Security and composability of randomness expansion from Bell inequalities

Serge Fehr,[1] Ran Gelles,[2] and Christian Schaffner[1,3]
[1]*Centrum Wiskunde & Informatica, Amsterdam, Netherlands*
[2]*Department of Computer Science, University of California, Los Angeles, California 90095, USA*
[3]*Institute for Logic, Language and Computation (ILLC), University of Amsterdam, Amsterdam, Netherlands*
(Received 20 September 2012; published 30 January 2013)

The nonlocal behavior of quantum mechanics can be used to generate guaranteed fresh randomness from an untrusted device that consists of two nonsignalling components; since the generation process requires some initial fresh randomness to act as a catalyst, one also speaks of randomness expansion. R. Colbeck and A. Kent [J. Phys. A **44**, 095305 (2011)] proposed the first method for generating randomness from untrusted devices, but without providing a rigorous analysis. This was addressed subsequently by S. Pironio *et al.* [Nature (London) **464**, 1021 (2010)], who aimed at deriving a lower bound on the min-entropy of the data extracted from an untrusted device based only on the observed nonlocal behavior of the device. Although that article succeeded in developing important tools for reaching the stated goal, the proof itself contained a bug, and the given formal claim on the guaranteed amount of min-entropy needs to be revisited. In this paper we build on the tools provided by Pironio *et al.* and obtain a meaningful lower bound on the min-entropy of the data produced by an untrusted device based on the observed nonlocal behavior of the device. Our main result confirms the essence of the (improperly formulated) claims of Pironio *et al.* and puts them on solid ground. We also address the question of composability and show that different untrusted devices can be composed in an alternating manner under the assumption that they are not entangled. This enables superpolynomial randomness expansion based on two untrusted yet unentangled devices.

DOI: 10.1103/PhysRevA.87.012335 PACS number(s): 03.67.Dd, 03.65.Ud

I. INTRODUCTION

A. Background

One of the counterintuitive features of quantum mechanics is its *nonlocality*: measuring possibly far apart quantum systems in randomly selected bases (chosen out of some given class) may lead to correlations that are impossible to obtain classically. Anticipated by Einstein, Rosen, and Podolsky [1], it was Bell [2] who put this property on firm ground by proposing an inequality that is satisfied by any classical correlation but is violated when the correlation is obtained from measuring entangled quantum states. Such inequalities are called *Bell inequalities*.

An important example of such a Bell inequality was proposed by Clauser, Horne, Shimony, and Holt (CHSH) [3] and states that if X and Y are independent uniformly distributed bits and if bit A is obtained by "processing" X without knowing Y and bit B is obtained by "processing" Y without knowing X, then the probability that $A \oplus B = X \wedge Y$ is at most 75%. This bound on the probability holds if the processing is done classically with shared randomness but can be violated when the processing involves measuring an entangled quantum state; in this latter case, a probability of roughly 85% can be achieved.

Violating a Bell inequality necessarily means that there must be some amount of fresh randomness in the outputs A and B (given the inputs X and Y). More formally, consider an untrusted device $\mathfrak{D}$, prepared by an adversarial manufacturer Eve. The device consists of two components, set up by Eve, which on respective inputs X and Y produce respective outputs A and B *without communicating*. No matter how the two components work, as long as a given Bell inequality is violated during n sequential interactions with $\mathfrak{D}$ (which can be observed by doing statistics), there must be a certain amount of uncertainty in the n output pairs $(A_1, B_1), \ldots, (A_n, B_n)$, even given the n input pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$, and thus it should be possible to apply a randomness extractor to obtain nearly random bits.

This kind of randomness expansion from untrusted devices was first suggested by Colbeck [4] and Colbeck and Kent [5], who presented a scheme that uses Greenberger-Horne-Zeilinger (GHZ) states and reaches a linear expansion, but without providing a rigorous security analysis. The main point missing in these works is a method to rigorously bound the min-entropy of a device's output. The work of Pironio *et al.* [6] addresses this issue, and they propose a technique to numerically compute a lower bound on the min-entropy of the output pair AB (conditioned on X and Y) as a function of the Bell value of the device $\mathfrak{D}$ (which quantifies the violation of Bell inequality). For the special case of CHSH, they also show an analytical bound.

The authors of [6] also consider the case of n sequential interactions with $\mathfrak{D}$, and they show how to *estimate* the average Bell value of $\mathfrak{D}$ over n rounds by doing statistics over the observed data. This is nontrivial because the Bell value of $\mathfrak{D}$ may change over the different rounds, and for each round, it may depend on the behavior of the previous rounds. In other words, the Bell value of $\mathfrak{D}$ during round $i+1$ depends on the history $(A_1, B_1, X_1, Y_1), \ldots, (A_i, B_i, X_i, Y_i)$. Pironio *et al.* then claim to have a bound on the min-entropy

of $(A_1,B_1),\ldots,(A_n,B_n)$ conditioned on $(X_1,Y_1),\ldots,(X_n,Y_n)$, where the bound is actually a function of the observed data, i.e., a function of $(A_1,B_1,X_1,Y_1),\ldots,(A_n,B_n,X_n,Y_n)$.

However, the claimed bound in [6] does not hold in general; there is a flaw in its derivation, which is without an obvious fix.[1] Thus, even though the necessary tools are provided in [6], they are not put together in the right way to be able to control the min-entropy of $(A_1,B_1),\ldots,(A_n,B_n)$ produced by an untrusted device $\mathfrak{D}$.

B. Our result

In this paper, we make up for this shortfall in [6]. Specifically, we put together the tools provided in [6] in order to obtain a correct bound on the min-entropy of $(A_1,B_1),\ldots,(A_n,B_n)$, conditioned on $(X_1,Y_1),\ldots,(X_n,Y_n)$, by means of the observed data. The trick is to consider and bound the min-entropy *conditioned* on the event that the estimator for the average Bell value lies in some interval. This gives us some control over the average Bell value of the device but, as we show, still leaves enough uncertainty in the data to get a good bound on its min-entropy.

We also address the question of the composability of untrusted devices. We show that under the assumption that different devices are not entangled, the output of one device, after privacy amplification, can be used as the input for a second device, and the resulting output of the second device, after privacy amplification, can again be fed into the first device and so on. Using an extractor with a short seed to do the privacy amplification, this allows for a superpolynomial randomness-expansion scheme using two untrusted (but guaranteed to be unentangled) devices.

C. Concurrent and related work

In concurrent and independent work, Vazirani and Vidick [7] as well as Pironio and Massar [8] came up with results that overlap with ours. We briefly discuss here the similarities and the differences between our results and those of Vazirani and Vidick and of Pironio and Massar. We encourage the reader to also look at the comparisons given in [7,8].

Vazirani and Vidick obtain a randomness-expansion scheme with superpolynomial expansion and security against quantum side information. We do not achieve security against quantum side information, and our superpolynomial randomness-expansion scheme requires two unentangled devices in an iterative way, whereas their scheme works with just *one* single device. On the other hand, their result is tailored to CHSH, while our result is generic and holds for any Bell inequality. Pironio and Massar's results, on the other hand, are very similar to ours and only differ in some minor details.[2]

In a very recent paper, Barrett *et al.* point out the possibility of Trojan-horse attacks on device-independent randomness-expansion protocols (see the Supplemental Material of [9]). It seems impossible to prevent Eve from programming devices (that are used multiple times) to release in later rounds information about previous outputs. We note that although such an attack seems unavoidable, in a *single* activation of our randomness-expansion scheme (see Sec. IV for details), we can reuse the same devices over and over again and still prevent such a Trojan-horse attack by only releasing the output of the very last round (and aborting if things go wrong before the last round is reached).

II. PRELIMINARIES

We assume the reader is familiar with quantum information processing, and we merely fix our notation and some basic concepts in this section.

A. Quantum states

The *state* of a quantum system $\mathcal{A}$ is given by a *density matrix* $\rho_{\mathcal{A}}$, i.e., a positive-semidefinite trace-1 matrix acting on some Hilbert space $\mathcal{H}_{\mathcal{A}}$. We denote the set of all such matrices, acting on $\mathcal{H}_{\mathcal{A}}$, by $D(\mathcal{H}_{\mathcal{A}})$. The state space of the joint quantum systems $\mathcal{AB}$, which consist of two (or more) subsystems $\mathcal{A}$ and $\mathcal{B}$, is given by the tensor product $\mathcal{H}_{\mathcal{AB}}=\mathcal{H}_{\mathcal{A}}\otimes\mathcal{H}_{\mathcal{B}}$. If the state of the joint system is given by $\rho_{\mathcal{AB}}$, then the state of subsystem $\mathcal{A}$ when considered as a "stand-alone" system is given by the reduced density matrix $\rho_{\mathcal{A}}=\mathrm{tr}_{\mathcal{B}}(\rho_{\mathcal{AB}})\in D(\mathcal{H}_{\mathcal{A}})$, obtained by tracing out system $\mathcal{B}$.

A random variable X over a finite set $\mathbb{X}$ with probability distribution P_X can be represented by means of the density matrix as $\rho_X=\sum_x P_X(x)|x\rangle\langle x|\in D(\mathcal{H}_X)$, where $\{|x\rangle\}_{x\in\mathbb{X}}$ forms a basis of $\mathcal{H}_X=\mathbb{C}^{|\mathbb{X}|}$. Thus, we may view X as a quantum system, and we say that its state, ρ_X, is *classical*. If the state of a quantum system $\mathcal{E}$ depends on the random variable X, in that the state of $\mathcal{E}$ is given by $\rho_{\mathcal{E}}^x\in D(\mathcal{H}_{\mathcal{E}})$ if $X=x$, then we can view the pair $X\mathcal{E}$ as a bipartite quantum system in state $\rho_{X\mathcal{E}}=\sum_x P_X(x)|x\rangle\langle x|\otimes\rho_{\mathcal{E}}^x\in D(\mathcal{H}_X\otimes\mathcal{H}_{\mathcal{E}})$. This naturally extends to multiple random variables and quantum systems.

The distance between two states $\rho_{\mathcal{E}},\tilde{\rho}_{\mathcal{E}}\in D(\mathcal{H}_{\mathcal{E}})$ is measured by their *trace distance* $\frac{1}{2}\|\rho_{\mathcal{E}}-\tilde{\rho}_{\mathcal{E}}\|_1$, where $\|\cdot\|_1$ is the L_1 norm.[3] In case of classical states ρ_X and $\tilde{\rho}_X$, corresponding to distributions P_X and $\tilde{P}_X$, the trace distance coincides with the statistical distance $\frac{1}{2}\sum_x|P_X(x)-\tilde{P}_X(x)|$.

B. Closeness to uniform, min-entropy, and extractors

In the following definitions, we consider a bipartite system $X\mathcal{E}$ with classical X, given by $\rho_{X\mathcal{E}}$. X is said to be *random*

[1] As a matter of fact, the formal statement is already suspicious since the min-entropy is a fixed number, determined by the underlying probability distribution, whereas the claimed bound is a random variable. This is like saying that we can lower bound the min-entropy of throwing a fair die by the result of the throw. The former equals $\log_2(6)\approx 2.6$, whereas the latter is a random number in $\{1,\ldots,6\}$. We also point out that trying to bound the min-entropy *conditioned on the observed outcome* makes no sense either because this conditional min-entropy obviously vanishes.

[2] As a historical note, previous versions of their paper and our paper claimed security against quantum side information, but both proofs were incorrect.

[3] Defined by $\|A\|_1:=\mathrm{tr}(\sqrt{A^\dagger A})$, where $A^\dagger$ denotes the Hermitian transpose.

and independent from $\mathcal{E}$ if $\rho_{X\mathcal{E}} = \rho_U \otimes \rho_{\mathcal{E}}$, where ρ_U is the fully mixed state on $\mathcal{H}_X$ (i.e., U is classical and, as a random variable, uniformly distributed).

Definition 1. The *distance to uniform* of X given $\mathcal{E}$ is $d(X|\mathcal{E}) := \frac{1}{2}\|\rho_{X\mathcal{E}} - \rho_U \otimes \rho_{\mathcal{E}}\|_1$.

If Ω is some event, determined by the random variable X, then $d(X \mid \mathcal{E},\Omega)$ is naturally defined by means of replacing the distribution P_X by $P_{X|\Omega}$. The same applies to the next two definitions.

Definition 2. The *guessing probability* of X given $\mathcal{E}$ is

$$\mathrm{Guess}(X \mid \mathcal{E}) := \sup_{\{M_x\}_x} \sum_x P_X(x)\mathrm{tr}\big(M_x\, \rho_{\mathcal{E}}^x\big),$$

where the supremum is over all positive operator-valued measures $\{M_x\}_x$ on $\mathcal{H}_{\mathcal{E}}$.

Definition 3. The *min-entropy* of X given $\mathcal{E}$ is

$$H_{\min}(X \mid \mathcal{E}) := -\log_2 \mathrm{Guess}(X \mid \mathcal{E}).$$

This definition was shown in [10] to coincide with the definition originally introduced by Renner [11], which also coincides with the classical definition of conditional min-entropy in the case where $\mathcal{E}$ is classical.

Definition 4. A function $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^\xi$ is a $(k,\varepsilon_{\mathrm{ext}})$-*strong extractor* if, for any bipartite quantum system $X\mathcal{E}$ with classical X and with $H_{\min}(X \mid \mathcal{E}) \geqslant k$ and for a uniform and independent seed Y, we have $d(\mathrm{Ext}(X,Y) \,\big|\, Y\mathcal{E}) \leqslant \varepsilon_{\mathrm{ext}}$.

Note that we find "extractor against quantum adversaries" too cumbersome a term; thus we just call Ext a (strong) extractor, even though it is a stronger notion than the standard notion of a (strong) extractor.

C. Bell inequality and CHSH

For given finite sets $\mathbb{A},\mathbb{B},\mathbb{X},\mathbb{Y}$, consider a conditional probability distribution $P_{AB|XY}$, specified as follows. There exists $\rho_{\mathcal{AB}} \in D(\mathcal{H}_{\mathcal{A}}\otimes\mathcal{H}_{\mathcal{B}})$ for an arbitrary (finite) dimensional two-partite quantum system $\mathcal{AB}$ and families of measurements $\{M_x^a\}$ and $\{N_y^b\}$, indexed by $x \in \mathbb{X}$ and $y \in \mathbb{Y}$, acting on $\mathcal{A}$ and $\mathcal{B}$, and with measurement outcomes $a \in \mathbb{A}$ and $b \in \mathbb{B}$, respectively, such that

$$P_{AB|XY}(a,b \mid x,y) = \mathrm{tr}\big[\big(M_x^a \otimes N_y^b\big)\, \rho_{\mathcal{AB}}\, \big(M_x^a \otimes N_y^b\big)^\dagger\big]$$

for all $(a,b,x,y) \in \mathbb{A} \times \mathbb{B} \times \mathbb{X} \times \mathbb{Y}$.

Definition 5. For any set $\mathcal{C} = \{c_{abxy}\}$ of *Bell coefficients*, the *Bell value* of $P_{AB|XY}$ (with respect to $\mathcal{C}$) is defined as

$$I(P_{AB|XY}) = \sum_{abxy} c_{abxy} P_{AB|XY}(a,b \mid x,y).$$

$P_{AB|XY}$ is called *classical* (or *local*) if there exist (conditional) probability distributions P_R, $P_{A|XR}$, and $P_{B|YR}$ such that

$$P_{AB|XY}(a,b \mid x,y) = \sum_r P_R(r) P_{A|XR}(a \mid x,r) P_{B|YR}(b \mid y,r)$$

for all a,b,x,y; this is equivalent to requiring that $P_{AB|XY}$ can be specified by means of a *separable state* $\rho_{\mathcal{AB}}$. We let I_0 denote the maximal Bell value achievable (for a given set of Bell coefficients) with a classical $P_{AB|XY}$. We speak of a *violation of Bell inequality* if there exists a quantum system resulting in conditional probability distribution with a Bell value greater than I_0. For instance, for so-called CHSH Bell coefficients [3], given by $c_{abxy} = (-1)^{xy}(-1)^{a\oplus b}$ for $a,b,x,y \in \{0,1\}$, it is known that $I_0 = 2$, but $I = 2\sqrt{2}$ is possible for a quantum system.

III. FRESH RANDOMNESS FROM UNTRUSTED DEVICES

In this section, we recall (some of) the findings of [6] and also discuss and fix some subtle issue not discussed there. Throughout this and the following sections, we consider fixed finite sets $\mathbb{A},\mathbb{B},\mathbb{X},\mathbb{Y}$ and a fixed set $\mathcal{C} = \{c_{abxy}\}$ of Bell coefficients. The reader may think of CHSH, but our results hold generally.

A. A single interaction

We consider an untrusted device $\mathfrak{D}$, prepared by an adversary Eve. As discussed in the Introduction, $\mathfrak{D}$ consists of two components, which, on respective inputs $x \in \mathbb{X}$ and $y \in \mathbb{Y}$, produce respective outputs $a \in \mathbb{A}$ and $b \in \mathbb{B}$ *without communicating*.[4] Formally, $\mathfrak{D}$'s behavior is given by an unknown conditional probability distribution $P_{AB|XY}$, which is specified by an unknown quantum state $\rho_{\mathcal{AB}} \in D(\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}})$ of unknown dimension and unknown families of measurements $\{M_x^a\}$ and $\{N_y^b\}$, acting on the respective systems $\mathcal{A}$ and $\mathcal{B}$. We are interested in the guaranteed amount of uncertainty in A and B (conditioned on X and Y) under the promise that $P_{AB|XY}$ has some given Bell value greater than I_0. This motivates the following definition.

Definition 6. For a given set of Bell coefficients, we define h_0 to be the function

$$h_0(I) = \inf_{\substack{\mathcal{H}_{\mathcal{A}},\mathcal{H}_{\mathcal{B}},\rho_{\mathcal{AB}} \\ \{M_x^a\},\{N_y^b\}}} \min_{\substack{x\in\mathbb{X} \\ y\in\mathbb{Y}}} H_{\min}(AB \mid X{=}x, Y{=}y),$$

where the outer infimum is over all finite-dimensional Hilbert spaces $\mathcal{H}_{\mathcal{A}}$ and $\mathcal{H}_{\mathcal{B}}$, all states $\rho_{\mathcal{AB}} \in D(\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}})$, and all families of measurements $\{M_x^a\}$ and $\{N_y^b\}$ such that $P_{AB|XY}(a,b \mid x,y) = \mathrm{tr}(M_x^a \otimes M_y^b\, \rho_{\mathcal{AB}}\, M_x^{a\dagger} \otimes M_y^{b\dagger})$ has a Bell value of at least I. Also, we define h to be the *convex closure* of h_0, i.e., the maximal convex function that does not exceed h_0.[5]

Pironio *et al.* [6] show that by means of a hierarchy of semidefinite programs (SDPs) [12,13], $h_0(I)$ can be numerically computed up to arbitrary precision (by means of a possibly expensive computation). They also show an analytical lower bound of $1 - \log_2(1 + \sqrt{2 - I^2/4})$ for $h_0(I)$ in the case of CHSH, which reaches 1 for $I = I_{\max} = 2\sqrt{2}$ [whereas the numerical calculation gives $h_0(2\sqrt{2}) \approx 1.23$] and monotonically decreases to 0 as I goes down to $I_0 = 2$; see Fig. 2 in [6]. Since this lower bound is convex, it is also a

[4]The results derived here apply also to devices with three or more components, including the three-component devices used in [5].

[5]Formally, $h(I) = \max f(I)$, where the maximum is over all convex functions f, which are upper bounded by h_0.

lower bound on h; we will need this later on.[6] For now, we can conclude that if an unknown bipartite quantum system (with fixed measurements $\{M_x^a\}$ and $\{N_y^b\}$) is promised to have a CHSH value of $I = 2\sqrt{2}$, then the joint min-entropy in the measurement outcomes A and B is lower bounded by approximately 1.23 bits (one bit if one wants to rely on the analytical bound).

B. Sequential repetitions

In order to get more uncertainty and in order to be able to *estimate* the Bell value, we consider a sequential repetition of extracting uncertainty from an untrusted device $\mathfrak{D}$ as above. Informally, rather than interacting with $\mathfrak{D}$ once [i.e., inputting $(x,y) \in \mathbb{X} \times \mathbb{Y}$ and observing $(a,b) \in \mathbb{A} \times \mathbb{B}$], we interact with $\mathfrak{D}$ n times in sequence by inputting $(x_1,y_1) \in \mathbb{X} \times \mathbb{Y}$ and observing $(a_1,b_1) \in \mathbb{A} \times \mathbb{B}$, inputting $(x_2,y_2) \in \mathbb{X} \times \mathbb{Y}$ and observing $(a_2,b_2) \in \mathbb{A} \times \mathbb{B}$, etc. This procedure is formalized as follows.

a. Modeling. We consider an arbitrary but fixed bipartite state $\rho_{\mathcal{AB}} \in D(\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B})$ of an arbitrary finite-dimensional bipartite quantum system $\mathcal{AB}$ and a sequence of n arbitrary but fixed pairs of families of measurements $(\{M_{x_1}^{a_1}\},\{N_{y_1}^{b_1}\}), \ldots, (\{M_{x_n}^{a_n}\},\{N_{y_n}^{b_n}\})$.

For each pair, $\{M_{x_j}^{a_j}\}$ is a family of measurements, indexed by $x_j \in \mathbb{X}$, acting on $\mathcal{A}$, with measurement outcomes $a_j \in \mathbb{A}$; similarly, $\{N_{y_j}^{b_j}\}$ is a family of measurements acting on $\mathcal{B}$. We allow the two components of the device $\mathfrak{D}$ to communicate *between* rounds; this is captured by a sequence $U_2, \ldots, U_n$ of unitary transformations acting on $\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$, where U_j is applied to the (collapsed) state before the jth interaction. For $j \in \{1, \ldots, n\}$, denote with a^j the concatenation of the first j rounds $a^j = a_1 \cdots a_j$ and do the same for b, x, and y. Let A^j, B^j, X^j, Y^j be the corresponding random variables. To ease notation, we use bold letters as shortcuts for the concatenation of all n rounds, e.g., $\boldsymbol{a} = a^n, \boldsymbol{A} = A^n$, etc.

Formally, the conditional probability distribution $P_{\boldsymbol{AB}|\boldsymbol{XY}}$ is defined as

$$P_{\boldsymbol{AB}|\boldsymbol{XY}}(\boldsymbol{a},\boldsymbol{b} \mid \boldsymbol{x},\boldsymbol{y}) = \prod_{j=1}^{n} P_{A_jB_j|X_jY_jT_j}(a_j,b_j \mid x_j,y_j,t_j), \quad (1)$$

where $T_j = (A^{j-1}, B^{j-1}, X^{j-1}, Y^{j-1})$ and $t_j = (a^{j-1}, b^{j-1}, x^{j-1}, y^{j-1})$ denote the transcripts up to round $j-1$, and

$$P_{A_jB_j|X_jY_jT_j}(a_j,b_j \mid x_j,y_j,t_j) = \mathrm{tr}\left[\left(M_{x_j}^{a_j} \otimes N_{y_j}^{b_j}\right) \rho_{\mathcal{AB}|T_j=t_j} \left(M_{x_j}^{a_j} \otimes N_{y_j}^{b_j}\right)^\dagger\right], \quad (2)$$

where $\rho_{\mathcal{AB}|T_j=t_j}$ is inductively defined for $j = 1, \ldots, n$ as follows. $\rho_{\mathcal{AB}|T_1=t_1} = \rho_{\mathcal{AB}}$, and for $1 \leqslant j < n$,

$$\rho_{\mathcal{AB}|T_{j+1}=t_{j+1}} = U_{j+1} \frac{\left(M_{x_j}^{a_j} \otimes N_{y_j}^{b_j}\right) \rho_{\mathcal{AB}|T_j=t_j} \left(M_{x_j}^{a_j} \otimes N_{y_j}^{b_j}\right)^\dagger}{P_{A_jB_j|X_jY_jT_j}(a_j,b_j \mid x_j,y_j,t_j)} U_{j+1}^\dagger \quad (3)$$

is the state obtained by applying U_{j+1} to the state to which $\rho_{\mathcal{AB}|T_j=t_j}$ collapses when $\mathcal{A}$ and $\mathcal{B}$ are measured by $\{M_{x_j}^{a_j}\}$ and $\{N_{y_j}^{b_j}\}$, respectively, and a_j and b_j are observed.

What is important to realize is that before every round j, the situation is exactly as in Sec. III A, with a fixed state $\rho_{\mathcal{AB}|T_j=t_j}$ and fixed measurements $\{M_{x_j}^{a_j}\}$ and $\{N_{y_j}^{b_j}\}$ in the device $\mathfrak{D}$, and thus $P_{A_jB_j|X_jY_jT_j}(\cdot,\cdot \mid \cdot,\cdot,t_j)$ here behaves as $P_{AB|XY}$ does in Sec. III A.

We would like to point out that there is no need to make $\{M_{x_j}^{a_j}\}$ (or $\{N_{y_j}^{b_j}\}$) dependent on previous inputs and outputs, i.e., on t_j using the above notation, because we may assume that the measurement $\{M_{x_j}^{a_j}\}$ encodes x_j and a_j into the postmeasurement state of $\mathcal{A}$ and that the subsequent unitary U_{j+1} copies this (classical) information into the state of $\mathcal{B}$. The subsequent measurements can then be control measurements, which perform a measurement depending on the history. Similarly, we may assume the $\{M_{x_j}^{a_j}\}$ to be identical for different j (and the same for $\{N_{y_j}^{b_j}\}$) since the quantum system $\mathcal{A}$ may maintain a counter that is increased by every unitary U_j, and $\{M_{x_j}^{a_j}\}$ can then be chosen as a control measurement that is controlled by the counter.[7] Given the conditional probability distribution $P_{\boldsymbol{AB}|\boldsymbol{XY}}$ as specified above, which describes the input-output behavior of the n sequential interactions with the device $\mathfrak{D}$, once a distribution $P_{\boldsymbol{XY}}$ is decided upon, which specifies how the inputs x_j and y_j are chosen in each round, the joint probability distribution $P_{\boldsymbol{ABXY}}$ is determined as $P_{\boldsymbol{ABXY}} = P_{\boldsymbol{XY}} P_{\boldsymbol{AB}|\boldsymbol{XY}}$.

b. Estimating the Bell value. Once the device $\mathfrak{D}$ is given, i.e., the state $\rho_{\mathcal{AB}}$, the measurements $(\{M_{x_1}^{a_1}\},\{N_{y_1}^{b_1}\})$, $\ldots$, $(\{M_{x_n}^{a_n}\},\{N_{y_n}^{b_n}\})$, and the unitaries $U_2, \ldots, U_n$ are fixed, $P_{A_1B_1|X_1Y_1}$ and thus the Bell value of the first round of interaction, $I_1 = I(P_{A_1B_1|X_1Y_1})$, are determined. For the other rounds, this is slightly more subtle. The reason is that the state $\rho_{\mathcal{AB}|T_2=t_2}$ before the second round, and thus the probability distribution $P_{A_2B_2|X_2Y_2,T_2=t_2}$, depends on what happened in the first round, i.e., depends on $t_2 = (a_1,b_1,x_1,y_1)$. Thus, the Bell value of the second round, $I_2 = I(P_{A_2B_2|X_2Y_2,T_2=t_2})$, is a function of t_2. Similarly, the Bell value of the jth round, $I_j = I(P_{A_jB_j|X_jY_j,T_j=t_j})$, is a function of t_j. We let

$$\bar{I} = \frac{1}{n} \sum_{j=1}^{n} I_j \quad (4)$$

be the average Bell value, averaged over n rounds, and we write $\bar{I} = \bar{I}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y})$ to make its dependency on the $\boldsymbol{a}$, $\boldsymbol{b}$, etc., explicit.[8]

Pironio *et al.* show in [6] that the average Bell value $\bar{I}$ can be estimated by analyzing the data collected over n rounds.

[6] Actually, the numerical computations for CHSH suggest that $h = h_0$; we do not know if this holds generally.

[7] These observations on the independence of the measurements on the history and the round are not crucial for our proofs; they merely simplify the notation.

[8] Actually, it only depends on $(a^{n-1}, b^{n-1}, x^{n-1}, y^{n-1})$.

Specifically, defining

$$\hat{I} = \hat{I}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}) = \frac{1}{n}\sum_{j=1}^{n}\sum_{abxy} c_{abxy}\frac{\chi(a_j=a,b_j=b,x_j=x,y_j=y)}{P_{XY}(x,y)}, \quad (5)$$

where $\chi(e)$ is the indicator of the event e [that is, $\chi(e)=1$ if the event e occurs and 0 otherwise], the following proposition holds.

Proposition 1 [6]. For $\bar{I}$ and $\hat{I}$ as above, for arbitrary but identically and independently distributed (iid) $(\boldsymbol{X},\boldsymbol{Y})$, meaning that $P_{\boldsymbol{XY}} = \prod_j P_{X_jY_j}$ with $P_{X_jY_j} = P_{XY}$ for all j, and for any $\varepsilon > 0$,

$$P[\bar{I}(\boldsymbol{A},\boldsymbol{B},\boldsymbol{X},\boldsymbol{Y}) \leqslant \hat{I}(\boldsymbol{A},\boldsymbol{B},\boldsymbol{X},\boldsymbol{Y}) - \varepsilon] \leqslant \exp\left(-\frac{\varepsilon^2 n}{2\left(\frac{c_{\max}}{p_{\min}}+I_{\max}\right)^2}\right),$$

where $I_{\max}$ is the maximal value of I achievable by means of a quantum system, $p_{\min} = \min_{x,y} P_{XY}(x,y)$, and $c_{\max} = \max\{c_{abxy}\}$.

Thus, except with small probability, the estimated value $\hat{I}$ for the average Bell value is not much smaller than the real average Bell value $\bar{I}$. For a fixed choice of Bell coefficients $\mathcal{C} = \{c_{abxy}\}$, which uniquely determines $I_{\max}$, we write $c(p_{\min}) = \frac{1}{2\ln 2}(\frac{c_{\max}}{p_{\min}} + I_{\max})^{-2}$, so that the probability in Proposition 1 can be written as $2^{-c(p_{\min})\varepsilon^2 n}$.

We stress that for Proposition 1 to hold, it is crucial that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are chosen *independently* of the internal state of $\mathfrak{D}$; this is implicit in the statement of Proposition 1 by having modeled the internal state $\rho_{\mathcal{AB}}$ of $\mathfrak{D}$ to be fixed and independent of $\boldsymbol{X}$ and $\boldsymbol{Y}$: $\rho_{\boldsymbol{XY}\mathcal{AB}} = \rho_{\boldsymbol{XY}} \otimes \rho_{\mathcal{AB}}$. Obviously, if $\mathfrak{D}$ knows $\boldsymbol{X}$ and $\boldsymbol{Y}$ in advance, then it can easily be programmed to have a large Bell value while, for instance, being classical.

c. Bounding the min-entropy. It remains to argue that if $\bar{I}$ is nontrivial, i.e., sufficiently greater than I_0, which can be learned by observing $\hat{I}$ (except with small probability), then the pair $(\boldsymbol{A},\boldsymbol{B})$ contains a linear (in n) amount of min-entropy. To this end, Pironio *et al.* show [see Eq. (A5) in [6]] that

$$P_{\boldsymbol{AB}|\boldsymbol{XY}}(\boldsymbol{a},\boldsymbol{b}\mid\boldsymbol{x},\boldsymbol{y}) \leqslant 2^{-h(\bar{I}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y})n)} \quad (6)$$

for all $\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}$. In the derivation, they use the fact that h is convex. From (6), they conclude [see Eq. (A9) in [6]] that $H_{\min}(\boldsymbol{AB}\mid\boldsymbol{X}=\boldsymbol{x},\boldsymbol{Y}=\boldsymbol{y}) \geqslant nh(\bar{I})$ and thus $\geqslant nh(\hat{I}-\varepsilon)$ except with small probability. However, this conclusion does not seem correct. What follows from (6) is that

$$H_{\min}(\boldsymbol{AB}\mid\boldsymbol{X}=\boldsymbol{x},\boldsymbol{Y}=\boldsymbol{y}) \geqslant nh\big(\bar{I}\big(a_0^n,b_0^n,\boldsymbol{x},\boldsymbol{y}\big)\big) \quad (7)$$

for the values a_0^n and b_0^n that *minimize* the right-hand side of (7), but then the right-hand side of (7) is likely to be smaller than $nh(\bar{I}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}))$ or $nh(\hat{I}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}))$ for the values $\boldsymbol{a}$ and $\boldsymbol{b}$ actually *observed*.[9]

For the remainder of this section, we propose and discuss a possible way to get a meaningful and useful statement on the min-entropy of $(\boldsymbol{A},\boldsymbol{B})$ in terms of $h(\bar{I})$ and thus of $h(\hat{I})$ except with small probability. We partition the interval $[I_0,I_{\max}] \subset \mathbb{R}$, ranging from the trivial, meaning classical, Bell value I_0 to the maximal value $I_{\max}$, into m disjoint blocks: $[I_0,I_{\max}] = \Omega_0 \cup \cdots \cup \Omega_{m-1}$, where Ω_ℓ is of the form $\Omega_\ell = [J_\ell, J_{\ell+1})$, with the exception that $\Omega_{m-1} = [J_{m-1}, I_{\max}]$ for some boundary points $I_0 = J_0 \leqslant J_1 \leqslant \cdots \leqslant J_{m-1} \leqslant I_{\max}$. The value of $m \in \mathbb{N}$ and the (possibly different) sizes of Ω_ℓ are arbitrary but fixed.

For any parameter $\varepsilon > 0$, given the random variables $\boldsymbol{A},\boldsymbol{B},\boldsymbol{X},\boldsymbol{Y}$, describing the n interactions with the device $\mathfrak{D}$, we can now define the random variable L_ε to be the unique random variable that satisfies $\hat{I}(\boldsymbol{A},\boldsymbol{B},\boldsymbol{X},\boldsymbol{Y}) - \varepsilon \in \Omega_{L_\varepsilon}$ (with natural adjustments outside of the range $[I_0,I_{\max}]$).[10]

Theorem 1. Let $(\boldsymbol{X},\boldsymbol{Y})$ be iid. Then, for any ε, $\delta > 0$, there exists a "good" event $\mathcal{G}$ with

$$P[\mathcal{G}] \geqslant 1 - m2^{-\delta n} - 3(2^{-c(p_{\min})\varepsilon^2 n}),$$

such that

$$\mathrm{Guess}(\boldsymbol{AB}\mid\boldsymbol{X}=\boldsymbol{x},\boldsymbol{Y}=\boldsymbol{y},L_\varepsilon=\ell,\mathcal{G}) \leqslant 2^{-nh(J_\ell)+\delta n+1}$$

and thus

$$H_{\min}(\boldsymbol{AB}\mid\boldsymbol{X}=\boldsymbol{x},\boldsymbol{Y}=\boldsymbol{y},L_\varepsilon=\ell,\mathcal{G}) \geqslant nh(J_\ell) - \delta n - 1$$

for all $\boldsymbol{x}\in\mathbb{X}^n$, $\boldsymbol{y}\in\mathbb{Y}^n$, and $\ell\in\{0,\ldots,m-1\}$ with $P_{\boldsymbol{XY}L_\varepsilon|\mathcal{G}}(\boldsymbol{x},\boldsymbol{y},\ell) > 0$.

We would like to point out that for the bound on $P[\mathcal{G}]$ to hold, it is crucial that $\rho_{\mathcal{AB}}$ is independent of $(\boldsymbol{X},\boldsymbol{Y})$ [and (X_i,Y_i) are iid]: clearly, the device can fool you if it knows the inputs it will get in advance. However, for event $\mathcal{G}$ as defined in the proof below, the bound on the guessing probability holds *irrespective* of the distribution of $\boldsymbol{X}$ and $\boldsymbol{Y}$. Indeed, the value of $\mathrm{Guess}(\boldsymbol{AB}\mid\boldsymbol{X}=\boldsymbol{x},\boldsymbol{Y}=\boldsymbol{y},L_\varepsilon=\ell,\mathcal{G})$ is determined by the conditional probability distribution $P_{\boldsymbol{AB}|\boldsymbol{XY}}(\cdot,\cdot\,|\boldsymbol{x},\boldsymbol{y})$ alone (which is determined by $\rho_{\mathcal{AB}}$, the family of measurements and the unitaries); this holds because L_ε as well as $\mathcal{G}$ (we will see this below) are uniquely determined by $\boldsymbol{A},\boldsymbol{B},\boldsymbol{X}$, and $\boldsymbol{Y}$.

Proof. Let $\mathcal{B}^{\mathrm{guess}}$ be the bad event $\bar{I}(\boldsymbol{A},\boldsymbol{B},\boldsymbol{X},\boldsymbol{Y}) \leqslant \hat{I}(\boldsymbol{A},\boldsymbol{B},\boldsymbol{X},\boldsymbol{Y}) - \varepsilon$ where the estimated Bell value $\hat{I}$ is significantly larger than the average Bell value $\bar{I}$, and let $\mathcal{G}^{\mathrm{guess}}$ be its complement (which we understand as a good event); by Proposition 1, we know that $P[\mathcal{B}^{\mathrm{guess}}] \leqslant 2^{-c(p_{\min})\varepsilon^2 n}$. We define $\mathcal{B}_1$ to be the set of all "bad inputs" $(\boldsymbol{x},\boldsymbol{y})$ with the property

$$P[\mathcal{B}^{\mathrm{guess}}|\boldsymbol{X}=\boldsymbol{x},\boldsymbol{Y}=\boldsymbol{y}] \geqslant \tfrac{1}{2}; \quad (8)$$

it is straightforward to show that $P[(\boldsymbol{X},\boldsymbol{Y})\in\mathcal{B}_1] \leqslant 2(2^{-c(p_{\min})\varepsilon^2 n})$. Finally, we define $\mathcal{B}_2$ to be the set of all $(\boldsymbol{x},\boldsymbol{y},\ell)$ with the property

$$P_{L_\varepsilon|\boldsymbol{XY}\mathcal{G}^{\mathrm{guess}}}(\ell|\boldsymbol{x},\boldsymbol{y}) \leqslant 2^{-\delta n}. \quad (9)$$

It follows from the definition of $\mathcal{B}_2$ that

$$P[(\boldsymbol{X},\boldsymbol{Y},L_\varepsilon)\in\mathcal{B}_2|\mathcal{G}^{\mathrm{guess}}] \leqslant m2^{-\delta n}.$$

We slightly abuse notation and identify the *set* $\mathcal{B}_1$ with the bad *event* $(\boldsymbol{X},\boldsymbol{Y})\in\mathcal{B}_1$, and we write $\mathcal{G}_1$ for its complementary

[9] We note that the authors of [6], in a work independent of ours [8], fix this issue in a similar manner as we do here.

[10] The definition of L_ε simply captures that if $\hat{I}$ is too close to the lower end of an interval, then we take the next lower interval to be on the safe side.

good event and correspondingly for $\mathcal{B}_2$ and $\mathcal{G}_2$. We now define the good event $\mathcal{G}$ as $\mathcal{G} := \mathcal{G}^{\text{guess}} \wedge \mathcal{G}_1 \wedge \mathcal{G}_2$. Using the union bound over the bad events, it is not too hard to show that $P[\mathcal{G}] \geqslant 1 - m2^{-\delta n} - 3(2^{-c(p_{\min})\varepsilon^2 n})$.

It remains to argue the bound on the min-entropy. Let $\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}$ be such that $\bar{I}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}) > \hat{I}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}) - \varepsilon$; i.e., they have positive probability conditioned on the good event $\mathcal{G}^{\text{guess}}$. Furthermore, let ℓ be the unique value with

$$\hat{I}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}) - \varepsilon \in \Omega_\ell.$$

If $(\boldsymbol{x},\boldsymbol{y}) \notin \mathcal{B}_1$, then $P[\mathcal{G}^{\text{guess}} | \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Y} = \boldsymbol{y}] \geqslant \frac{1}{2}$, and hence, conditioning on the event $\mathcal{G}^{\text{guess}}$ can increase the probabilities by at most a factor of 2. For those $(\boldsymbol{x},\boldsymbol{y}) \notin \mathcal{B}_1$, it then follows from (6) that

$$\begin{aligned} P_{\boldsymbol{AB}|\boldsymbol{XY},\mathcal{G}^{\text{guess}}}(\boldsymbol{a},\boldsymbol{b} \mid \boldsymbol{x},\boldsymbol{y}) &\leqslant 2(2^{-nh(\bar{I}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}))}) \\ &\leqslant 2(2^{-nh(\hat{I}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y})-\varepsilon)}) \\ &\leqslant 2(2^{-nh(J_\ell)}). \end{aligned} \tag{10}$$

If additionally we have $(\boldsymbol{x},\boldsymbol{y},\ell) \notin \mathcal{B}_2$, then

$$\begin{aligned} P_{\boldsymbol{AB}|\boldsymbol{XY}L_\varepsilon,\mathcal{G}^{\text{guess}}}(\boldsymbol{a},\boldsymbol{b} \mid \boldsymbol{x},\boldsymbol{y},\ell) &\leqslant \frac{P_{\boldsymbol{AB}|\boldsymbol{XY},\mathcal{G}^{\text{guess}}}(\boldsymbol{a},\boldsymbol{b} \mid \boldsymbol{x},\boldsymbol{y})}{P_{L_\varepsilon|\boldsymbol{XY}\mathcal{G}^{\text{guess}}}(\ell \mid \boldsymbol{x},\boldsymbol{y})} \\ &\leqslant 2(2^{-nh(J_\ell)})(2^{\delta n}). \end{aligned} \tag{11}$$

Note that additionally conditioning on $\mathcal{G}_1$ and $\mathcal{G}_2$ does not change the above conditional probability distribution if $(\boldsymbol{x},\boldsymbol{y}) \notin \mathcal{B}_1$ and $(\boldsymbol{x},\boldsymbol{y},\ell) \notin \mathcal{B}_2$. Thus, the same bound also applies to $P_{\boldsymbol{AB}|\boldsymbol{XY}L_\varepsilon,\mathcal{G}}(\boldsymbol{a},\boldsymbol{b} \mid \boldsymbol{x},\boldsymbol{y},\ell)$ for all $\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}$, and ℓ with $P_{\boldsymbol{ABXY}L_\varepsilon|\mathcal{G}}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y},\ell) > 0$. By definition of the guessing probability and the min-entropy, this proves the claim. ■

d. A specific example. Consider CHSH, so that the Bell value of a given device is expected to be in the range from $I_0 = 2$ to $I_{\max} = 2\sqrt{2} \approx 2.828$. Let us divide this range into $J_0 = I_0 < J_1 = 2.2 < J_2 = 2.4 < J_3 = 2.6 < I_{\max}$, and let us take a q-biased input distribution $P_{\boldsymbol{XY}} = \prod_j P_{X_jY_j}$ with $P_{X_jY_j}(0,0) = 1 - 3q$ and $P_{X_jY_j}(x,y) = q$ for all $(x,y) \in \{0,1\}^2 \setminus \{(0,0)\}$, where $0 < q \leqslant 1/4$ is some parameter. Finally, let us fix some small parameters $\varepsilon,\delta > 0$; for concreteness, say that $\varepsilon = 0.05$ and $\delta = 0.01$.

Consider now n sequential interactions with an untrusted device $\mathfrak{D}$, where in each round x_j and y_j are chosen (according to $P_{X_jY_j}$) and input into $\mathfrak{D}$, and a_j and b_j are obtained as output from $\mathfrak{D}$. Let us say that from the collected data, we get $\hat{I}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}) = 2.7 \in \Omega_3$ as the estimation for the average Bell value. By Theorem 1, we have that, given $\boldsymbol{x}$ and $\boldsymbol{y}$ and $L_\varepsilon = 3$, the min-entropy of $\boldsymbol{a}$ and $\boldsymbol{b}$ is at least $n[h(2.6) - \delta] - 1 \approx n(0.36 - \delta) > n/3$ bits, except with probability $4(2^{-\delta n}) + 3(2^{-c(q)\varepsilon^2 n})$.[11] Thus, when applying a suitable randomness extractor to $\boldsymbol{a},\boldsymbol{b}$, we can extract, say, $n/4$ bits that are exponentially close to uniformly distributed (given $\boldsymbol{x}$ and $\boldsymbol{y}$ and $L_\varepsilon = 3$).

In order to sample the inputs according to the biased input distribution $P_{\boldsymbol{XY}}$, as suggested in [6], it is known to be sufficient (on average) to have access to $nO(q \log_2(1/q))$ random bits [14]. Since $q \log_2(1/q)$ converges to 0 for $q \to 0$, if q is chosen to be a small enough constant, then, say, $n/4$ random bits are sufficient. Thus, by starting off with $n/4$ random bits, we obtained another $n/4$ almost-random bits and thus hold now $n/2$ random bits.[12] Thus, we have expanded the randomness by a factor 2. Choosing $q = O(n^{-1/3})$, one obtains an expansion factor $O(n^{2/3}/\log_2 n)$ while still being negligibly close to perfect randomness [since $c(n^{-1/3}) = \Omega(n^{-2/3})$].

Having generated fresh randomness from an untrusted device $\mathfrak{D}$, one is now tempted to use the newly obtained randomness to generate even more fresh randomness from the device $\mathfrak{D}$ and so on. This does not work. The reason is that the generated randomness is not random *to the device* $\mathfrak{D}$, or, more formally, not independent of the internal state of $\mathfrak{D}$; indeed, $\mathfrak{D}$ has already observed $\boldsymbol{x}$ and $\boldsymbol{y}$, and it has itself produced $\boldsymbol{a}$ and $\boldsymbol{b}$. We argue below, however, that we can use the fresh randomness to generate even more randomness from *another* device, as long as the devices are not entangled with each other or with the adversary.

e. Classical side information. The case where the adversarial producer of the devices Eve holds classical side information about the device $\mathfrak{D}$ can be reduced to the case without side information by conditioning on particular values of the side information.

f. Quantum side information? Ideally, one would like to obtain similar results in the case where Eve holds *quantum* side information about the device $\mathfrak{D}$, i.e., where Eve maintains a quantum state $\mathcal{E}$ that is entangled with state $\mathcal{AB}$ contained in $\mathfrak{D}$. It is not too hard to see that if the function $2^{-h(\cdot)}$ is concave (which is, e.g., satisfied for CHSH), then the min-entropy bound $H_{\min}(AB|X = x, Y = y) \geqslant h(I)$ for a *single* interaction, which holds by definition of the function h, extends to the case of quantum side information in that also $H_{\min}(AB|X = x, Y = y, \mathcal{E}) \geqslant h(I)$ holds, where I is the Bell value of the distribution obtained by tracing out $\mathcal{E}$. This seems to suggest that quantum side information is useless for Eve.

Unfortunately, the techniques we use for analyzing sequential repetitions do not appear applicable in the case of quantum side information. The main technical problem we run into is to show that a lower bound on the *smooth* min-entropy of a random variable conditioned on a quantum state can be obtained from a lower bound on the *smooth* min-entropy of the random variable conditioned on the *measurement outcome* for every possible measurement (as is the case for the nonsmooth version [10]).[13] It is not clear if such a statement is actually true.

Thus, whether quantum side information is useless for Eve also in the case of sequential repetition, where the average Bell value is estimated from the observed data, remains an open problem. We refer to [7] for a different proof technique which does work against quantum side information.

[11] This probability is an average value over $\boldsymbol{A}\ \boldsymbol{B}$, $\boldsymbol{X}\ \boldsymbol{Y}$.

[12] We are ignoring here the randomness needed for the extractor.

[13] Essentially, the notion of *smooth* min-entropy takes into account that we may condition on a good event that has a probability close to 1.

Let $\varepsilon, \delta, e > 0$ be the security parameters. Let $t > 0$ and assume long enough initial randomness $K_{\text{init}} = (K_0, K_{\text{Bank}})$, split into parts of size $|K_0| = \eta$ and $|K_{\text{Bank}}| = O(t\eta^2 \log_2 \eta)$. Denote the second part by $K_{\text{Bank}} = S_1 \cdots S_t$ where $|S_1| = \cdots = |S_t| = O(\eta^2 \log_2 \eta)$.

1. For $i = 1$ to t do:
   (a) Set $R_i = \text{Sample}\,(K_{i-1})$.
   (b) Input R_i to $\mathfrak{D}_{(i \mod 2)}$ and denote the output by W_i.
   (c) Estimate the Bell value $\hat{I}_i$ of the ith interaction with the devices [Eq. (5)], and determine the unique value ℓ_i that satisfies $\hat{I}_i - \varepsilon \in \Omega_{\ell_i}$
   (d) Calculate the guaranteed min-entropy in W_i, $\text{H}^i_{\min} := \|W_i\|(\text{h}(J_{\ell_i}) - \delta) - 1$, and let $\xi_i = \lfloor \text{H}^i_{\min} - 4\log_2 \varepsilon_{\text{ext}}^{-1} - c_{\text{ext}} \rfloor$, with a constant c_{ext} specified in Lemma 1.
   (e) If $\xi_i \le \eta$, set $K_i = K_{i+1} = \ldots = K_t = \perp$ (the empty string) and go to step 2. Otherwise set $K_i = \text{Ext}\,(W_i, S_i)$ (with $\varepsilon_{\text{ext}} = 2^{-e\eta}$ and $\xi_i = |K_i|$).
2. Output K_t.

FIG. 1. Randomness expansion scheme with two independent devices.

IV. COMPOSABILITY

Consider two (or more) untrusted devices $\mathfrak{D}$ and $\mathfrak{D}'$, prepared by the adversary Eve. We assume that $\mathfrak{D}$ and $\mathfrak{D}'$ cannot communicate and are not entangled with each other. The case when Eve holds classical side information about the devices can be treated as described in the previous section. We can then apply Theorem 1 to argue that the output $\boldsymbol{AB}$ produced by $\mathfrak{D}$ has high min-entropy (except with small probability) given the internal state of $\mathfrak{D}'$ (because $\mathfrak{D}'$ is independent of $\mathfrak{D}$), assuming that a large enough average Bell value is observed. It thus follows that by applying an extractor (with suitable parameters and a freshly chosen seed) to $\boldsymbol{AB}$, we obtain a bit string K that is close to random and independent of the internal state of $\mathfrak{D}'$. This in particular implies that if we use the randomness K to sample the input $\boldsymbol{X}'\boldsymbol{Y}'$ to $\mathfrak{D}'$ (according to a prescribed distribution), then $\boldsymbol{X}'\boldsymbol{Y}'$ is close to *independent* of the internal state of $\mathfrak{D}'$. As the dependency between the internal (quantum) state of $\mathfrak{D}$ and the inputs and outputs of $\mathfrak{D}'$ is purely classical, we can condition on this classical information and apply Theorem 1 to argue that the output $\boldsymbol{A}'\boldsymbol{B}'$ produced by $\mathfrak{D}'$ has high min-entropy given the current internal state of $\mathfrak{D}$. Therefore, we are in the same situation as above and so can use the randomness extracted from $\boldsymbol{A}'\boldsymbol{B}'$ to sample again inputs for $\mathfrak{D}$, and we can keep on going like this as long as a large enough Bell value is observed. We stress that the above line of reasoning only works because we assumed the devices $\mathfrak{D}, \mathfrak{D}'$ to be unentangled to start with.

A. Randomness expansion with independent devices

We increase the level of abstraction and from now on consider an untrusted device $\mathfrak{D}$ as an abstract object that takes some input $R = (\boldsymbol{X}, \boldsymbol{Y}) = (X^n, Y^n)$ and produces some output $W = (\boldsymbol{A}, \boldsymbol{B}) = (A^n, B^n)$ and maintains some internal quantum state $\mathcal{E}_{\mathfrak{D}}$. It is guaranteed that, if R is properly distributed and W satisfies some statistical property, then W has lower-bounded min-entropy given $R = r$ (except with small probability), even given *classical* side information. Let $|R|$ denote the length of the bit string R, that is, $|R| = n \log_2 |\mathbb{X}||\mathbb{Y}|$ (similarly, $|W| = n \log_2 |\mathbb{A}||\mathbb{B}|$). Let $\|R\| = n$ be the number of pairs (x, y) in R, that is, the number of sequential interactions with the device caused by the input R (respectively, $\|W\| = n$).

Instead of using v devices, our scheme alternates between two devices $\mathfrak{D}_0, \mathfrak{D}_1$. Formally (Fig. 1), we split the initial randomness K_{init} into two parts $K_{\text{init}} = (K_0, K_{\text{Bank}})$ such that $|K_0| = \eta$ and $|K_{\text{Bank}}| = O(t\eta^2 \log_2 \eta)$ for some constant $t > 0$. The first part K_0 is used to generate an input to the devices, and the other part K_{Bank} is used as a "bank" of t seeds for the extractor. For $i = 1, \ldots, t$, let R_i denote the input given to the device on the ith activation.[14] For the first iteration, the input is the biased string $R_1 = \text{Sample}(K_0)$. Let W_i denote the output of the ith activation. In order to generate the input for the next iteration, R_{i+1}, we use an extractor on the bit string W_i with a fresh random seed S_i from the K_{Bank}. Denote the extracted string by $K_i = \text{Ext}(W_i, S_i)$. Finally, K_i is used to generate a biased string $R_{i+1} = \text{Sample}(K_i)$, which is the input for the next iteration. Note that each iteration achieves at most quadratic expansion; hence the strings R_i, W_i, K_i are of length poly(η).

In order to extract randomness out of W_i, we use the extractor of De *et al.* [15] with $\varepsilon_{\text{ext}} = 2^{-e\eta}$, where the constant $e > 0$ is taken arbitrarily small.

Lemma 1 (Corollary 5.2 in [15]). There exists an extractor $\text{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^\xi$ which is

$$(\xi + 4\log_2 \varepsilon_{\text{ext}}^{-1} + c_{\text{ext}}\,,\ \varepsilon_{\text{ext}})$$

strong against quantum adversaries, with $c_{\text{ext}} = O(1)$ and a seed of size $d = O(\log_2^2(n/\varepsilon_{\text{ext}}) \log_2 \xi)$.

[14] We alternate between the devices; thus the ith activation is done with $\mathfrak{D}_{(i \mod 2)}$, which is also referred to as the ith device.

The seed for the extractor is of length

$$\begin{aligned} &O\left(\log_2^2(|W_i|/\varepsilon_{\text{ext}})\log_2 \xi_i\right) \\ &= O\left(\left[\log_2^2 \text{poly}(\eta) + e^2\eta^2\right]\log_2 \text{poly}(\eta)\right) \\ &= O\left(e^2\eta^2 \log_2 \eta\right). \end{aligned} \tag{12}$$

V. CONCLUSION AND OPEN PROBLEMS

An interesting extension to our result is to generalize Theorem 1 to the setting of quantum side information. This would allow a composition theorem for the more general case in which the devices can be entangled with each other and with Eve. Although we can allow quantum side information *for a single interaction* with the device, we are currently unable to give a rigorous proof of security against quantum side information for multiple interactions and leave it as the main open question.

ACKNOWLEDGMENTS

R.G. is grateful to CWI, Amsterdam, for hosting him while part of this work was done. C.S. is supported by a NWO VENI grant. This manuscript was made open-access thanks to a grant by NWO.

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[2] J. Bell, Physics **1**, 195 (1964).
[3] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[4] R. Colbeck, Ph.D. thesis, University of Cambridge, 2009.
[5] R. Colbeck and A. Kent, J. Phys. A **44**, 095305 (2011).
[6] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Nature (London) **464**, 1021 (2010).
[7] U. Vazirani and T. Vidick, Philos. Trans. R. Soc. A **370**, 3432 (2012).
[8] S. Pironio and S. Massar, Phys. Rev. A **87**, 012336 (2013).
[9] J. Barrett, R. Colbeck, and A. Kent, Phys. Rev. Lett. **110**, 010503 (2013).
[10] R. König, R. Renner, and C. Schaffner, IEEE Trans. Inf. Theory **55**, 4337 (2009).
[11] R. Renner, Ph.D. thesis, ETH Zürich, 2005.
[12] M. Navascués, S. Pironio, and A. Acín, Phys. Rev. Lett. **98**, 010401 (2007).
[13] M. Navascués, S. Pironio, and A. Acín, New J. Phys. **10**, 073013 (2008).
[14] D. Knuth and A. Yao, in *Algorithms and Complexity: New Directions and Recent Results* (Academic, New York, 1976), pp. 357–428.
[15] A. De, C. Portmann, T. Vidick, and R. Renner, SIAM J. Comput. **41**, 915 (2012).