each containing $n$ boxes (so that $M_j = nN_j$).
3. The parties choose at random one block of boxes of size $n$ from each device, using bits from the $\varepsilon$-SV source.
4. The parties perform an estimation of the violation of the Bell inequality in the chosen block from each of the two devices by computing the empirical average $L_j := \frac{1}{n} \sum_{i=1}^{n} B(x^j_i, u^j_i)$. The protocol is aborted unless for both of them, $L_j \leq \left( \frac{1}{2} - \varepsilon \right)^4 \frac{\delta}{2} (1 - \mu)$ (with fixed constants $\delta, \mu > 0$).
5. Conditioned on not aborting in the previous step, the parties apply the extractor from part (i) of Lemma 5 to the sequence of outputs from the chosen block in each device.

FIG. 2: Protocol for device-independent randomness amplification from a two devices

Protocol II

1. The $\varepsilon$-SV source is used to choose the measurement settings $u_{1 \leq M_1}, \ldots, u_{k \leq M_k}$ for the $k$ devices. The devices produce output bits $x_{1 \leq M_1}, \ldots, x_{k \leq M_k}$.
2. The measurements in device $j$ are partitioned into $N_j$ blocks of boxes each containing $n$ boxes (so that $M_j = nN_j$).
3. The parties choose at random one block of boxes of size $n$ from each device for $1 \leq j \leq k$, using bits from the $\varepsilon$-SV source.
4. The parties perform an estimation of the violation of the Bell inequality in the chosen block from each device $1 \leq j \leq k$ by computing the empirical average $L_j := \frac{1}{n} \sum_{i=1}^{n} B(x^j_i, u^j_i)$. The protocol is aborted unless $\forall j, L_j \leq \left( \frac{1}{2} - \varepsilon \right)^4 \frac{\delta}{2} (1 - \mu)$ (with fixed constants $\delta, \mu > 0$).
5. Conditioned on not aborting in the previous step, the parties apply a deterministic extractor from part (ii) of Lemma 5 to the sequence of outputs from the chosen block in each device.

FIG. 3: Protocol for device-independent randomness amplification from $k$ devices for $k > 2$.

Proof. Recall that the superscripts denote the device and the subscripts denote the block of boxes, while capital letters indicate the set of boxes within the block so that $U^j_{1 \leq N_j}$ and $X^j_{1 \leq N_j}$ denote the inputs and outputs of the block of boxes numbered 1 to $N_j$ in the $j$-th device.

Under the Markov assumption that the SV source and the box are uncorrelated (see Section II B), Lemma 15 gives that the block of boxes from the first device is approximately uncorrelated with the block of boxes chosen from the second device with the SV source. The test for the Bell value from Section IV A is performed on the two blocks of boxes. When both blocks of boxes pass the test, Lemmas 14 and 10 give that each such block is a source with linear min-entropy in the size of the block $n$. Applying the independent-source extractors from Section II A, one obtains a string of nearly uniform random bits.

In more detail, assume the test performed on the chosen block of boxes from each device ($j =$
1, 2) accepts. Then

$$L_j := \frac{1}{n} \sum_{i=1}^{n} B(x_j^i, u_j^i) \leq \left( \frac{1}{2} - \varepsilon \right)^4 (1 - \mu) \frac{\delta}{2}$$

(18)

From Lemma 12, we know that for each device, with probability at least $1 - 2 \exp \left( -n \frac{\delta^2(1-\mu)^2(1/2-\varepsilon)^8}{8} \right)$, the fraction of the $n$ boxes with uniform Bell value $B_j' \geq \frac{\delta}{|U|}$ is less than $(1 - \mu)$. Then, Lemma 10 gives that for any no-signaling box with uniform Bell value less than $\frac{\delta}{|U|}$, for any measurement setting $u \in U$ and any output $x$, $\Pr(x|u, Z, W) \leq \left( \frac{1+2\delta}{3} \right)$. Therefore, from Lemma 14, we have that for each device for any chosen set of inputs $(u_1, \ldots, u_n)$, the output distribution of the chosen $n$ boxes gives with probability at least $1 - 4\sqrt{\delta'}$ a min-entropy source with entropy:

$$H_{\text{min}}(P) \geq -\log \left[ \frac{1}{1 - \delta'} \left( \frac{1 + 2\delta}{3} \right)^\mu n + \sqrt{\delta'} \right]$$

(19)

where $\delta' = 2 \exp \left( -n \frac{\delta^2(1-\mu)^2(1/2-\varepsilon)^8}{8} \right)$.

By Lemma 15, it follows that the the distributions on outputs of the chosen boxes in the two devices is close to independent. We can then apply the results on independent randomness extractors for two independent $(n, H_{\text{min}}(S))$ sources in [2] where it was shown that a random hash function can be used to obtain $\Theta(n)$ bits $S'_n$ with error $2^{-\Omega(n)}$. This only shows that $d(S'_n|Z, W) \leq 2^{-\Omega(n)} + 2 \exp \left( -\Omega \left( n \frac{\delta^2(1-\mu)^2(1/2-\varepsilon)^8}{8} \right) \right)$.

(20)

To obtain Eq. (17) we throw away a $(1 - \nu)$ fraction of the bits in $S'$ (for $\nu$ sufficiently small), obtaining $S$, and apply Lemma 6.

It remains to estimate the error in the de Finetti bound in Lemma 15. Let us apply Lemma 15 with the number of blocks as specified in Eq. (16). Then we obtain that

$$\Pr_{A_2 \sim \nu([N_2])} \left( T \geq \frac{1}{t} \right) \leq \frac{1}{t}.$$

(21)

with $T$ given by Eq. (58). Thus with probability larger than $1 - 1/t$ over $A_2$ we have $T \leq \frac{1}{t}$. By Markov inequality we have that for such good choice of $A_2$

$$\Pr_{U_{2,A_2} \sim \nu_{A_2} U_{2,A_2} \sim P} (T' \geq \eta) \leq \frac{T}{\eta} \leq \frac{1}{t\eta}.$$

(22)

with

$$T' := \mathbb{E}[U; U_{2,A_2}] \left\| q(X^1, X^2_{A_2}|U^1, U^2_{A_2}) - q(X^1|U^1) \otimes q(X^2_{A_2}|U^2_{A_2}) \right\|_1.$$

(23)

Choosing $\eta = 1/\sqrt{t}$ we find that with probability larger than $1 - 1/t - 1/\sqrt{t} \geq 1 - 2/\sqrt{t}$, the boxes from the first device, the chosen block of boxes from the second device, and the chosen inputs will be such that $T' \leq 1/\sqrt{t}$.

The robustness of the protocol follows from Lemma 17. □
IV. TOOLS FOR RANDOMNESS AMPLIFICATION

In this section we give the tools we will employ proving the correctness of Protocol 2.

A. Verification of the Bell value

The lemma below states that, with high probability, the arithmetic average of mean values for conditional boxes is close to the observed value. The variables $W_i$ will be later interpreted as settings and outputs, i.e. $W_i = (u_i, x_i)$, and the expectation value of variable $B_i$ as the Bell value with inputs taken from a SV-source. It has a similar flavour to previous results [19-21] obtained in the context of the related problem of randomness expansion.

Lemma 8. Consider arbitrary random variables $W_i$ for $i = 1, \ldots, n$, and binary random variables $B_i$ that are functions of $W_i$, i.e. $B_i = f_i(W_i)$ for some functions $f_i$. Let us denote $B_1 = E(B_1)$ (i.e. $B_i$ are conditional means). Define for $k = 2, \ldots, n$, the empirical average $Z_k = \frac{1}{k} \sum_{i=1}^{k} B_i$ and the arithmetic average of conditional means $\mathbb{Z}_k = \frac{1}{k} \sum_{i=1}^{k} B_i$. Then we have

\[ \Pr(|Z_n - \mathbb{Z}_n| \geq s) \leq 2e^{-\frac{n s^2}{2}} \]

Before we prove the lemma, we need to state Azuma inequality. Let $X_0, \ldots, X_k$ and $Y_0, \ldots, Y_k$ be two sequences of random variables. Then $X_0, \ldots, X_k$ is said to be a martingale with respect to $W_0, \ldots, W_k$ if for all $0 \leq i \leq k$, $\mathbb{E}|X_i| < \infty$ and $\mathbb{E}(X_i|W_0, \ldots, W_{i-1}) = X_{i-1}$.

Lemma 9. (Azuma-Hoeffding) Suppose $X_0, \ldots, X_k$ is a martingale with respect to $W_0, \ldots, W_k$, and that $|X_{l+1} - X_l| \leq c_l$ for all $0 \leq l \leq k - 1$. Then for all positive reals $t$,

\[ \Pr(|X_k - X_0| \geq t) \leq 2 \exp \left(-\frac{t^2}{2 \sum_{l=1}^{k} c_l^2} \right) \]

Now we can prove the lemma.

Proof. Define $X_0 = 0, X_l = l(Z_l - \overline{Z}_l)$ and $W_0 = 0$. Let us show that $\{X_i\}_{i=0}^{k}$ and $\{W_i\}_{i=0}^{k}$ satisfy the assumptions of Lemma 9. First,

\[ |X_l - X_{l-1}| = |\overline{B}_l - B_l| \leq 1, \]
since $B_i$ is binary, and $0 \leq B_i \leq 1$. Let us now check that $\{X_i\}_{i=0}^{k}$ is a martingale with respect to $\{W_i\}_{i=0}^{k-1}$. We have $\mathbb{E}|X_l| \leq l < \infty$. Moreover for $l \geq 2$ we have

$$
\mathbb{E}(X_l|W_{l-1}, \ldots, W_0) = \mathbb{E}(X_l|W_{l-1}, \ldots, W_1) = \sum_{i=1}^{l} \mathbb{E}(B_i|W_{l-1}, \ldots, W_1) - \sum_{i=1}^{l} \mathbb{E}(\overline{B}_i|W_{l-1}, \ldots, W_1) = \sum_{i=1}^{l-1} \mathbb{E}(B_i|W_{l-1}, \ldots, W_1) - \sum_{i=1}^{l-1} \mathbb{E}(\overline{B}_i)
$$

(29)

where we used property $\mathbb{E}(\mathbb{E}(A|B)|BC) = \mathbb{E}(A|B)$, where $A, B, C$ are random variables. Now let us note that since $B_i$ is a function of $W_i$, we get for $i \leq l - 1$

$$
\mathbb{E}(B_i|W_{l-1}, \ldots, W_1) = B_i
$$

(30)

Thus the last line of Eq. (32) is equal to $X_{l-1}$, so that we have

$$
\mathbb{E}(X_l|W_{l-1}, \ldots, W_1) = X_{l-1}
$$

(31)

for $l = 2, \ldots, k$. For $l = 1$, Eq. (31) follows from $\mathbb{E}(X_1|W_0) = \mathbb{E}(X_1) = 0 = X_0$, since $X_1 = \mathbb{E}(B_1|W_0) - B_1 = \mathbb{E}(B_1) - B_1$. Now we apply Lemma 9 with $c_l = 1$, and obtain the inequality (26). □

B. From good Bell value to randomness of conditional boxes

In section IV C we will show that to obtain a min-entropy source we need to ensure that a linear fraction of the conditional boxes has randomness. In this section we will argue that the randomness of a box is ensured if the value of the Bell value with inputs taken from a SV source is small (we call the largest such value, over SV sources, the SV Bell value). Let $U$ denote all the settings appearing in the Bell expression. We consider first the uniform Bell value

$$
\overline{B}^U = \frac{1}{|U|} \mathbb{B}\{P(x|u)\} = \frac{1}{|U|} \sum_{u, x} B(x, u) P(x|u)
$$

(32)

where $|U|$ denotes the total number of settings in the Bell expression. If the Bell function $B(x, u)$ is properly chosen, one can prove using linear programming that if $\overline{B}^U$ is small, the probabilities of any output are bounded away from 1. However, since our inputs to each device are chosen using a SV source, we will be only able to estimate the value of the following expression

$$
\overline{B}^{SV} = \sum_{u, x} \nu_{SV}(u) B(x, u) P(x|u)
$$

(33)

where $\nu_{SV}(u)$ is the distribution from the SV source. Again, for a suitably chosen Bell function, when the latter expression is small the former is also small, and implies randomness.
1. Bounding output probabilities by linear programming

Let us show that for the specific Bell inequality we consider, when the value of the Bell expression is small there is weak randomness. Consider a four-partite no-signaling box \( P(x|u) \) that obtains a value \( \delta \) for the Bell expression in Eq. (6). The following lemma shows that for any measurement setting \( u \), the probability of any outcome \( x \) is bounded from above by a function of \( \delta \).

**Lemma 10.** Consider a four-partite no-signaling box \( P(x|u) \) satisfying

\[
B^U := \frac{1}{|U|} B.\{P(x|u)\} \leq \delta / |U|, \tag{37}
\]

for some \( \delta \geq 0 \), with \( B \) the indicator vector for the Bell expression in Eq. (6) and \( |U| = 16 \) the number of settings in the Bell expression. For any measurement setting \( u^* \) and any output \( x^* \), we have

\[
\Pr (x^*|u^*, Z, W) \leq \left( \frac{1 + 2\delta}{3} \right). \tag{38}
\]

**Proof.** Consider any measurement setting \( u^* \) and any corresponding output \( x^* \) for this setting. Then \( \Pr (x^*|u^*, Z, W) \) can be computed by the following linear program

\[
\Pr (x^*|u^*, Z, W) = \max_{\{P\}} : M^T_{u^*,x^*}\{P(x|u)\} \tag{39}
\]

s.t. \( A\{P(x|u)\} \leq c \).

Here, the indicator vector \( M_{u^*,x^*} \) is an \( 2^4 \times 2^4 \) element vector with entries \( M_{u^*,x^*}(x, u) = I_{u=u^*}I_{x=x^*} \). The constraint on the box \( \{P(x|u)\} \) written as a vector with \( 2^4 \times 2^4 \) entries is given by the matrix \( A \) and the vector \( c \). These encode the no-signaling constraints between the four parties, the normalization and the positivity constraints on the probabilities \( P(x|u) \). In addition, \( A \) and \( c \) also encode the condition that \( B.\{P(x|u)\} \leq \delta \) with \( \delta \) the bound on the Bell value for the box. Analogous programs can be formulated for each of the \( 2^4 \) measurement settings appearing in the Bell inequality in Eq. (6) and each of the \( 2^4 \) corresponding outputs.

The solution to the primal linear program in Eq. (39) can be bounded by any feasible solution to the dual program which is written as

\[
\min_{\lambda_{u,x}} : c^T\lambda_{u,x} \tag{40}
\]

s.t. \( A^T\lambda_{u,x} = M_{u,x} \), \( \lambda_{u,x} \geq 0 \).

For each \( \{u^*,x^*\} \), we find a feasible \( \lambda_{u^*,x^*} \) satisfying the constraints to the dual program above that gives \( c^T\lambda_{u^*,x^*} \leq \left( \frac{1 + 2\delta}{3} \right) \). We therefore obtain by the duality theorem of linear programming that

\[
\Pr (x^*|u^*, Z, W) \leq \left( \frac{1 + 2\delta}{3} \right), \tag{41}
\]

which is the required bound. □
2. Obtaining randomness of individual conditional boxes

Let us start with a simple lemma relating the SV Bell value with the true Bell value.

**Lemma 11.**

\[
\frac{1}{\left(\frac{1}{2} + \varepsilon\right)^4 |U|} B_{SV} \leq B^U \leq \frac{1}{\left(\frac{1}{2} - \varepsilon\right)^4 |U|} B_{SV}
\]  

(42)

**Proof.** It follows directly from the definition of a SV source, since

\[
\left(\frac{1}{2} - \varepsilon\right)^4 \leq P(u) \leq \left(\frac{1}{2} + \varepsilon\right)^4.
\]  

(43)

Consider the test for the Bell value in Protocol 2, namely for each device \(j\), one rejects unless

\[
L_j = \frac{1}{n} \sum_{i=1}^{n} B(x_j^i, u_j^i) \leq \left(\frac{1}{2} - \varepsilon\right)^4 \frac{\delta}{2} (1 - \mu)
\]  

(44)

for constants \(\mu\) and \(\delta\). We now show that when the test accepts for device \(j\), with high probability at most a \((1 - \mu)\) fraction of the \(n\) boxes have the uniform Bell value \(B_U \geq \delta |U|\).

**Lemma 12.** Assume that the test given by Eq. (44) for the box \(P(x^1, \ldots, x^n | u^1, \ldots, u^n)\) accepts (for fixed constants \(\mu\) and \(\delta\)). Consider the set \(I := \left\{i : B_{i}^{U} \geq \frac{\delta}{|U|}\right\}\). With probability at least

\[
1 - 2 \exp\left(-n \frac{\delta^2 (1 - \mu)^2 (\frac{1}{2} - \varepsilon)^8}{8}\right), |I| \leq (1 - \mu)n.
\]

**Proof.** If the test accepts, we have that \(L = \frac{1}{n} \sum_{i=1}^{n} B(x_i, u_i) \leq \left(\frac{1}{2} - \varepsilon\right)^4 (1 - \mu)^2 \frac{\delta}{2}\). Let \(\bar{L} = \frac{1}{n} \sum_{i=1}^{n} B_{i}^{SV}\). Then from Eq. (26) in Section IV A we have

\[
\Pr\left(L \geq \left(\frac{1}{2} - \varepsilon\right)^4 \delta (1 - \mu)\right) \leq 2e^{-n \frac{\delta^2 (1 - \mu)^2 (\frac{1}{2} - \varepsilon)^8}{8}}.
\]  

(45)

Consider the set \(I := \left\{i : B_{i}^{U} \geq \frac{\delta}{|U|}\right\}\). Then from Eq. (42), with probability at least

\[
1 - 2 \exp\left(-n \frac{\delta^2 (1 - \mu)^2 (\frac{1}{2} - \varepsilon)^8}{8}\right), \text{ we have } \frac{|I|}{n} \delta \leq (1 - \mu) \delta \text{ and the fraction of boxes with average value } B^U \geq \frac{\delta}{|U|} \text{ is less than } (1 - \mu).
\]

C. From randomness of conditional boxes to min-entropy source

In this section we show that if a device is such that a linear amount of conditional boxes have randomness (in the weak sense that probability of their outputs is bounded away from one), then the distribution on outputs constitutes a min-entropy source. The considerations in this section will be applicable to any of the devices \(j \in [k]\) and any chosen block. Therefore we will skip the indices for simplicity. We will show, that if, with large probability over sequences \(x_1, u_1, \ldots, x_n, u_n\), a linear fraction of those boxes have, for any setting, probability of any output bounded away from 1, then the total probability distribution is close in variational norm to a min-entropy source (See [19–21] for a similar result in the context of randomness expansion).

We first prove, that if this happens for all sequences (i.e. with probability 1), then total box is a min-entropy source itself and subsequently consider the case when the probability is close to 1.
Lemma 13. Fix any measure $P$ on space of sequences $(x_1, u_1, \ldots, x_n, u_n)$. Suppose that for a sequence $(x_1, u_1, \ldots, x_n, u_n)$, there exists $K \subseteq [n]$ of size larger than $\mu n$, such that for all $l \in K$ the conditional boxes $P_{x,l,u<l}(x_l|u_l)$ satisfy

$$P_{x,l,u<l}(x_l|u_l) \leq \gamma,$$  \hfill (46)

for some $0 \leq \gamma < 1$. Then for any sequence $(u_1, \ldots, u_n)$, $P(x_1, \ldots, x_n|u_1, \ldots, u_n)$ satisfies

$$P(x_1, \ldots, x_n|u_1, \ldots, u_n) \leq \gamma^{\mu n}$$ \hfill (47)

Proof. The proof proceeds by successive application of the Bayes rule and the time-ordered non-signaling structure.

$$P(x_1, \ldots, x_n|u_1, \ldots, u_n) = \Pr(x_1|u_1) \Pr(x_2|u_2, x_1) \ldots \Pr(x_n|u_n, x_{n-1}, u_{n-1}, \ldots, x_1, u_1)$$ \hfill (48)

where we have used the fact that the outputs of the $l$-th box can depend only upon the inputs and outputs of the previous boxes due to the time-ordered structure of the boxes. Now, due to the assumption that at least $\mu n$ of the conditional boxes $P_{x,l,u<l}(x_l|u_l)$, satisfy Eq. (46), we have that

$$P(x_1, \ldots, x_n|u_1, \ldots, u_n) \leq \gamma^{\mu n}.$$ \hfill (49)

\hfill \Box

Let us now deal with the case, where the linear number of conditional boxes satisfy (50) on some subset of sequences $x_1, u_1, \ldots, x_n, u_n$ whose probability is close to 1.

Lemma 14. Consider arbitrary probability distribution $P$ on space of sequences $x_1, u_1, \ldots, x_n, u_n$. Suppose that for any sequence $x_1, u_1, \ldots, x_n, u_n$ from a set $S$ of measure $Q > 1 - \delta'$, with $\delta' \leq 1/2$, there exists $K$, a subset of $[n]$ of size larger than $\mu n$, such that for all $l \in K$ the conditional boxes $P_{x,l,u<l}(x_l|u_l)$ satisfy

$$P_{x,l,u<l}(x_l|u_l) \leq \gamma.$$ \hfill (50)

for $0 \leq \gamma < 1$. Denote $P_{(u_1, \ldots, u_n)}$ to be conditional probability distribution over output sequences $(x_1, \ldots, x_n)$ given by $P(x_1, \ldots, x_n|u_1, \ldots, u_n)$. Then with probability $1 - 4\sqrt{\delta'}$, for any $(u_1, \ldots, u_n)$, $P_{(u_1, \ldots, u_n)}$ forms a min-entropy source with min-entropy given by

$$H_{\text{min}}(P_{(u_1, \ldots, u_n)}) \geq -\log \left( \frac{\gamma^{\mu n}}{1 - \delta'} + \sqrt{\delta'} \right).$$ \hfill (51)

Proof. We define a distribution $P'$ on the space of sequences $x_1, u_1, \ldots, x_n, u_n$ as follows: it is equal to $P/P(S)$ if $u_1, \ldots, u_n \in S$, and zero otherwise. Since $P(S) \geq 1 - \delta'$ we have

$$||P - P'|| \leq 4\delta'.$$ \hfill (52)

Now, we have that $||P - P'|| = \sum_{u_1, \ldots, u_n} P(u_1, \ldots, u_n)||P_{(u_1, \ldots, u_n)} - P'_{(u_1, \ldots, u_n)}||$, hence by the above inequality we have:

$$\sum_{u_1, \ldots, u_n} P(u_1, \ldots, u_n)||P_{(u_1, \ldots, u_n)} - P'_{(u_1, \ldots, u_n)}|| \leq 4\delta'.$$ \hfill (53)

Using the Markov inequality, we obtain that

$$\Pr_{u_1, \ldots, u_n \sim P(u_1, \ldots, u_n)} (||P_{(u_1, \ldots, u_n)} - P'_{(u_1, \ldots, u_n)}|| \geq \sqrt{\delta'}) \leq 4\sqrt{\delta'}. \hfill (54)$$
Eq. (47) of lemma 13 implies that for any \( \mathbf{u} \in \mathbf{S} \) we have \( P(x_1, \ldots, x_n|u_1, \ldots, u_n) \leq \gamma^{\mu_n} \) hence
\[
P'(x_1, \ldots, x_n|u_1, \ldots, u_n) \leq \frac{\gamma^{\mu_n}}{1-\delta'}
\] (55)
for all \( (x_1, \ldots, x_n|u_1, \ldots, u_n) \). Now, with probability at least \( 1 - 4\sqrt{\delta'} \) we have \( ||P(u_1, \ldots, u_n) - P'(u_1, \ldots, u_n)|| \leq \sqrt{\delta'} \) so that with this probability for all \( (x_1, \ldots, x_n|u_1, \ldots, u_n) \) we have
\[
P(x_1, \ldots, x_n|u_1, \ldots, u_n) \leq \frac{\gamma^{\mu_n}}{1-\delta'} + \sqrt{\delta'}.
\] (56)
The bound on the min-entropy \( H_{\min}(P) \) in Eq. (51) follows.

\[\Box\]

**D. Imposing independence between devices: de Finetti bounds**

Consider two devices, the first consisting of \( n \) boxes and the second consisting of \( N_2 \) blocks of \( n \) boxes each. In this section, we show that for suitable choice of \( N_2 \), the boxes from the first device are close to being uncorrelated with the boxes in a block chosen from the second device using an \( \varepsilon \)-SV source. The lemma follows a similar formulation to the Lemma 6 in [3], which considered the case \( n = 1 \) for an arbitrary number of devices, and is based on the information-theoretic approach of [30][31] for proving de Finetti theorems for quantum states and non-signaling distributions.

We denote the box by \( P(X^1, X^2_{\leq N_2}|U^1, U^2_{\leq N_2}) \), where the superscript denotes the device and the subscript denotes the block of boxes. Capital letters denote the inputs and outputs for a set of \( n \) boxes so that \( X^1 = (x^1_1, \ldots, x^1_t) \) and \( X^2_{\leq N_2} = (x^2_{1,1}, \ldots, x^2_{n,1}, \ldots, x^2_{1,N_2}, \ldots, x^2_{n,N_2}) \) with the second subscript denoting the block.

**Lemma 15.** Let \( P(X^1, X^2_{\leq N_2}|U^1, U^2_{\leq N_2}) \) be a \((2; n, N_2 n)\) time-ordered non-signaling distribution, with output and input alphabets \( \Sigma \) and \( \Delta \), respectively (i.e. \( P : \Sigma^{(N_2+1)n} \times \Delta^{(N_2+1)n} \rightarrow \mathbb{R}^+ \)). The distribution \( P \) represents two devices with the first containing \( n \) boxes and the second \( N_2 \) blocks of \( n \) boxes each. Let \( A_2 \in [N_2] \) and \((U^1, U^2_{\leq N_2})\) be chosen from an \( \varepsilon \)-SV source \( \nu(A_2, U^1, U^2_{\leq N_2}) \). Then for every positive real \( t \), we have
\[
P_{A_2 \sim \nu} \left( T \geq \sqrt{\frac{2 \ln (2) n \log |\Sigma|} {N_2^{1-\log (1+2\varepsilon)}}} \right) \leq \frac{1}{t},
\] (57)
with
\[
T := \mathbb{E}_{U^1, U^2_{\leq N_2} \sim \nu, A_2} \mathbb{E}_{X^2_{< A_2} \sim P} \left\| q(X^1, X^2_{< A_2}|U^1, U^2_{A_2}) - q(X^1|U^1) \otimes q(X^2_{A_2}|U^2_{A_2}) \right\|_1
\] (58)
where \( q \) is the conditional box given the inputs \( U^2_{A_2} \) and outputs \( X^2_{< A_2} \) of all prior boxes to the ones in block \( A_2 \),
\[
q(X^1, X^2_{A_2}|U^1, U^2_{A_2}) : = P_{X^2_{< A_2}, U^2_{A_2}}(X^1, X^2_{A_2}|U^1, U^2_{A_2}),
\] (59)
and \( \nu_{A_2} \) is the probability \( \nu \) conditioned on the value of \( A_2 \).

**Proof.** Using the upper bound on mutual information \( I(A : B) \leq \min(\log |A|, \log |B|) \) and the chain rule \( I(A : BC) = I(A : B) + I(A : C|B) \), we have that for every \( A_2 \in [N_2] \),
\[
n \log |\Sigma| \geq \mathbb{E}_{U^1, U^2_{\leq N_2} \sim \nu, A_2} I(X^1 : X^2_{\leq N_2})_P \]
\[
= \mathbb{E}_{U^1, U^2_{\leq N_2} \sim \nu, A_2} \left[ I(X^1 : X^2_{\leq N_2})_P + \ldots + I(X^1 : X^2_{N_2} X^2_{< N_2})_P \right]
\] (60)
Therefore if \( A_2, U^1, U^2_{\leq N_2} \) are chosen from a \( \varepsilon \)-SV source \( \nu \),
\[
E_{A_2 \sim \nu} E_{U^1, U^2_{\leq N_2} \sim \nu_{A_2}} E_{X^1_{\leq A_2} \sim P} I(X^1 : X^2_{A_2}) \leq \left( \frac{1}{2} + \varepsilon \right) \log N_2 \log |\Sigma|.
\]  
(61)

We now use Pinsker’s inequality relating the mutual information and trace distance as
\[
I(A : B)_p \geq \frac{1}{2 \ln(2)} \| p_{A,B} - p_A \otimes p_B \|_1^2,
\]  
(62)
and the convexity of \( x \mapsto x^2 \) to obtain
\[
E_{A_2 \sim \nu} E_{U^1, U^2_{\leq N_2} \sim \nu_{A_2}} E_{X^1_{\leq A_2} \sim P} \| q(X^1, X^2_{A_2} | U^1, U^2_{A_2}) - q(X^1 | U^1) \otimes q(X^2_{A_2} | U^2_{A_2}) \|_1 \leq \sqrt{2 \ln(2) N_2 \log(1 + 2 \varepsilon) \log |\Sigma| / N_2}
\]  
(63)

By Markov inequality,
\[
\Pr_{A_2 \sim \nu} \left( T \geq \sqrt{\frac{2 n \ln(2) t \log |\Sigma|}{N_2}} \right) \leq \frac{\sqrt{\Lambda_2 \log(1 + 2 \varepsilon)}}{t},
\]  
(64)

with
\[
T := E_{U^1, U^2_{\leq N_2} \sim \nu_{A_2}} E_{X^1_{\leq A_2} \sim P} \| q(X^1, X^2_{A_2} | U^1, U^2_{A_2}) - q(X^1 | U^1) \otimes q(X^2_{A_2} | U^2_{A_2}) \|_1
\]

Finally, to obtain Eq. (57) we replace \( t \) by \( t^2 N_2 \log(1 + 2 \varepsilon) \) in the equation above.

The lemma above can be generalized to the scenario where one considers \( k > 2 \) devices, which will be used in Protocol II in Fig. 3. Consider \( k \) devices with the \( j \)-th device consisting of \( N_j \) blocks of \( n \) boxes each, for \( j \in [k] \). It shows that for suitable choices of \( N_j \), \( k \) blocks of boxes, chosen one from each device using an \( \varepsilon \)-SV source are close to being uncorrelated with each other.

In the following, we consider the time-ordered non-signaling box \( P(X^1_{\leq N_1}, \ldots, X^k_{\leq N_k} | U^1_{\leq N_1}, \ldots, U^k_{\leq N_k}) \). Here, the superscripts denote the device and the subscripts denote the block of boxes, while capital letters indicate the set of boxes within the block. Therefore, \( U^j_{\leq N_j} \) and \( X^j_{\leq N_j} \) denote the inputs and outputs of the block of boxes numbered 1 to \( N_j \) in the \( j \)-th device with \( X^j_{\leq N_j} \) being short for the outputs \( x^j_1, \ldots, x^j_{N_j+1}, x^j_{2n+2}, \ldots, x^j_{(N_j-1)n+1}, \ldots, x^j_{N_j n, N_j} \) where the second subscript labels the block. The proof of the following lemma follows an analogous route to Lemma 15 above and a similar lemma from our previous paper [5].

**Lemma 16.** Let \( P(X^1_{\leq N_1}, \ldots, X^k_{\leq N_k} | U^1_{\leq N_1}, \ldots, U^k_{\leq N_k}) \) be a \( (k; N_1 n, N_2 m, \ldots, N_k n) \) time-ordered non-signaling distribution, with output and input alphabets \( \Sigma \) and \( \Lambda \), respectively (i.e. \( P : \Sigma^{k \times (\sum_{i=1}^k N_i)} \times \Lambda^{\sum_{i=1}^k N_i, n} \rightarrow [\mathbb{R}]^+ \)). The distribution \( P \) represents \( k \) devices with the \( j \)-th device containing \( nN_j \) boxes. Assume \((A_1, \ldots, A_k) \in [N_1] \times \ldots \times [N_k] \) are chosen using an \( \varepsilon \)-SV source \( \nu \) that is uncorrelated with the box \( P \). Then for any probability distribution \( \nu(U^1_{\leq N_1}, \ldots, U^k_{\leq N_k}) \) and for every set of positive reals \( \{t_2, \ldots, t_k\} \), we have
\[
\Pr_{(A_1, \ldots, A_k) \sim \nu} \left( T \geq \sum_{i=2}^k \sqrt{\frac{2 \log |\Sigma| \ln(2) n t_i^2 \sum_{j=1}^{i-1} N_j}{N_i^{1-\log(1+2 \varepsilon)}}} \right) \leq \sum_{i=2}^k \frac{1}{t_i},
\]  
(65)
with

\[
T := \mathbb{E}_{u^{1}_{N_{1}}, \ldots, u_{N_{k}}^{k}} \mathbb{E}_{x^{i}_{A_{1}} \ldots x^{i}_{A_{k}}} P
\left\| q(X_{A_{1}}^{i}, \ldots, X_{A_{k}}^{i} | U_{A_{1}}^{i}, \ldots, U_{A_{k}}^{i}) - q(X_{A_{1}}^{i}, U_{A_{1}}^{i}) \otimes \cdots \otimes q(X_{A_{k}}^{i}, U_{A_{k}}^{i}) \right\|_{1}
\]

(66)

where \( q \) is the conditional box given the inputs \( U_{A_{j}}^{i} \) and outputs \( X_{A_{j}}^{i} \) of all prior boxes for all \( 1 \leq j \leq k \):

\[
q(X_{A_{1}}^{i}, \ldots, X_{A_{k}}^{i} | U_{A_{1}}^{i}, \ldots, U_{A_{k}}^{i}) := P_{X_{A_{1}}^{i} \ldots x_{A_{k}}^{i}} \left( X_{A_{1}}^{i}, \ldots, X_{A_{k}}^{i} | U_{A_{1}}^{i}, \ldots, U_{A_{k}}^{i} \right),
\]

(67)

and \( \nu_{A_{1}, \ldots, A_{k}} \) is the probability \( \nu \) conditioned on measuring \((A_{1}, \ldots, A_{k})\). Moreover, there exist suitable choice of block sizes \( N_{j} \) and reals \( \{t_{i}\} \) such that

\[
\Pr_{(A_{1}, \ldots, A_{k}) \sim \nu} \left( T \geq \frac{1}{l'} \right) \leq \frac{1}{l'}
\]

(68)

for some suitably large parameter \( l' \).

The proof is completely analogous to the proof of Lemma 6 of [8] and is given in the appendix for the convenience of the reader.

E. Robustness of the protocol

In the remainder of this section, we would like to estimate the amount of noise that the protocol can tolerate. Suppose we are given a box \( P(x_{1}, \ldots, x_{n} | u_{1}, \ldots, u_{n}) \) such that for every input and output and \( \forall i \in [n], \) \( B \{ P_{x_{i}, u_{i}} (x_{i} | u_{i}) \} \leq \delta \). This will be the case, for example, if all the the entangled states and measurements used to produce a box are only \( O(\delta) \)-close to the ones that would lead to a box violating maximally the Bell inequality (i.e. \( n \) copies of the entangled state given by Eq. (9), each measured in the bases given by Eqs. (10) and (11)).

**Lemma 17.** Consider the verification procedure applied to a box \( P(x_{1}, \ldots, x_{n} | u_{1}, \ldots, u_{n}) \) such that for every input and output and \( \forall i \in [n], \) \( B \{ P_{x_{i}, u_{i}} (x_{i} | u_{i}) \} \leq \delta \). Then the test accepts with probability \( 1 - 2 \exp \left( -\frac{n(1/2-\epsilon)^2(1-\mu)^2\delta^2}{32} \right) \) as long as

\[
\delta \leq \frac{(1-\mu)\delta(1/2-\epsilon)^{4}}{4(\frac{1}{2}+\epsilon)^{4}},
\]

(69)

**Proof.** We have that for all \( 1 \leq i \leq n \),

\[
\delta_{i} := B \{ P_{x_{1}, \ldots, x_{i-1}, u_{i}} (x_{i} | u_{i}) \} \leq \delta,
\]

(70)

Denoting as before, \( B^{SV}_{i} := \mathbb{E}_{u_{i} \sim \nu} \mathbb{E}_{x_{i} \sim P_{x_{i}, u_{i}}} B(x_{i}, u_{i}), \) \( L = \frac{1}{n} \sum_{i=1}^{n} B(x_{i}, u_{i}) \) and \( L = \frac{1}{n} \sum_{i=1}^{n} B^{SV}_{i}, \) we know from Lemma 9 that

\[
\Pr \left( L \geq \overline{L} + s \right) \leq 2 e^{-s^{2}n/2}.
\]

(71)

From Eq. (70) and the relation between the uniform and non-uniform Bell values in Eq. (42), we have

\[
\overline{L} \leq \frac{(1/2+\epsilon)^{4}}{n} \sum_{i=1}^{n} \delta_{i} \leq \left( \frac{1}{2} + \epsilon \right)^{4} \overline{\delta}
\]

(72)
Choosing \( s = \left( \frac{1}{2} - \varepsilon \right)^4 (1 - \mu)^4 \), we find that when
\[
\varepsilon \leq \frac{(1 - \mu) \delta (1/2 - \varepsilon)^4}{4(1 + \varepsilon)^4},
\]
(73)
we have
\[
\Pr \left( L \leq \frac{(1 - \mu) \delta (1/2 - \varepsilon)^4}{2} \right) \geq 1 - 2 \exp \left( -n \frac{(1 - \mu)^8 (1 - \mu)^2 \delta^2}{32} \right),
\]
(74)
so that the test is passed with high probability for suitably large \( n \). □

V. CONCLUSION AND OPEN QUESTIONS

We presented a protocol for obtaining secure random bits from an arbitrarily (but not fully) deterministic Santha-Vazirani source. The protocol uses a finite number (as few as eight for the four-partite Bell inequality considered) of non-signaling devices, and works even with correlations attainable by noisy quantum mechanical resources. Moreover the correctness of the protocol is not based on quantum mechanics and only requires the no-signaling principle.

We leave the following open questions to future research:

- The most pressing open question is whether there is an efficient protocol for device-independent randomness amplification tolerating a constant rate of noise and giving a constant rate of random bits (per use of the devices). The protocol of this paper has zero-rate due to the use of the de Finetti bound (Lemma 15).

- Is there a bipartite Bell inequality with the property that it is algebraically violated by quantum correlations, and for all settings in the Bell expression, the probabilities of the outputs are bounded away from one? Our protocol could be applied with such a Bell expression with a significant reduction in the number of no-signaling devices required.

- Is there a protocol that can tolerate a higher level of noise? What if we assume the validity of quantum mechanics?

- Can we amplify randomness from other different types of sources? A particularly interesting case is the min-entropy source [13, 14].

- A more technical question is to improve the de Finetti theorem given in [30, 31]. What are the limits of de Finetti type results when the subsystems are selected from a Santha-Vazirani source?

- Finally suppose one would like to realize device-independent quantum key distribution with only an imperfect SV source as the randomness source. Is there an efficient protocol for that tolerating a constant rate of noise and gives a constant rate of key? Here the question is open for both quantum-mechanical and non-signalling adversaries.

Acknowledgments. The paper is supported by ERC AdG grant QOLAPS and by Foundation for Polish Science TEAM project co-financed by the EU European Regional Development Fund. FB acknowledges support from EPSRC. Part of this work was done in National Quantum Information Center of Gdańsk. Part of this work was done when F. B., R. R., K. H. and M. H. attended the program “Mathematical Challenges in Quantum Information” at the Isaac Newton Institute for Mathematical Sciences in the University of Cambridge.
VI. APPENDIX: PROOF OF LEMMA 16

Proof. Again, the proof of the above lemma follows the same steps as that of Lemma 15 with the inputs $U_{j < A_j}$ and outputs $X_{j < A_j}$ now understood to have occurred earlier in time than $X_{A_j}$ for each $1 \leq j \leq k$.

We first analyze the case when the $k$ blocks $(A_1, \ldots, A_k)$ are chosen from the uniform distribution over $[N_1] \times \ldots \times [N_k]$ and then consider the scenario where they are chosen from an $\varepsilon$-SV source. With $(A_1, \ldots, A_k)$ chosen uniformly, we first show that each of the $j$ blocks is approximately in a product state with the previous $j - 1$ blocks of boxes. The product form of all blocks will then follow by application of the triangle inequality.

As in Lemma 15, applying the upper bound on mutual information $I(A : B) \leq \min(\log |A|, \log |B|)$ and the chain rule $I(A : BC) = I(A : B) + I(A : C | B)$, we have

$$\log |\Sigma| n \sum_{j=1}^{k-1} N_j \geq \mathbb{E}_{U_{N_1}^{k}, \ldots, U_{N_k}^{k}} \mathbb{E}_{X_{N_1}^{k}, \ldots, X_{N_k}^{k}} I(X_{N_1}^{k}, \ldots, X_{N_k}^{k} : X_{A_k}^{k}) \delta = \mathbb{E}_{U_{N_1}^{k}, \ldots, U_{N_k}^{k}} \mathbb{E}_{X_{N_1}^{k}, \ldots, X_{N_k}^{k}} I(X_{N_1}^{k}, \ldots, X_{N_k}^{k} : X_{A_k}^{k}) \delta + \ldots + I(X_{N_1}^{k}, \ldots, X_{N_k}^{k} : X_{N_k}^{k}) \delta$$

(75)

When $A_k$ is chosen from the uniform distribution $U[N_k]$, we get

$$\mathbb{E}_{A_k \sim U[N_k]} \mathbb{E}_{U_{N_1}^{k}, \ldots, U_{N_k}^{k}} \mathbb{E}_{X_{N_1}^{k}, \ldots, X_{N_k}^{k}} I(X_{N_1}^{k}, \ldots, X_{N_k}^{k} : X_{A_k}^{k}) \delta \leq \frac{\log |\Sigma| n \sum_{j=1}^{k-1} N_j}{N_k}.$$  

(76)

Using Pinsker’s inequality from Eq. (62) relating the mutual information and trace distance and the convexity of $x^2$ we obtain

$$\mathbb{E}_{A_k \sim U[N_k]} \mathbb{E}_{U_{N_1}^{k}, \ldots, U_{N_k}^{k}} \mathbb{E}_{X_{N_1}^{k}, \ldots, X_{N_k}^{k}} \left\| q(X_{N_1}^{k}, \ldots, X_{N_k}^{k} : U_{N_1}^{k}, \ldots, U_{N_k}^{k} | U_{A_k}^{k}) - q(X_{N_1}^{k}, \ldots, X_{N_k}^{k} : U_{N_1}^{k}, \ldots, U_{N_k}^{k}) \otimes q(X_{A_k}^{k} | U_{A_k}^{k}) \right\|_1 \leq \sqrt{2 \ln (2) \log |\Sigma| n \sum_{j=1}^{k-1} N_j}.$$  

(77)

The above argument can also be applied to the box $q(X_{N_1}^{k}, \ldots, X_{N_k}^{k} : U_{N_1}^{k}, \ldots, U_{N_k}^{k})$ to give

$$\mathbb{E}_{A_k \sim U([N_k - 1] \times [N_k])} \mathbb{E}_{U_{N_1}^{k}, \ldots, U_{N_k}^{k}} \mathbb{E}_{X_{N_1}^{k}, \ldots, X_{N_k}^{k}} \left\| q(X_{N_1}^{k}, \ldots, X_{N_k}^{k} : U_{N_1}^{k}, \ldots, U_{N_k}^{k} | U_{A_k}^{k}) - \right\| q(X_{N_1}^{k}, \ldots, X_{N_k}^{k} : U_{N_1}^{k}, \ldots, U_{N_k}^{k} \otimes q(X_{A_k}^{k} | U_{A_k}^{k}) \right\|_1 \leq \sqrt{2 \ln (2) \log |\Sigma| n \sum_{j=1}^{k-2} N_j}.$$  

(78)

Using Eqs. (77) and (78), the triangle inequality and the monotonicity of the 1-norm under
discarding subsystems, we obtain

\[
\mathbb{E}_{A_{k-1}, A_k \sim U([N_{k-1}] \times [N_k])} \mathbb{E}_{U_{N_1}^{k}} \cdots \mathbb{E}_{U_{N_k}^{k}} \mathbb{E}_{X_{A_{k-1}}^{k} \sim P} \mathbb{E}_{X_{A_k}^{k} \sim P} q(x_{N_1}^{1}, \ldots, x_{N_k}^{k-2}, A_{k-1}, x_{A_k}^{k} | U_{N_1}^{k}, U_{N_2}^{k-2}, A_{k-1}, U_{A_k}) - q(x_{N_1}^{1}, \ldots, x_{N_k}^{k-2} | U_{N_1}^{k}, U_{N_2}^{k-2}) \otimes q(x_{A_{k-1}}^{k-1} | U_{A_{k-1}}^{k-1}) \otimes q(x_{A_k}^{k} | U_{A_k}^{k}) \leq \sqrt{\frac{2 \ln (2) \log \left( \sum |n_j \sum_{j=1}^{N_j} N_j \right)}{N_k}} + \sqrt{\frac{2 \ln (2) \log \left( \sum |n_j \sum_{j=1}^{N_j} N_j \right)}{N_k}} \tag{79}
\]

Following the above reasoning for the \((k - 2)^{th}\) box up to the second we find

\[
\mathbb{E}_{A_1, \ldots, A_k \sim U([N_1] \times \cdots \times [N_k])} \mathbb{E}_{U_{N_1}^{1}} \cdots \mathbb{E}_{U_{N_k}^{1}} \mathbb{E}_{X_{A_{k-1}}^{1} \sim P} \mathbb{E}_{X_{A_k}^{1} \sim P} \left\| q(x_{A_1}^{1}, \ldots, x_{A_k}^{k} | U_{A_1}^{1}, \ldots, U_{A_k}^{k}) - q(x_{A_1}^{1} | U_{A_1}^{1}) \otimes \cdots \otimes q(x_{A_k}^{k} | U_{A_k}^{k}) \right\|_1 \leq \sqrt{\frac{2 \ln (2) \log \left( \sum |n_j \sum_{j=1}^{N_j} N_j \right)}{N_i}} \tag{80}
\]

Let us now show how to extend the argument to the case when \((A_1, \ldots, A_k)\) are chosen from a \(\varepsilon\)-SV source. Let

\[
R_i := I(x_{N_1}^{1}, \ldots, x_{N_k}^{1} ; x_{A_i}^{i})q.
\]

From the chain-rule argument presented before we have

\[
\mathbb{E}_{A_i \sim U[N_i]} \mathbb{E}_{(A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_k) \sim \nu_{A_i, A_{i-1}, A_{i+1}, \ldots, A_k}} \mathbb{E}_{U_{N_1}^{i}, \ldots, U_{N_k}^{i} \sim \nu_{A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_k}} \mathbb{E}_{X_{A_{i-1}}^{i} \sim P} \mathbb{E}_{X_{A_{i+1}}^{i} \sim P} R_i \leq \log |\Sigma| n \sum_{j=1}^{i-1} N_j \tag{82}
\]

Then from the definition of a \(\varepsilon\)-SV source,

\[
\mathbb{E}_{(A_1, \ldots, A_k) \sim \nu} \mathbb{E}_{U_{N_1}^{i}, \ldots, U_{N_k}^{i} \sim \nu_{A_1, \ldots, A_k}} \mathbb{E}_{X_{A_{i-1}}^{i} \sim P} \mathbb{E}_{X_{A_{i+1}}^{i} \sim P} R_i \leq \mathbb{E}_{A_i \sim U[N_i]} \mathbb{E}_{(A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_k) \sim \nu_{A_i, A_{i-1}, A_{i+1}, \ldots, A_k}} \mathbb{E}_{U_{N_1}^{i}, \ldots, U_{N_k}^{i} \sim \nu_{A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_k}} \mathbb{E}_{X_{A_{i-1}}^{i} \sim P} \mathbb{E}_{X_{A_{i+1}}^{i} \sim P} R_i \leq N_i \left( \frac{1}{2} + \varepsilon \right)^{\log (N_i)} n \log |\Sigma| \sum_{j=1}^{i-1} N_j \tag{83}
\]

Then by Pinsker’s inequality and the convexity of \(x^2\),

\[
\mathbb{E}_{(A_1, \ldots, A_k) \sim \nu} \mathbb{E}_{U_{N_1}^{i}, \ldots, U_{N_k}^{i} \sim \nu_{A_1, \ldots, A_k}} \mathbb{E}_{X_{A_{i-1}}^{i} \sim P} \mathbb{E}_{X_{A_{i+1}}^{i} \sim P} \left\| q(x_{N_1}^{1}, \ldots, x_{N_k}^{k-1} | A_{i}^{1}, \ldots, U_{A_i}^{1} | U_{N_1}^{i}, \ldots, U_{N_i}^{i-1}) - q(x_{N_1}^{1}, \ldots, x_{N_k}^{k-1} | U_{N_1}^{i}, \ldots, U_{N_i}^{i-1}) \otimes q(x_{A_i}^{i} | U_{A_i}^{i}) \right\|_1 \leq \sqrt{2 \ln (2) N_i \log (1+2\varepsilon)} \frac{n \log |\Sigma| \sum_{j=1}^{i-1} N_j}{N_i} \tag{84}
\]

By Markov inequality,

\[
\Pr_{(A_1, \ldots, A_k) \sim \nu} \left( T_i \geq \sqrt{2 \ln (2) n t_i \log |\Sigma| \sum_{j=1}^{i-1} N_j} \right) \leq \sqrt{\frac{N_i \log (1+2\varepsilon)}{t_i}}, \tag{85}
\]
with
\[
T_i := \mathbb{E}_{U_{\leq N_1}^1, \ldots, U_{\leq N_k}^k \sim \nu, A_1, \ldots, A_k} \mathbb{E}_{X_{<A_1}^1, \ldots, X_{<A_k}^k \sim \mathcal{P}} \left\| q(X_{\leq N_1}^1, \ldots, X_{<N_i}^i \mid U_{\leq N_i}^1, \ldots, U_{\leq N_{i-1}}^{i-1}, A_i) - q(X_{\leq N_1}^1, \ldots, X_{\leq N_{i-1}}^{i-1} \mid U_{\leq N_1}^1, \ldots, U_{\leq N_{i-1}}^{i-1}) \otimes q(X_{A_i}^i \mid U_{A_i}^i) \right\|_1
\]

However, by the triangle inequality and the monotonicity of the 1-norm under discarding subsystems, we have that if for all \(2 \leq i \leq k\),
\[
T_i < \sqrt{\frac{2 \ln 2 t_i \log |\Sigma| \sum_{j=1}^{i-1} N_j}{N_i}} =: r_i,
\]
then
\[
T < \sum_{i=2}^k r_i.
\]

Therefore,
\[
\Pr_{(A_1, \ldots, A_k) \sim \nu([N_1] \times \cdots \times [N_k])} \left( T < \sum_{i=2}^k r_i \right) \geq \Pr_{(A_1, \ldots, A_k) \sim \nu([N_1] \times \cdots \times [N_k])} \left( T_2 < T_2 \cap \cdots \cap T_k < r_k \right)
\]
\[
= 1 - \Pr_{(A_1, \ldots, A_k) \sim \nu([N_1] \times \cdots \times [N_k])} \left( T_2 \geq r_2 \cup \cdots \cup T_k \geq r_k \right)
\]
\[
\geq 1 - \sum_{i=2}^k \sqrt{\frac{N_i \log(1+2\varepsilon)}{t_i}},
\]
where the final inequality follows from the union bound and Eq. (85). This gives
\[
\Pr_{(A_1, \ldots, A_k) \sim \nu([N_1] \times \cdots \times [N_k])} \left( T \geq \sum_{i=2}^k \sqrt{\frac{2 \ln 2 nt_i \log |\Sigma| \sum_{j=1}^{i-1} N_j}{N_i}} \right) \leq \sum_{i=2}^k \sqrt{\frac{N_i \log(1+2\varepsilon)}{t_i}}.
\]

Finally, to obtain Eq. (57), we replace \(t_i\) by \(t_i^2 N_i \log(1+2\varepsilon)\) in the equation above. With the choice \(t_i = t\) for all \(i, N_i = 1\),
\[
N_i^{1-\log(1+2\varepsilon)} = 2 \ln 2 \log |\Sigma| n^t \sum_{j=1}^{i-1} N_j,
\]
and denoting \(t' := kt\), we obtain Eq. (68).

\[\square\]

[1] M. Santha and U. V. Vazirani. Generating Quasi-Random Sequences from Slightly-Random Sources. Proceedings of the 25th IEEE Symposium on Foundations of Computer Science (FOCS’84), 434 (1984).
[2] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. SIAM Journal on Computing, 17(2): 230 (1988).
[3] J. Barrett and N. Gisin. How Much Measurement Independence Is Needed to Demonstrate Nonlocality? Phys. Rev. Lett. 106, 100406 (2011).
[4] M. J. W. Hall. Local Deterministic Model of Singlet State Correlations Based on Relaxing Measurement Independence. Phys. Rev. Lett. 105, 250404 (2010).
[5] R. Colbeck and R. Renner. Free randomness can be amplified. Nature Physics 8, 450 (2012).
[6] Xin-Li. Extractors for a constant number of independent sources with polylogarithmic min-entropy. (to appear in FOCS 2013).
[7] R. Gallego, L. Masanes, G. de la Torre, C. Dhara, L. Aolita and A. Acin. Full randomness from arbitrarily deterministic events. [arXiv:1210.6514] (2012).
[8] R. Ramanathan, F. G. S. L. Brandao, A. Grudka, K. Horodecki, M. Horodecki and P. Horodecki. Robust Device-Independent Randomness Amplification. arXiv: 1308.4635 (2013).
[9] O. Guhne, G. Toth, P. Hyllus and H. Briegel, Phys. Rev. Lett. 95, 120405 (2005).
[10] P. Mironowicz and M. Pawłowski. Amplification of arbitrarily weak randomness. [arXiv:1301.7722] (2013).
[11] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, M. Pawłowski and R. Ramanathan. Free randomness amplification using bipartite chain correlations. [arXiv:1303.5591] (2013).
[12] J. E. Pope and A. Kay. Limited Free Will in Multiple Runs of a Bell Test. [arXiv:1304.4904] (2013).
[13] L. P. Thinh, L. Sheridan and V. Scarani. Bell tests with min-entropy sources. [arXiv:1304.3598] (2013).
[14] M. Plesch and M. Pivoluska. Single Min-Entropy Random Source can be Amplified. [arXiv:1305.0990] (2013).
[15] R. Augusiak, M. Demianowicz, M. Pawłowski, J. Tura and A. Acin. Monogamies of correlations and amplification of randomness. arxiv: 1307.6390 (2013).
[16] R. Colbeck and A. Kent. Private Randomness Expansion With Untrusted Devices. Journal of Physics A: Mathematical and Theoretical 44(9), 095305 (2011).
[17] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, and A. Wigderson. Simulating independence: New constructions of condensers, Ramsey graphs, dispersers, and extractors. Proceedings of the 37th Annual ACM Symposium on Theory of Computing, pp 110 (2005).
[18] A. Rao. Extractors for a constant number of polynomially small min-entropy independent sources. Proceedings of the 38th Annual ACM Symposium on Theory of Computing (2006).
[19] S. Pironio et al. Random numbers certified by Bell’s theorem. Nature 464, 1021 (2010).
[20] S. Pironio and S. Massar. Security of practical private randomness generation. Phys. Rev. A 87, 012336 (2013).
[21] S. Fehr, R. Gelles, and C. Schaffner. Security and Composability of Randomness Expansion from Bell Inequalities. [arXiv:1111.6052] (2011).
[22] R. Colbeck, PhD dissertation. Quantum And Relativistic Protocols For Secure Multi-Party Computation. University of Cambridge, arXiv:0911.3814 (2009).
[23] A. Acin, S. Massar and S. Pironio. Randomness versus Nonlocality and Entanglement. Phys. Rev. Lett. 108, 100402 (2012).
[24] J. Barrett, L. Hardy and A. Kent. No Signaling and Quantum Key Distribution. Phys. Rev. Lett. 95, 010503 (2005).
[25] Ll. Masanes. Universally Composable Privacy Amplification from Causality Constraints. Phys. Rev. Lett. 102, 140501 (2009).
[26] E. Hänggi, R. Renner and S. Wolf. Efficient Device-Independent Quantum Key Distribution. EUROCRYPT 2010, 216 (2010).
[27] U. Vazirani and T. Vidick. Certifiable Quantum Dice - Or, testable exponential randomness expansion. [arXiv:1111.6054] (2011).
[28] U. Vazirani and T. Vidick. Fully device independent quantum key distribution. [arXiv:1210.1810] (2012).
[29] M. Coudron, T. Vidick, and H. Yuen. Robust Randomness Amplifiers: Upper and Lower Bounds. [arXiv:1305.6626] (2013).
[30] F. G. S. L. Brandao and A. W. Harrow. Quantum de Finetti Theorems under Local Measurements with Applications. STOC 2013: 861-870. arXiv: 1210.6367 (2012).
[31] F. G. S. L. Brandao and A. W. Harrow. Product-state Approximations to Quantum Groundstates. STOC 2013: 871-880.