Seedless extractors for device-independent quantum cryptography

Cameron Foreman[1,2] and Lluís Masanes[2,3]

[1]Quantinuum, Partnership House, Carlisle Place, London SW1P 1BX, United Kingdom
[2]Department of Computer Science, University College London, United Kingdom
[3]London Centre for Nanotechnology, University College London, United Kingdom

Device-independent (DI) quantum cryptography aims at providing secure cryptography with minimal trust in, or characterisation of, the underlying quantum devices. A key step in DI protocols is randomness extraction (or privacy amplification), which typically requires a *seed* of additional bits with sufficient entropy and statistical independence from any bits generated during the protocol. In this work, we propose a method for extraction in DI protocols that does not require a seed and is secure against computationally unbounded quantum adversaries. The core idea is to use the Bell violation of the raw data, rather than its min-entropy, as the extractor promise. We present a complete security proof in a model where the experiment uses memoryless measurement devices acting on an arbitrary joint (across all rounds) state. Our results mark a first step in this alternative, seedless, approach to extraction in DI protocols.

1 Introduction

Device-independent (DI) quantum cryptography has emerged as the gold standard in terms of security for cryptographic applications [1, 2]. This paradigm provides security with minimal trust in, or characterisation of, the underlying quantum devices. This is achieved by exploiting quantum non-locality, that is, correlations which violate Bell inequalities [3]. Some applications of DI quantum cryptography include key distribution [4–8], randomness expansion [9, 10], randomness amplification [11], and many more [12–14].

A crucial step in numerous DI protocols is that of *randomness extraction* (similarly *privacy amplification*), which involves generating a near-perfect random output (a *final output*) by classically processing some imperfect, somewhat random input (a *extractor input*) derived from measurement outcomes. To date, randomness extraction in DI tasks has required the consumption of a *seed* of bits that must, at a minimum, be sufficiently statistically independent of the quantum hardware [15, 16] and sufficiently random from the adversary's perspective [17]. In practice, this requirement is often more stringent, such as requiring a perfectly uniform and independent seed [18, 19].

In this work, we demonstrate that randomness extraction in DI quantum cryptography can be achieved using a deterministic algorithm (i.e., without a seed), while preserving security against computationally unbounded quantum adversaries – and, by extension, against computationally unbounded classical adversaries. The intuition behind our results

Cameron Foreman: cameron.foreman@quantinuum.com

is that the violation of Bell inequalities not only guarantees a lower bound on the min-entropy of the outcomes, but certain statistical independence between the outcomes of different rounds of the experiment. This concept can be understood from results in *self-testing*, such as those showing that maximal violation of the CHSH inequality [20] implies the measured bipartite state is essentially pure [21], indicating that there is no correlation between rounds.

Prior to this work, only the min-entropy promise has been exploited in DI protocols, which necessitates the use of randomised (seeded or multi-source) extractors. It is shown in [22] that it is impossible to deterministically extract even a single bit from an input characterised solely by min-entropy. However, our approach involves designing extractors that take advantage of the promise of Bell violation instead. This stronger promise allows us to eliminate the need for a seed. In our proofs, we focus on the CHSH inequality due to its simple experimental setup (a quantum device with two isolated parts) and widespread use in DI protocols, although our proof techniques can be readily applied to other Bell inequalities.

The results of this paper are proven for the case where the experimental state can be arbitrary (e.g., with correlations across rounds), while the two measurement devices are memoryless (or equivalently, where each round's measurement is performed using a separate, non-communicating device). Although not fully general, they mark an important step in initiating a new *seedless* approach to randomness extraction in DI quantum cryptography protocols, with numerous problems to explore (see 'Conclusion and discussion' in Section 4). From a fundamental perspective, by exploiting the full power of Bell inequality violation, we identify a new class of distributions that can be both deterministically extracted from and generated by a realisable experimental process. This contributes to a long line of research in computer science exploring deterministic randomness extraction, e.g. [23–29] (see [30] for an overview). Moreover, we observe that arbitrary violation of the CHSH Bell inequality is sufficient for seedless extraction – which possibly speaks to the power of Bell non-locality as a resource [31–35].

It is important to note that DI protocols include a step to test the degree of Bell violation, which requires random numbers to choose the measurement settings in each round. Additionally, several other subroutines rely on randomness; for example, key distribution protocols need preshared randomness for public channel authentication. Therefore, the seedless extraction we present in this work does not completely eliminate the need for initial randomness. However, we expect that the initial randomness required for a Bell test must satisfy weaker statistical conditions than that for both a Bell test and seeded extraction. Thus, we hope that future contributions using the techniques of this work will improve the capabilities of DI protocols, particularly randomness amplification [33, 36–40].

In Section 2 we review the relationship between Bell-inequality violation and randomness, we introduce the notation and main definitions of the paper, and prove our two theorems on seedless extraction. In Section 3 we apply these results to a spot-checking based DI protocol, in which we show that our seedless extractors allow for extraction in the case of arbitrary low CHSH violation, and have unity rate in the limit of maximal CHSH violation. In Section 4 we conclude and list the most important open problems on DI seedless extractors. The Appendix contains the proof of two lemmas and a theorem.

2 Device-independent seedless extractors

2.1 Randomness and violation of Bell inequalities

Suppose Alice has a quantum system with Hilbert space $\mathcal{A}$, which can be measured with two observables labelled $x \in \{0,1\}$ with outcomes $a \in \{0,1\}$ represented by the positive operator-valued measure (POVM) elements $A(a|x)$. Analogously, Bob has a system $\mathcal{B}$ and two observables $y \in \{0,1\}$ with outcomes $b \in \{0,1\}$ and POVMs $B(b|y)$. The joint state of $\mathcal{A} \otimes \mathcal{B}$ is denoted by $\rho_{\mathcal{AB}}$. With this notation, we can write the CHSH inequality [20] as

$$\text{CHSH} = \sum_{x,y,a,b=0}^{1} \text{tr}[\rho_{\mathcal{AB}} A(a|x) B(b|y)] \, (-1)^{a+b+xy} \leq 2 \ . \tag{1}$$

If this inequality is violated then no locally causal model can explain the observed correlations [20]. Note that, in our notation, the operator $A(a|x)$ is meant to act trivially on $\mathcal{B}$, so we can write $A(a|x)B(b|y)$ instead of $A(a|x) \otimes B(b|y)$.

The predictability of outcome a when measuring x can be quantified by the bias of the probability distribution of a in the following sense

$$|\text{P}(a=0|x) - \text{P}(a=1|x)| = |\, \text{tr}(\rho_{\mathcal{AB}} \, [A(0|x) - A(1|x)]) \,| \ . \tag{2}$$

The following theorem proven in [41] tells us that the stronger the CHSH violation, the less predictable the outcome a.

Theorem 1. For any pair of Hilbert spaces $\mathcal{A}, \mathcal{B}$ and measurements $\{A(0|x), A(1|x)\}$ on $\mathcal{A}$ and $\{B(0|y), B(1|y)\}$ on $\mathcal{B}$, we define the shifted CHSH operator as

$$S = \mu_s \mathbb{1} - \nu_s \sum_{x,y,a,b=0}^{1} A(a|x) B(b|y) (-1)^{a+b+xy} \ , \tag{3}$$

with coefficients

$$\mu_s = 2 \left(2 - \frac{s^2}{4}\right)^{-1/2} , \qquad \nu_s = \frac{s}{4} \left(2 - \frac{s^2}{4}\right)^{-1/2} \ . \tag{4}$$

The following two semi-definite inequalities

$$\pm \, [A(0|0) - A(1|0)] \leq S \ , \tag{5}$$

hold for all $s \in (2, 2\sqrt{2})$.

The shifted CHSH operator (3) includes an identity term with a positive coefficient and the CHSH operator from (1) with a negative coefficient. This specific parametrisation, defined using s, is known to yield a tight family of inequalities when s corresponds to the CHSH violation. Notably, inequality (5) indicates that greater CHSH violation corresponds to reduced predictability, since

$$|\text{tr}(\rho_{\mathcal{AB}} \, [A(0|0) - A(1|0)])| \leq \text{tr}(\rho_{\mathcal{AB}} S) \ . \tag{6}$$

This fact, along with its generalisation to other Bell inequalities, is central to DI quantum cryptography and crucial for the results in this section. Our proofs rely on having a family of semi-definite inequalities which relate the predictability of generated outcomes to Bell inequality violation. Since Equation (6) is known to be tight and Equation (3) has a simple form, we focus on the CHSH inequality when presenting our results. However, our proof can be easily adapted to other Bell inequalities.

2.2 Security in the presence of a quantum adversary

The process for generating the extractor input $\mathbf{a} = (a_1, \ldots, a_{n_r})$ consists of n_r rounds labelled by $i \in \{1, \ldots, n_r\}$. In round i Alice performs the measurement $\{A_i(0|0), A_i(1|0)\}$ on the system with Hilbert space $\mathcal{A}_i$, and obtains the outcome a_i. Specifically, this means that when Alice is generating the extractor input, she uses the input $x_i = 0$ in every round. Similarly, in round i Bob has a system with Hilbert space $\mathcal{B}_i$, but our protocol does not require Bob to make measurements for extractor input generation. The adversary (Eve) holds an arbitrary quantum system with Hilbert space $\mathcal{E}$. We use the notation $\mathcal{A}_r = \bigotimes_{i=1}^{n_r} \mathcal{A}_i$ and $\mathcal{B}_r = \bigotimes_{i=1}^{n_r} \mathcal{B}_i$ and understand that the action of $A_i(a|x)$ on $\mathcal{A} \otimes \mathcal{B}$ is trivial on all factors but $\mathcal{A}_i$. The factorisation of the total Hilbert space $\mathcal{A}_r \otimes \mathcal{B}_r \otimes \mathcal{E}$ enforces the assumption of no-signalling between Alice, Bob and Eve. Additionally, the fact that we model every round i with a different Hilbert space $\mathcal{A}_i \otimes \mathcal{B}_i$ enforces the assumption that devices have no memory.

The global state shared among Alice, Bob and Eve is $\rho_{\mathcal{A}_r \mathcal{B}_r \mathcal{E}}$, and the reduced state of Alice and Eve $\rho_{\mathcal{A}_r \mathcal{E}}$. The final output $\mathbf{k} = (k_1, \ldots, k_m) \in \{0,1\}^m$ is produced by applying the (deterministic) function $g : \{0,1\}^{n_r} \to \{0,1\}^m$ to the extractor input $\mathbf{a} \to \mathbf{k} = g(\mathbf{a})$. In the following sections we characterise the functions g. Although the final output $\mathbf{k}$ is a classical system, it is convenient to associate to it a Hilbert space $\mathcal{K} = \mathbb{C}^{2^m}$ and to represent its values by an orthonormal basis $|\mathbf{k}\rangle \in \mathcal{K}$. After Alice measures all her systems $\mathcal{A}_i$ for $i \in \{1, \ldots, n_r\}$ and generates the final output on system $\mathcal{K}$, the joint state of systems $\mathcal{K} \otimes \mathcal{E}$ is

$$\rho_{\mathcal{K}\mathcal{E}} = \sum_{\mathbf{k} \in \{0,1\}^m} \sum_{\mathbf{a} \in \{0,1\}^{n_r}} |\mathbf{k}\rangle\langle\mathbf{k}|_{\mathcal{K}} \, \delta^{\mathbf{k}}_{g(\mathbf{a})} \, \mathrm{tr}_{\mathcal{A}_r} \left[\rho_{\mathcal{A}_r \mathcal{E}} \prod_{i=1}^{n_r} A_i(a_i|0) \right] . \tag{7}$$

We include the labels $\mathcal{K}, \mathcal{E}, \mathcal{A}_r$ to the operators to specify on which Hilbert space they act, hence we don't need to use the symbol $\otimes$. The goal of Alice and Bob is to produce a state $\rho_{\mathcal{K}\mathcal{E}}$ that is indistinguishable from an ideal final output $u_{\mathcal{K}} \rho_{\mathcal{E}}$, where we define the uniform state (sometimes called maximally mixed) as

$$u_{\mathcal{K}} = 2^{-m} \sum_{\mathbf{k} \in \{0,1\}^m} |\mathbf{k}\rangle\langle\mathbf{k}|_{\mathcal{K}} \ . \tag{8}$$

This indistinguishability can be formalised as a bound on the trace norm

$$\| \rho_{\mathcal{K}\mathcal{E}} - u_{\mathcal{K}} \rho_{\mathcal{E}} \|_1 \leq \epsilon \ , \tag{9}$$

which is defined as $\|M\|_1 = \mathrm{tr} \sqrt{M^\dagger M}$ for any operator M. The bound in Equation (9) implies that any cryptographic task which requires an ideal final output $u_{\mathcal{K}} \rho_{\mathcal{E}}$ as a resource, is also secure if performed using the real final output $\rho_{\mathcal{K}\mathcal{E}}$, up to an error of probability ϵ. Hence, we say that our seedless extraction protocol can be composed with any other cryptographic protocol: it is universally composable [42].

2.3 XOR is a seedless extractor

The XOR function allows to extract a single bit $k \in \{0,1\}$ with an error that can be made exponentially small in n_r. The computational cost of XOR is $O(n_r)$, showing that seedless extractors do not need to be computationally hard to implement.

Theorem 2 (XOR)**.** After measuring the n_{r}-round state $\rho_{\mathcal{A}_{\mathsf{r}}\mathcal{B}_{\mathsf{r}}\mathcal{E}}$ with the observables $\{A_i(a_i|0) : a_i = 0, 1\}$ for all rounds $i = 1, \ldots, n_{\mathsf{r}}$ and applying the XOR function

$$g(\mathbf{a}) = \sum_{i=1}^{n_{\mathsf{r}}} a_i \bmod 2 \in \{0, 1\} \tag{10}$$

to the outcomes $k = g(a_1, \ldots, a_{n_{\mathsf{r}}})$, the resulting state $\rho_{\mathcal{K}\mathcal{E}}$ written in Equation (7) satisfies

$$\|\rho_{\mathcal{K}\mathcal{E}} - u_{\mathcal{K}}\rho_{\mathcal{E}}\|_1 \leq \operatorname{tr}\left[\rho_{\mathcal{A}_{\mathsf{r}}\mathcal{B}_{\mathsf{r}}} \prod_{i=1}^{n_{\mathsf{r}}} S_i\right] , \tag{11}$$

for all $s \in [2, 2\sqrt{2}]$.

The above result shows that the larger the violation of CHSH (i.e. the smaller the expectation of $\prod_i S_i$), the smaller the distance between the real and the ideal final outputs.

Proof. We start by defining the operator

$$C_i = A_i(0|0) - A_i(1|0) , \tag{12}$$

and noting that

$$A_i(a_i|0) = \frac{1}{2}\left(\mathbb{1} + (-1)^{a_i} C_i\right) . \tag{13}$$

As proven in [43], there is no loss of generality in assuming that the operators $A_i(a_i|0)$ are projectors, which implies that C_i is full-rank.

Next, we substitute Equation (13) into the joint state after Alice generates the final output (7) and expand the product $\prod_i (\mathbb{1} + (-1)^{a_i} C_i)$ into $2^{n_{\mathsf{r}}}$ terms labelled by the vectors $\mathbf{r} \in \{0, 1\}^{n_{\mathsf{r}}}$,

$$\begin{aligned} \rho_{\mathcal{K}\mathcal{E}} &= \sum_{k,\mathbf{a}} |k\rangle\langle k|_{\mathcal{K}}\, \delta^k_{g(\mathbf{a})} \operatorname{tr}_{\mathcal{A}_{\mathsf{r}}}\left[\rho_{\mathcal{A}_{\mathsf{r}}\mathcal{E}} \prod_i \frac{1}{2}\left(\mathbb{1} + (-1)^{a_i} C_i\right)\right] \\ &= \sum_{k,\mathbf{a}} |k\rangle\langle k|_{\mathcal{K}}\, \delta^k_{g(\mathbf{a})} \operatorname{tr}_{\mathcal{A}_{\mathsf{r}}}\left[\rho_{\mathcal{A}_{\mathsf{r}}\mathcal{E}}\, 2^{-n_{\mathsf{r}}} \sum_{\mathbf{r}} \prod_i (-1)^{a_i r_i} C_i^{r_i}\right] , \end{aligned} \tag{14}$$

where we have used the power identities $C_i^0 = \mathbb{1}$ and $C_i^1 = C_i$ for full-rank operators. Next, we write the XOR function as a scalar product $g(\mathbf{a}) = \mathbf{a} \cdot \mathbf{1} \bmod 2$ with the vector $\mathbf{1} = (1, \ldots, 1)$. This allows us to write the Kronecker delta as

$$\delta^k_{g(\mathbf{a})} = \frac{1}{2}\left[1 + (-1)^{\mathbf{a}\cdot\mathbf{1}+k}\right] , \tag{15}$$

and perform the summation

$$\begin{aligned} 2^{-n_{\mathsf{r}}} \sum_{\mathbf{a}} \delta^k_{g(\mathbf{a})} (-1)^{\mathbf{a}\cdot\mathbf{r}} &= 2^{-n_{\mathsf{r}}-1} \sum_{\mathbf{a}} \left[(-1)^{\mathbf{a}\cdot\mathbf{r}} + (-1)^{\mathbf{a}\cdot(\mathbf{r}-\mathbf{1})+k}\right] \\ &= 2^{-1}\left[\delta^{\mathbf{0}}_{\mathbf{r}} + (-1)^k \delta^{\mathbf{1}}_{\mathbf{r}}\right] , \end{aligned} \tag{16}$$

where $\mathbf{0} = (0, \ldots, 0)$. Substituting Equation (16) into the state (14) gives

$$\rho_{\mathcal{K}\mathcal{E}} = \sum_k \frac{1}{2} |k\rangle\langle k|_{\mathcal{K}} \left(\rho_{\mathcal{E}} + (-1)^k \operatorname{tr}_{\mathcal{A}_{\mathsf{r}}}[\rho_{\mathcal{A}_{\mathsf{r}}\mathcal{E}} \textstyle\prod_i C_i]\right) , \tag{17}$$

which then can be used in the trace norm yielding

$$\begin{aligned}\|\rho_{\mathcal{K}\mathcal{E}} - u_{\mathcal{K}}\rho_{\mathcal{E}}\|_1 &= \sum_k \frac{1}{2}\left\|(-1)^k \operatorname{tr}_{\mathcal{A}_\mathsf{r}}(\rho_{\mathcal{A}_\mathsf{r}\mathcal{E}}\textstyle\prod_i C_i)\right\|_1 \\ &= \|\operatorname{tr}_{\mathcal{A}_\mathsf{r}}(\rho_{\mathcal{A}_\mathsf{r}\mathcal{E}}\textstyle\prod_i C_i)\|_1 \end{aligned} \tag{18}$$

For any Hermitian operator X, its trace norm satisfies $\|X\|_1 = \max_H \operatorname{tr}(HX)$, where H is constrained to be Hermitian and have eigenvalues ± 1. Therefore, there is an Hermitian operator H acting on $\mathcal{E}$, with spectrum ± 1, which satisfies

$$\|\operatorname{tr}_{\mathcal{A}_\mathsf{r}}(\rho_{\mathcal{A}_\mathsf{r}\mathcal{E}}\textstyle\prod_i C_i)\|_1 = \operatorname{tr}(\rho_{\mathcal{A}_\mathsf{r}\mathcal{E}}\left[(\textstyle\prod_i C_i)\, H\right]) \ . \tag{19}$$

We can write the spectral decomposition $H = H_+ - H_-$ with some projectors $H_\pm$ on $\mathcal{E}$ and obtain

$$\|\operatorname{tr}(\rho_{\mathcal{A}_\mathsf{r}\mathcal{E}}\textstyle\prod_i C_i)\|_1 = \operatorname{tr}(\rho_{\mathcal{A}_\mathsf{r}\mathcal{E}}\left[\textstyle\prod_i C_i\right] H_+) - \operatorname{tr}(\rho_{\mathcal{A}_\mathsf{r}\mathcal{E}}\left[\textstyle\prod_i C_i\right] H_-) \ . \tag{20}$$

At this point we recall Theorem 1, which can be written as $\pm C_i \leq S_i$. Note that here, operator C_i acts trivially on $\mathcal{B}_i$ while S_i does not. Using Lemma 6 we can generalise Theorem 1 to $\pm\prod_i C_i \leq \prod_i S_i$, which implies $\pm[\prod_i C_i]H_\pm \leq [\prod_i S_i]H_\pm$ and

$$\pm \operatorname{tr}(\rho_{\mathcal{A}_\mathsf{r}\mathcal{E}}\left[\textstyle\prod_i C_i\right] H_\pm) \leq \operatorname{tr}(\rho_{\mathcal{A}_\mathsf{r}\mathcal{B}_\mathsf{r}\mathcal{E}}\left[\textstyle\prod_i S_i\right] H_\pm)\,. \tag{21}$$

Finally, substituting this in Equation (20) and using the fact that $H_+ + H_- = \mathbb{1}$, we obtain

$$\|\operatorname{tr}_{\mathcal{A}_\mathsf{r}}(\rho_{\mathcal{A}_\mathsf{r}\mathcal{E}}\textstyle\prod_i C_i)\|_1 \leq \operatorname{tr}(\rho_{\mathcal{A}_\mathsf{r}\mathcal{B}_\mathsf{r}}\textstyle\prod_i S_i) \ , \tag{22}$$

which concludes the proof. □

2.4 Seedless extractors with arbitrary output length

In this section, we analyse seedless extractors that produce an extractor output $\mathbf{k}$ of arbitrary length m. The proof relies on randomised methods, meaning we do not obtain explicit constructions and the resulting extractors are likely computationally hard to implement. In future work, we will show that certain linear functions are seedless extractors and can be implemented in computation time $O(n_\mathsf{r} \log n_\mathsf{r})$ [44]. The following lemma is proven in Appendix A.

Lemma 3. There exist functions $g : \{0,1\}^{n_\mathsf{r}} \to \{0,1\}^m$, for $n_\mathsf{r} > 5$ and $n_\mathsf{r} - m > 0$, satisfying

$$\left|\sum_{\mathbf{a}} \left(\delta^{\mathbf{k}}_{g(\mathbf{a})} - 2^{-m}\right)(-1)^{\mathbf{a}\cdot\mathbf{r}}\right| \leq n_\mathsf{r}^2 \sqrt{2}^{n_\mathsf{r}-m} \ , \tag{23}$$

for all $\mathbf{k} \in \{0,1\}^m$ and all $\mathbf{r} \in \{0,1\}^{n_\mathsf{r}}$. We call such functions *m-bit extractor functions*.

This property is crucial for proving the following theorem, as it enables us to bound each coefficient after expanding the trace distance between the real and ideal output using the triangle inequality. Specifically: (1) For $\mathbf{r} = \mathbf{0}$, this property ensures that the sizes of the sets $\{\mathbf{a} \mid g(\mathbf{a}) = \mathbf{k}\}$ are similar for all $\mathbf{k}$. (2) For $\mathbf{r} \neq \mathbf{0}$, it ensures that the sizes of the sets $\{\mathbf{a} \mid g(\mathbf{a}) = \mathbf{k} \text{ and } \mathbf{a}\cdot\mathbf{r} = 0\}$ and $\{\mathbf{a} \mid g(\mathbf{a}) = \mathbf{k} \text{ and } \mathbf{a}\cdot\mathbf{r} = 1\}$ are similar for all $\mathbf{k}$.

Theorem 4. Let $g : \{0,1\}^{n_r} \to \{0,1\}^m$ be a function satisfying the condition (23). After measuring the n_r-round state $\rho_{\mathcal{A}_r\mathcal{B}_r\mathcal{E}}$ with the observables $\{A_i(a_i|0) : a_i = 0,1\}$ for all $i = 1,\dots,n_r$ and applying the function $\mathbf{k} = g(\mathbf{a})$ to the outcomes $\mathbf{a} = (a_1,\dots,a_{n_r})$, the resulting state $\rho_{\mathcal{K}\mathcal{E}}$ written in Equation (7) satisfies

$$\|\rho_{\mathcal{K}\mathcal{E}} - u_{\mathcal{K}}\rho_{\mathcal{E}}\|_1 \leq n_r^2\sqrt{2}^{m-n_r}\,\mathrm{tr}\left[\rho_{\mathcal{A}_r\mathcal{B}_r}\prod_{i=1}^{n_r}(\mathbb{1}+S_i)\right] , \tag{24}$$

for all $s \in [2, 2\sqrt{2}]$.

The above result shows that, the smaller the expectation of $\prod_i(\mathbb{1}+S_i)$ (i.e. the larger the violation of CHSH), the smaller the distance between the real and the ideal final outputs. If we fix this distance to a specific value $\|\rho_{\mathcal{K}\mathcal{E}} - u_{\mathcal{K}}\rho_{\mathcal{E}}\|_1 = \epsilon$, then the larger the violation of CHSH, the larger the length m of the extractor output.

Proof. We start by substituting Equation (13) in the joint state after Alice generates the final output (7) and expanding the product $\prod_i(\mathbb{1}+(-1)^{a_i}C_i)$ into 2^{n_r} terms labelled by the vectors $\mathbf{r} \in \{0,1\}^{n_r}$,

$$\begin{aligned}\rho_{\mathcal{K}\mathcal{E}} &= \sum_{\mathbf{k},\mathbf{a}} |\mathbf{k}\rangle\langle\mathbf{k}|_{\mathcal{K}}\,\delta^{\mathbf{k}}_{g(\mathbf{a})}\,\mathrm{tr}_{\mathcal{A}_r}\left[\rho_{\mathcal{A}_r\mathcal{E}}\prod_i\frac{1}{2}\left(\mathbb{1}+(-1)^{a_i}C_i\right)\right] \\ &= \sum_{\mathbf{k},\mathbf{a}} |\mathbf{k}\rangle\langle\mathbf{k}|_{\mathcal{K}}\,\delta^{\mathbf{k}}_{g(\mathbf{a})}\,\mathrm{tr}_{\mathcal{A}_r}\left[\rho_{\mathcal{A}_r\mathcal{E}}\,2^{-n_r}\sum_{\mathbf{r}}\prod_i(-1)^{a_i r_i}C_i^{r_i}\right] ,\end{aligned} \tag{25}$$

where we have used the power identities $C_i^0 = \mathbb{1}$ and $C_i^1 = C_i$ for full-rank operators. After substituting this in the left-hand side of Equation (24), applying the triangular inequality and using promise (23), we obtain

$$\begin{aligned}\|\rho_{\mathcal{K}\mathcal{E}} - u_{\mathcal{K}}\rho_{\mathcal{E}}\|_1 &= 2^{-n_r}\sum_{\mathbf{k}}\left\|\sum_{\mathbf{r}}\left(\sum_{\mathbf{a}}\left(\delta^{\mathbf{k}}_{g(\mathbf{a})}-2^{-m}\right)(-1)^{\mathbf{a}\cdot\mathbf{r}}\right)\mathrm{tr}_{\mathcal{A}_r}\left[\rho_{\mathcal{A}_r\mathcal{E}}\prod_i C_i^{r_i}\right]\right\|_1 \\ &\leq 2^{-n_r}\sum_{\mathbf{k}}\sum_{\mathbf{r}}\left|\sum_{\mathbf{a}}\left(\delta^{\mathbf{k}}_{g(\mathbf{a})}-2^{-m}\right)(-1)^{\mathbf{a}\cdot\mathbf{r}}\right|\left\|\mathrm{tr}_{\mathcal{A}_r}\left[\rho_{\mathcal{A}_r\mathcal{E}}\prod_i C_i^{r_i}\right]\right\|_1 \\ &\leq 2^{-n_r}\sum_{\mathbf{k}}\sum_{\mathbf{r}} n_r^2\sqrt{2}^{n_r-m}\left\|\mathrm{tr}_{\mathcal{A}_r}\left[\rho_{\mathcal{A}_r\mathcal{E}}\prod_i C_i^{r_i}\right]\right\|_1 \\ &= n_r^2\sqrt{2}^{m-n_r}\sum_{\mathbf{r}}\left\|\mathrm{tr}_{\mathcal{A}_r}\left[\rho_{\mathcal{A}_r\mathcal{E}}\prod_i C_i^{r_i}\right]\right\|_1 .\end{aligned} \tag{26}$$

Following the same steps as in the proof of Theorem 2: there are two complementary projectors $H_\pm$ acting on $\mathcal{E}$ such that

$$\|\mathrm{tr}(\rho_{\mathcal{A}_r\mathcal{E}}\textstyle\prod_i C_i^{r_i})\|_1 = \mathrm{tr}(\rho_{\mathcal{A}_r\mathcal{E}}\,[\prod_i C_i^{r_i}]\,H_+) - \mathrm{tr}(\rho_{\mathcal{A}_r\mathcal{E}}\,[\prod_i C_i^{r_i}]\,H_-) \ . \tag{27}$$

Lemma 6 allows us to generalise Theorem 1 to

$$\pm\textstyle\prod_i C_i^{r_i} \leq \prod_i S_i^{r_i} \ , \tag{28}$$

for any vector $\mathbf{r}$, which in turn implies

$$\pm\,\mathrm{tr}(\rho_{\mathcal{A}_r\mathcal{E}}\,[\textstyle\prod_i C_i^{r_i}]\,H_\pm) \leq \mathrm{tr}(\rho_{\mathcal{A}_r\mathcal{B}_r\mathcal{E}}\,[\textstyle\prod_i S_i^{r_i}]\,H_\pm) \ . \tag{29}$$

Substituting this in Equation (27) and using the fact that $H_+ + H_- = \mathbb{1}$, we obtain

$$\|\mathrm{tr}_{\mathcal{A}_\mathrm{r}}(\rho_{\mathcal{A}_\mathrm{r}\mathcal{E}} \prod_i C_i^{r_i})\|_1 \leq \mathrm{tr}(\rho_{\mathcal{A}_\mathrm{r}\mathcal{B}_\mathrm{r}} \prod_i S_i^{r_i}) \ . \tag{30}$$

Finally, we substitute this in Equation (26) and conclude the proof

$$\begin{aligned} \|\rho_{\mathcal{K}\mathcal{E}} - u_{\mathcal{K}}\rho_{\mathcal{E}}\|_1 &\leq n_\mathrm{r}^2 \sqrt{2}^{m-n_\mathrm{r}} \textstyle\sum_{\mathbf{r}} \mathrm{tr}(\rho_{\mathcal{A}_\mathrm{r}\mathcal{B}_\mathrm{r}} \prod_i S_i^{r_i}) \\ &= n_\mathrm{r}^2 \sqrt{2}^{m-n_\mathrm{r}} \, \mathrm{tr}(\rho_{\mathcal{A}_\mathrm{r}\mathcal{B}_\mathrm{r}} \textstyle\prod_i [\mathbb{1} + S_i]) \ . \end{aligned} \tag{31}$$

□

3 Estimation

Theorem 2 and Theorem 4 provide a relationship between the error ϵ, the length of the final output m, and the Bell-inequality violation quantified by $\langle \prod_i S_i \rangle$ or $\langle \prod_i (\mathbb{1}+S_i) \rangle$. This Bell-violation quantifier is different than the one used with standard seeded extractors, that is $\langle \sum_i S_i \rangle$. For large n, the statistical fluctuations of $\sum_i S_i$ are small, which allows us to relate the average $\langle \sum_i S_i \rangle$ to the particular value of $\sum_i S_i$ corresponding to the estimation data (a_j, b_j, x_j, y_j). Unfortunately, the quantities $\prod_i S_i$ and $\prod_i (\mathbb{1} + S_i)$ appearing in our bounds have strong fluctuations and cannot be bounded with the usual techniques.[1]

In this section, we present a proof technique for bounding $\langle \prod_i (\mathbb{1}+S_i) \rangle$ and $\langle \prod_i S_i \rangle$ with the estimation data. For this, we use the *spot-checking* procedure frequently used in DI protocols and, therefore, widely applicable. In this procedure, we randomly select a subset of rounds used for the estimation of the Bell violation, the rest of the rounds are used to generate the extractor input. This random selection limits the malicious behaviour of the devices. In order to implement this procedure and prove security, the required assumptions are:

- The devices and adversary operate according to quantum theory.
- The classical computer used for processing and statistical analysis is trusted and functions correctly.
- The quantum device comprises two isolated parts that do not exchange information during each round of the experiment (i.e., they are no-signaling).
- The quantum device does not signal to the adversary.
- The measurement devices are memoryless, meaning that the measurements in each round act on separate Hilbert spaces.

Importantly, (1) we make no assumptions about the state, which can be arbitrary and may exhibit correlations between rounds, and (2) although the measurements are modelled to act on a separate Hilbert space in each round, they do not need to be identical.

[1] We note that if the experimental state is product across the rounds, the estimator simplifies significantly and can be computed using standard techniques.

3.1 SPOT-CHECKING PROTOCOL FOR XOR EXTRACTION

In what follows, we provide a complete description of this estimation protocol and show that our XOR seedless extractor is able to extract a bit with arbitrary small error.

Set parameters: n, the total number of rounds, $p_{\mathsf{e}} \in (0,1)$, the probability for a round to be used for estimation, and $\epsilon > 0$, the tolerable error.

Data generation. For each round $l \in \{1,\dots,n\}$ repeat the following steps:

1. Generate the random variable $t_l \in \{\texttt{estimation}, \texttt{rawbit}\}$ with probabilities p_{e} and $p_{\mathsf{r}} = 1 - p_{\mathsf{e}}$ respectively.
2. If $t_l = \texttt{estimation}$ then:
   (a) Generate the random variables $x_l, y_l \in \{0,1\}$ with uniform distribution $\mathrm{P}(x_l, y_l) = 1/4$,
   (b) Perform the bi-local measurement $A_l(a_l|x_l)B_l(b_l|y_l)$ with outcomes $a_l, b_l \in \{0,1\}$,
   (c) Record the variable $z_l = a_l + b_l + x_l y_l \bmod 2$, which will be used to evaluate the CHSH inequality.
3. If $t_l = \texttt{rawbit}$ then perform the local measurement $A_l(a_l|0)$ and keep the outcome a_l as part of the extractor input.

Data processing. Denote by $n_{\mathsf{e}} \in \{0,n\}$ the number of rounds l with $t_l = \texttt{estimation}$, assign an index $j \in \{1,\dots,n_{\mathsf{e}}\}$ to each of them, and compile the estimation data $\mathbf{z} = (z_1 \dots, z_{n_{\mathsf{e}}})$. Denote by $n_{\mathsf{r}} \in \{0,n\}$ the number of rounds l with $t_l = \texttt{rawbit}$, assign an index $i \in \{1,\dots,n_{\mathsf{r}}\}$ to each of them, and compile the extractor input $\mathbf{a} = (a_1,\dots,a_{n_{\mathsf{r}}})$. The numbers $n_{\mathsf{e}}, n_{\mathsf{r}}$ are a function of $\mathbf{t} = (t_1,\dots,t_n)$.

1. Calculate the length of the final output as a function of $\mathbf{t}, \mathbf{z}$ with the formula;

$$m(\mathbf{t},\mathbf{z}) = \begin{cases} 1 & \text{if} \quad \max\limits_{s,\alpha_0,\alpha_1,\beta} \sum_{j=1}^{n_{\mathsf{e}}} \alpha_{z_j} + (\beta-1)n_{\mathsf{r}} - 2\log_2(1/\epsilon) \geq 0 \\ 0 & \text{otherwise} \end{cases} , \tag{32}$$

   where the maximisation over the parameters $s \in [2, 2\sqrt{2}]$ and $\alpha_0, \alpha_1, \beta \in \mathbb{R}$ is constrained by

$$p_{\mathsf{r}}\sqrt{2}^{\beta-1}(\mu_s - 4\nu_s) + p_{\mathsf{e}}\sqrt{2}^{\alpha_0} = 1 , \tag{33}$$

$$p_{\mathsf{r}}\sqrt{2}^{\beta-1}(\mu_s + 4\nu_s) + p_{\mathsf{e}}\sqrt{2}^{\alpha_1} = 1 . \tag{34}$$

2. Generate the final output $\mathbf{k} = g(\mathbf{a}) \in \{0,1\}^{m(\mathbf{t},\mathbf{z})}$ by applying to the extractor input the XOR function $g : \{0,1\}^{n_{\mathsf{r}}} \to \{0,1\}^{m(\mathbf{t},\mathbf{z})}$, defined in Equation (10).

The expression for the extractor output length (32) may seem unintuitive, but it is the most general form that allow us to prove the following theorem.

Theorem 5. For any choice of $n, p_{\mathsf{r}}, \epsilon$, the above protocol generates a final output $\rho_{\mathcal{KE}|\mathbf{t},\mathbf{z}}$ satisfying the following security condition

$$\sum_{\mathbf{t}\in\{0,1\}^n} \sum_{\mathbf{z}\in\{0,1\}^{n_{\mathsf{e}}}} \mathrm{P}(\mathbf{t},\mathbf{z}) \left\| \rho_{\mathcal{KE}|\mathbf{t},\mathbf{z}} - u_{\mathcal{K}} \rho_{\mathcal{E}|\mathbf{t},\mathbf{z}} \right\|_1 \leq \epsilon , \tag{35}$$

where $\rho_{\mathcal{KE}|\mathbf{t},\mathbf{z}}$ is defined in Equation (7) with $m = m(\mathbf{t},\mathbf{z})$ defined in Equation (32) (XOR extraction protocol) and Equation (39) (m-bit extraction protocol).

This theorem (proven in Appendix B) demonstrates that the above protocol produces a secure final output. The expression for output length (32) provides a general form with free parameters α_0, α_1, and β which, importantly, allows a free parameter to be distributed across each of the factors in the error term when expanded (see Equation (67)). Then, the maximisation constraints in (33) and (34) are required to ensure that specific operator identities are maintained while incorporating this additional freedom (in particular, to recover the identities in Equations (68) and (69)).

Next, we analyse the noise tolerance of the XOR extractor in the spot-checking protocol presented above. Define the relative frequency of the estimation outcomes as

$$q_z = \frac{|\{z_j = z : j = 1, \ldots, n_{\mathsf{e}}\}|}{n_{\mathsf{e}}} \ , \tag{36}$$

for $z = 0, 1$. Which is related to the CHSH expression (1) via

$$\mathrm{CHSH} = 4(q_0 - q_1) = 8q_0 - 4 \ . \tag{37}$$

Note that, for any permutation σ of $(z_1, z_2, \ldots, z_{n_{\mathsf{e}}})$ we have $m(\mathbf{t}, \sigma\mathbf{z}) = m(\mathbf{t}, \mathbf{z})$, hence m depends on $\mathbf{z}$ through q_0. Also, m depends on $\mathbf{t}$ via its relative frequency n_{e}/n, which in the large-n limit is $n_{\mathsf{e}}/n \approx p_{\mathsf{e}}$ with high probability. Therefore, if we assume a constant error $\epsilon > 0$, in the large-n limit, the condition for positive yield (i.e. $m = 1$) is

$$\max_{s,\alpha_0,\alpha_1,\beta} \ p_{\mathsf{e}}(\alpha_0 q_0 + \alpha_1 q_1) + \beta p_{\mathsf{r}} \geq 0 \ . \tag{38}$$

subject to the constraints (33) and (34).

The minimum value of CHSH with positive yield as a function of p_e is shown in Figure 1. Note that for $p_{\mathsf{e}} \geq 0.74$ a bit can be generated with arbitrarily small Bell violation. Interestingly, as shown by Figure 1, a necessary requirement for the extraction of a single bit is $p_e > 0.5$. This contrasts with the fact that, in protocols with seeded extractors, only a very small portion is required for estimation, evidencing a limitation of our estimation method.

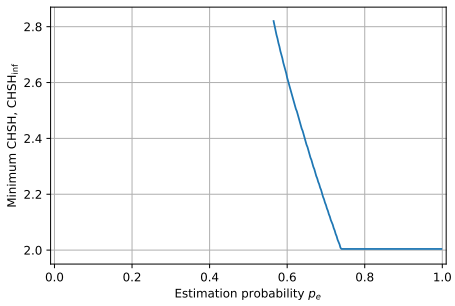

Figure 1: This plot shows the minimum CHSH value (38) that our XOR extractor (presented in Theorem 2) can produce a single bit with arbitrarily small error in the large-n regime.

3.2 SPOT-CHECKING PROTOCOL FOR m-BIT EXTRACTION

In what follows, we describe the spot-checking protocol for estimating the error of seedless extraction using our m-bit extractor functions. We then calculate the maximum efficiency and extraction rates. The protocol is identical to that of the XOR extractor in Section 3.1 except for the 'Data processing' step, which is replaced by the following:

Data processing. Define $\epsilon, \mathbf{t}, n_{\mathsf{e}}, n_{\mathsf{r}}, j, i, \mathbf{z}, \mathbf{a}$ as in Section 3.1.

1. Calculate the length of the final output as a function of $\mathbf{t}, \mathbf{z}$ with the formula

$$m(\mathbf{t}, \mathbf{z}) = \left\lfloor \max_{s,\alpha_0,\alpha_1,\beta} \sum_{j=1}^{n_{\mathsf{e}}} \alpha_{z_j} + \beta n_{\mathsf{r}} - 2\log_2 \frac{1}{\epsilon} - 4\log_2(n_{\mathsf{r}}) \right\rfloor . \tag{39}$$

   where the maximisation over the parameters $s \in [2, 2\sqrt{2}]$ and $\alpha_0, \alpha_1, \beta \in \mathbb{R}$ is constrained by

$$p_{\mathsf{r}}\sqrt{2}^{\beta-1}\left(1 + \mu_s - 4\nu_s\right) + p_{\mathsf{e}}\sqrt{2}^{\alpha_0} = 1 \ , \tag{40}$$

$$p_{\mathsf{r}}\sqrt{2}^{\beta-1}\left(1 + \mu_s + 4\nu_s\right) + p_{\mathsf{e}}\sqrt{2}^{\alpha_1} = 1 \ . \tag{41}$$

2. Generate the final output $\mathbf{k} = g(\mathbf{a}) \in \{0,1\}^{m(\mathbf{t},\mathbf{z})}$ by applying to the extractor input a function $g : \{0,1\}^{n_{\mathsf{r}}} \to \{0,1\}^{m(\mathbf{t},\mathbf{z})}$ satisfying (23).

This protocol also satisfies the security condition (35) in Theorem 5 (proven in Appendix B). Similar to the XOR case, the expression for the extractor output length in (39) is chosen as it provides the most general form that allows us to prove security. Specifically, this form includes free parameters that can be distributed across each factor in the error term upon expansion. The maximisation constraints in (40) and (41) ensure that specific operator identities are preserved under this added flexibility. For full details, see the proof in Appendix B.

Using the expression for final output length (39) we can obtain the efficiency rate $\mathcal{R}_{\text{eff}}$ and the extraction rate $\mathcal{R}_{\text{ext}}$. The extraction rate, the number of output bits per extractor input bit, is given by

$$\mathcal{R}_{\text{ext}} = \lim_{n\to\infty} \frac{m(\mathbf{t}, \mathbf{z})}{n_{\mathsf{r}}} = \frac{p_{\mathsf{e}}}{p_{\mathsf{r}}}(\alpha_0 q_0 + \alpha_1 q_1) + \beta \ , \tag{42}$$

and the efficiency rate, the number of output bits per round, is given by

$$\mathcal{R}_{\text{eff}} = \lim_{n\to\infty} \frac{m(\mathbf{t}, \mathbf{z})}{n} = p_{\mathsf{e}}(\alpha_0 q_0 + \alpha_1 q_1) + \beta p_{\mathsf{r}} \ . \tag{43}$$

The maximum value of $\mathcal{R}_{\text{ext}}$ and $\mathcal{R}_{\text{eff}}$ as a function of CHSH is depicted in Figure 2. The extraction rate approaches 1, demonstrating that our m-bit extractor performs optimally in the large-n regime and under high CHSH violation. However, the maximum efficiency rates are very low. This is because the estimation procedure consumes a significant proportion of rounds, as with the XOR extractor. Therefore, independently improving techniques for estimation of our Bell value quantifier's would significantly enhance the practicality of our seedless extractors.

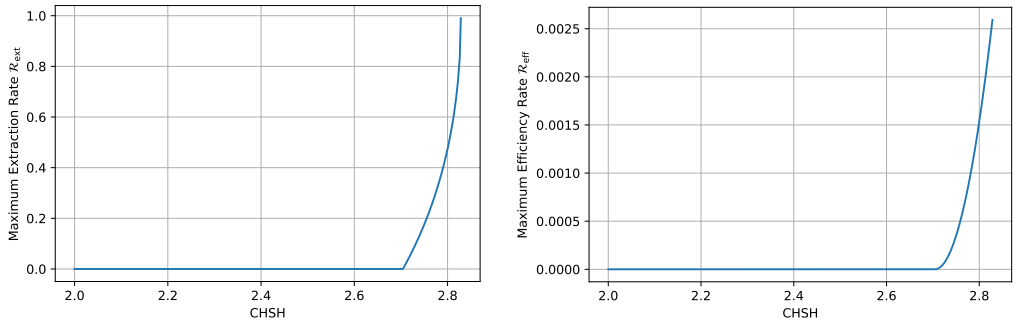

Figure 2: The left hand side plot shows the maximum extraction rate, $\mathcal{R}_{\text{ext}}$ (42), for different values of CHSH. The right hand side shows the maximum efficiency rate, $\mathcal{R}_{\text{eff}}$ (43), for our m-bit extractor functions, for different values of CHSH $= 4(q_0 - q_1) \in (2, 2\sqrt{2})$. The maximisation's are performed over the variables $s \in (2, 2\sqrt{2})$, $p_{\mathsf{e}} \in (0, 1)$ and $\alpha_0, \alpha_1, \beta \in \mathbb{R}$ such that constraints (40) and (41) are satisfied.

4 Conclusion and discussion

In this work we have proven that, in the context of DI protocols, randomness extraction and privacy amplification can be done without the need of a seed. These ideas and results constitute the first steps of a new paradigm for DI quantum cryptography, which might one day allow to minimise the resources required in various tasks. This new paradigm contains the following open problems, which we will address in future work.

- Analyse and generalise our spot-checking protocol to a DI protocol for randomness amplification. The protocol for seedless extraction presented here requires initial randomness to generate the variables t_l, x_l and y_l, so that the Bell violation can be tested. However, it is not known what the minimal statistical requirements are for this initial randomness, which could be weaker than those of a Santha-Vazirani source [22] (see [45] for necessary requirements).

- Characterise the seedless extractors consisting of linear functions. We have evidence that these extractors improve the efficiency rate and have a strong connection with linear error-correcting codes [44].

- The results presented in this work are restricted to the CHSH scenario: two parties with binary inputs and outputs. However, they can be generalised to arbitrary scenarios by using the NPA hierarchy [46]. In particular, we expect that increasing the number of inputs will cause the efficiency rate to improve substantially, as it happens in other scenarios [32, 47]. The efficiency rate can also be improved by using both outputs a_l, b_l as the extractor input, or by recycling the inputs used for estimation, since these also contain randomness [48].

- Characterise the optimal efficiency rates of seedless extractors. There is room for improvement in several parts of our scheme, one of them being the use of the measurement outcomes of all rounds for both: estimation and extractor input generation. Another being the development of new proof techniques for bounding $\langle \prod_i S_i \rangle$, or related quantities.

- Our results can be implemented with only two quantum devices (one for Alice and one for Bob). However, our theorems assume that the measurement devices have

no internal memory (as in [41]), which is an inconvenient assumption. Because of the link between Bell violation and independence of the rounds discussed in the Introduction, we expect that this assumption is unnecessary. We will analyse this hypothesis in future work.

Acknowledgements. We acknowledge useful discussions with Mafalda Almeida, Matty Hoban, Karan Khathuria, Simone Lin, Renato Renner, Ramona Wolf and Erik Woodhead. LM acknowledges financial support from the EPSRC Prosperity Partnership in Quantum Software for Simulation and Modelling (grant EP/S005021/1).

References

[1] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani. Device-independent security of quantum cryptography against collective attacks. *Physical Review Letters*, 98(23):230501, 2007. doi:https://doi.org/10.1103/PhysRevLett.98.230501.

[2] Stefano Pironio, Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani. Device-independent quantum key distribution secure against collective attacks. *New Journal of Physics*, 11(4):045021, 2009. doi:https://doi.org/10.1088/1367-2630/11/4/045021.

[3] John S Bell. On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika*, 1 (3):195, 1964.

[4] Dominic Mayers and Andrew Yao. Quantum cryptography with imperfect apparatus. In *Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No. 98CB36280)*, pages 503–509. IEEE, 1998. doi:https://doi.org/10.1109/SFCS.1998.743501.

[5] Renato Renner. Security of quantum key distribution. *International Journal of Quantum Information*, 6(01):1–127, 2008. doi:https://doi.org/10.1142/S0219749908003256.

[6] Umesh Vazirani and Thomas Vidick. Fully device independent quantum key distribution. *Communications of the ACM*, 62(4):133–133, 2019. doi:https://doi.org/10.1145/3310974.

[7] David P Nadlinger, Peter Drmota, Bethan C Nichol, Gabriel Araneda, Dougal Main, Raghavendra Srinivas, David M Lucas, Christopher J Ballance, Kirill Ivanov, Ernest Y-Z Tan, et al. Experimental quantum key distribution certified by Bell's theorem. *Nature*, 607(7920):682–686, 2022. doi:https://doi.org/10.1038/s41586-022-04941-5.

[8] Víctor Zapatero, Tim van Leent, Rotem Arnon-Friedman, Wen-Zhao Liu, Qiang Zhang, Harald Weinfurter, and Marcos Curty. Advances in device-independent quantum key distribution. *npj Quantum Information*, 9(1):10, 2023. doi:https://doi.org/10.1038/s41534-023-00684-x.

[9] Roger Colbeck and Adrian Kent. Private randomness expansion with untrusted devices. *Journal of Physics A: Mathematical and Theoretical*, 44(9):095305, 2011. doi:https://doi.org/10.1088/1751-8113/44/9/095305.

[10] Antonio Acín and Lluis Masanes. Certified randomness in quantum physics. *Nature*, 540(7632):213–219, 2016. doi:https://doi.org/10.1038/nature20119.

[11] Roger Colbeck and Renato Renner. Free randomness can be amplified. *Nature Physics*, 8(6):450–453, 2012. doi:https://doi.org/10.1038/nphys2300.

[12] Alexandru Gheorghiu, Elham Kashefi, and Petros Wallden. Robustness and device independence of verifiable blind quantum computing. *New Journal of Physics*, 17(8): 083040, 2015. doi:https://doi.org/10.1088/1367-2630/17/8/083040.

[13] Srijita Kundu and Ernest Y-Z Tan. Device-independent uncloneable encryption. *Quantum*, 9:1582, 2025. doi:https://doi.org/10.22331/q-2025-01-08-1582.

[14] Anne Broadbent and Peter Yuen. Device-independent oblivious transfer from the bounded-quantum-storage-model and computational assumptions. *New Journal of Physics*, 25(5):053019, 2023. doi:https://doi.org/10.1088/1367-2630/accf32.

[15] Rotem Arnon-Friedman, Christopher Portmann, and Volkher B. Scholz. Quantum-Proof Multi-Source Randomness Extractors in the Markov Model. In Anne Broadbent, editor, *11th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2016)*, volume 61 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 2:1–2:34, Dagstuhl, Germany, 2016. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. ISBN 978-3-95977-019-4. doi:https://doi.org/10.4230/LIPIcs.TQC.2016.2.

[16] Marshall Ball, Oded Goldreich, and Tal Malkin. Randomness extraction from somewhat dependent sources. In *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2022. doi:https://doi.org/10.4230/LIPIcs.ITCS.2022.12.

[17] Xin Li. Two source extractors for asymptotically optimal entropy, and (many) more. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1271–1281. IEEE, 2023. doi:https://doi.org/10.1109/FOCS57990.2023.00075.

[18] Xiongfeng Ma, Feihu Xu, He Xu, Xiaoqing Tan, Bing Qi, and Hoi-Kwong Lo. Postprocessing for quantum random-number generators: Entropy evaluation and randomness extraction. *Physical Review A*, 87(6):062327, 2013. doi:https://doi.org/10.1103/PhysRevA.87.062327.

[19] Cameron Foreman, Richie Yeung, Alec Edgington, and Florian J Curchod. Cryptomite: A versatile and user-friendly library of randomness extractors. *Quantum*, 9: 1584, 2025. doi:https://doi.org/10.22331/q-2025-01-08-1584.

[20] John F Clauser, Michael A Horne, Abner Shimony, and Richard A Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15): 880, 1969. doi:https://doi.org/10.1103/PhysRevLett.23.880.

[21] Yukun Wang, Xingyao Wu, and Valerio Scarani. All the self-testings of the singlet for two binary measurements. *New Journal of Physics*, 18(2):025021, 2016. doi:https://doi.org/10.1088/1367-2630/18/2/025021.

[22] Miklos Santha and Umesh V Vazirani. Generating quasi-random sequences from semi-random sources. *Journal of computer and system sciences*, 33(1):75–87, 1986. doi:https://doi.org/10.1016/0022-0000(86)90044-9.

[23] John Von Neumann. Various techniques used in connection with random digits. *John von Neumann, Collected Works*, 5:768–770, 1963.

[24] Ariel Gabizon and Ran Raz. Deterministic extractors for affine sources over large fields. *Combinatorica*, 28(4):415–440, 2008. doi:https://doi.org/10.1007/s00493-008-2259-3.

[25] Ariel Gabizon, Ran Raz, and Ronen Shaltiel. Deterministic extractors for bit-fixing sources by obtaining an independent seed. *SIAM Journal on Computing*, 36(4):1072–1094, 2006. doi:https://doi.org/10.1137/S0097539705447049.

[26] Jesse Kamp, Anup Rao, Salil Vadhan, and David Zuckerman. Deterministic extractors for small-space sources. In *Proceedings of the Thirty-Eighth*

Annual ACM Symposium on Theory of Computing, pages 691–700, 2006. doi:https://doi.org/10.1145/1132516.1132613.

[27] Xin Li. Improved two-source extractors, and affine extractors for polylogarithmic entropy. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 168–177. IEEE, 2016. doi:https://doi.org/10.1109/FOCS.2016.26.

[28] Eshan Chattopadhyay and Xin Li. Extractors for sumset sources. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, pages 299–311, 2016. doi:https://doi.org/10.1145/2897518.2897643.

[29] Salman Beigi, Andrej Bogdanov, Omid Etesami, and Siyao Guo. Optimal deterministic extractors for generalized santha-vazirani sources. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018. doi:https://doi.org/10.4230/LIPIcs.APPROX-RANDOM.2018.30.

[30] Eshan Chattopadhyay. Recent advances in randomness extraction. *Entropy*, 24(7): 880, 2022. doi:https://doi.org/10.3390/e24070880.

[31] Harry Buhrman, Richard Cleve, Serge Massar, and Ronald De Wolf. Nonlocality and communication complexity. *Reviews of modern physics*, 82(1):665, 2010. doi:https://doi.org/10.1103/RevModPhys.82.665.

[32] Lluis Masanes. Universally composable privacy amplification from causality constraints. *Physical Review Letters*, 102(14):140501, 2009. doi:https://doi.org/10.1103/PhysRevLett.102.140501.

[33] Rodrigo Gallego, Lluis Masanes, Gonzalo De La Torre, Chirag Dhara, Leandro Aolita, and Antonio Acín. Full randomness from arbitrarily deterministic events. *Nature Communications*, 4(1):2654, 2013. doi:https://doi.org/10.1038/ncomms3654.

[34] Lewis Wooltorton, Peter Brown, and Roger Colbeck. Device-independent quantum key distribution with arbitrarily small nonlocality. *Physical Review Letters*, 132(21): 210802, 2024. doi:https://doi.org/10.1103/PhysRevLett.132.210802.

[35] Máté Farkas. Unbounded device-independent quantum key rates from arbitrarily small nonlocality. *Physical Review Letters*, 132(21):210803, 2024. doi:https://doi.org/10.1103/PhysRevLett.132.210803.

[36] Ravishankar Ramanathan, Fernando GSL Brandão, Karol Horodecki, Michał Horodecki, Paweł Horodecki, and Hanna Wojewódka. Randomness amplification under minimal fundamental assumptions on the devices. *Physical Review Letters*, 117 (23):230501, 2016. doi:https://doi.org/10.1103/PhysRevLett.117.230501.

[37] Ravishankar Ramanathan, Michał Horodecki, Hammad Anwer, Stefano Pironio, Karol Horodecki, Marcus Grünfeld, Sadiq Muhammad, Mohamed Bourennane, and Paweł Horodecki. Practical no-signalling proof randomness amplification using Hardy paradoxes and its experimental implementation. *arXiv preprint arXiv:1810.11648*, 2018. doi:https://doi.org/10.48550/arXiv.1810.11648.

[38] Max Kessler and Rotem Arnon-Friedman. Device-independent randomness amplification and privatization. *IEEE Journal on Selected Areas in Information Theory*, 1 (2):568–584, 2020. doi:https://doi.org/10.1109/JSAIT.2020.3012498.

[39] Ravishankar Ramanathan, Michał Banacki, and Paweł Horodecki. No-signaling-proof randomness extraction from public weak sources. *arXiv preprint arXiv:2108.08819*, 2021. doi:https://doi.org/10.48550/arXiv.2108.08819.

[40] Cameron Foreman, Sherilyn Wright, Alec Edgington, Mario Berta, and Florian J Curchod. Practical randomness amplification and privatisation with implementations on quantum computers. *Quantum*, 7:969, 2023. doi:https://doi.org/10.22331/q-2023-03-30-969.

[41] Lluis Masanes, Stefano Pironio, and Antonio Acín. Secure device-independent quantum key distribution with causally independent measurement devices. *Nature Communications*, 2(1):238, 2011. doi:https://doi.org/10.1038/ncomms1244.

[42] Ran Canetti. Universally composable security: A new paradigm for cryptographic protocols. In *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*, pages 136–145. IEEE, 2001. doi:https://doi.org/10.1109/SFCS.2001.959888.

[43] Israel Gelfand and Mark Neumark. On the imbedding of normed rings into the ring of operators in Hilbert space. *Matematicheskii Sbornik*, 12(2):197–217, 1943.

[44] Cameron Foreman and Lluis Masanes. *In preparation.*

[45] Le Phuc Thinh, Lana Sheridan, and Valerio Scarani. Bell tests with min-entropy sources. *Physical Review A—Atomic, Molecular, and Optical Physics*, 87(6):062121, 2013. doi:https://doi.org/10.1103/PhysRevA.87.062121.

[46] Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, 10(7):073013, 2008. doi:https://doi.org/10.1088/1367-2630/10/7/073013.

[47] Lluis Masanes, Renato Renner, Matthias Christandl, Andreas Winter, and Jonathan Barrett. Full security of quantum key distribution from no-signaling constraints. *IEEE Transactions on Information Theory*, 60(8):4973–4986, 2014. doi:https://doi.org/10.1109/TIT.2014.2329417.

[48] Rutvij Bhavsar, Sammy Ragy, and Roger Colbeck. Improved device-independent randomness expansion rates using two sided randomness. *New Journal of Physics*, 25 (9):093035, 2023. doi:https://doi.org/10.1088/1367-2630/acf393.

A Proofs of Lemmas

Lemma 3. There exist functions $g : \{0,1\}^{n_r} \to \{0,1\}^m$, for $n_r > 5$ and $n_r - m > 0$, satisfying

$$\left| \sum_{\mathbf{a}} \left(\delta^{\mathbf{k}}_{g(\mathbf{a})} - 2^{-m} \right) (-1)^{\mathbf{a}\cdot\mathbf{r}} \right| \leq n_r^2 \sqrt{2}^{n_r - m} , \tag{44}$$

for all $\mathbf{k} \in \{0,1\}^m$ and all $\mathbf{r} \in \{0,1\}^{n_r}$.

Proof. Consider a random function $G : \{0,1\}^{n_r} \to \{0,1\}^m$, which assigns each $\mathbf{a} \in \{0,1\}^{n_r}$ to a uniformly random and independent element $\mathbf{k} \in \{0,1\}^m$. Define 2^{n_r} random variables indexed by $\mathbf{a} \in \{0,1\}^{n_r}$ as

$$X_{\mathbf{a}}(\mathbf{k}, \mathbf{r}) = \delta^{\mathbf{k}}_{G(\mathbf{a})} (-1)^{\mathbf{a}\cdot\mathbf{r}} , \tag{45}$$

where $\delta^{\mathbf{k}}_{G(\mathbf{a})} = 1$ if $G(\mathbf{a}) = \mathbf{k}$ and is otherwise 0. For any particular value $\mathbf{k}^*$ and $\mathbf{r}^*$, these random variables are independent (since G is a random function) and satisfy the probability assignment

$$\mathrm{P}(X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^*) = 0) = (1 - 2^{-m}) , \tag{46}$$

$$\mathrm{P}(X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^*) = (-1)^{\mathbf{a}\cdot\mathbf{r}^*}) = 2^{-m} . \tag{47}$$

Therefore, each $X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^*)$ has mean $\mathbb{E}(X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^*)) = (-1)^{\mathbf{a}\cdot\mathbf{r}^*} 2^{-m}$ and second moment $\mathbb{E}(X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^*)^2) = 2^{-m}$.

Bernstein's inequality states; for independent random variables $X_1, \dots X_n$,

$$\mathrm{P}(|\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mathbb{E}(X_i)| \geq C) \leq 2\exp\left(\frac{-C^2}{2(\sum_{i=1}^{n}\mathbb{E}(X_i^2) + \frac{1}{3}\max_i |X_i| C)}\right) . \tag{48}$$

Using our variables $X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^*)$, we obtain

$$\begin{aligned}\mathrm{P}(|\sum_{\mathbf{a}} X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^*) - \sum_{\mathbf{a}} (-1)^{\mathbf{a}\cdot\mathbf{r}^*} 2^{-m}| \geq C) &\leq 2\exp\left(\frac{-C^2}{2(\sum_{\mathbf{a}} 2^{-m} + \frac{1}{3}C)}\right) \\ &= 2\exp\left(\frac{-C^2}{2(2^{n_\mathsf{r}-m} + \frac{1}{3}C)}\right) . \end{aligned} \tag{49}$$

When $\mathbf{r}^* = \mathbf{0}$, $\sum_{\mathbf{a}} (-1)^{\mathbf{a}\cdot\mathbf{r}^*} 2^{-m} = \sum_{\mathbf{a}} 2^{-m} = 2^{n_\mathsf{r}-m}$, and so

$$\mathrm{P}(|\sum_{\mathbf{a}} \left(X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^* = \mathbf{0}) - 2^{-m}\right)| \geq C) \leq 2\exp\left(\frac{-C^2}{2(2^{n_\mathsf{r}-m} + \frac{1}{3}C)}\right) . \tag{50}$$

When $\mathbf{r}^* \neq \mathbf{0}$, $\sum_{\mathbf{a}} (-1)^{\mathbf{a}\cdot\mathbf{r}^*} 2^{-m} = 0$, and so

$$\mathrm{P}(|\sum_{\mathbf{a}} X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^* \neq \mathbf{0})| \geq C) \leq 2\exp\left(\frac{-C^2}{2(2^{n_\mathsf{r}-m} + \frac{1}{3}C)}\right) . \tag{51}$$

Together, this can be expressed for any $\mathbf{r}^* \in \{0,1\}^{n_\mathsf{r}}$ as

$$\mathrm{P}(|\sum_{\mathbf{a}} \left(X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^*) - (-1)^{\mathbf{a}\cdot\mathbf{r}^*} 2^{-m}\right)| \geq C) \leq 2\exp\left(\frac{-C^2}{2(2^{n_\mathsf{r}-m} + \frac{1}{3}C)}\right) , \tag{52}$$

and, setting $C = n_\mathsf{r}^2 \sqrt{2}^{n_\mathsf{r}-m}$, leads to

$$\mathrm{P}(|\sum_{\mathbf{a}} \left(X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^*) - (-1)^{\mathbf{a}\cdot\mathbf{r}^*} 2^{-m}\right)| \geq n_\mathsf{r}^2 \sqrt{2}^{n_\mathsf{r}-m}) \tag{53}$$

$$\leq p^* := 2\exp\left(\frac{-n_\mathsf{r}^4 2^{n_\mathsf{r}-m}}{2(2^{n_\mathsf{r}-m} + \frac{1}{3}n_\mathsf{r}^2\sqrt{2}^{n_\mathsf{r}-m})}\right) . \tag{54}$$

Next, we apply the union bound to lower bound the probability that Equation (44) holds for all $\mathbf{r}$ and $\mathbf{k}$, as

$$\begin{aligned}&\bigcap_{\mathbf{k},\mathbf{r}} \mathrm{P}(|\sum_{\mathbf{a}} \left(X_{\mathbf{a}}(\mathbf{k}, \mathbf{r}) - (-1)^{\mathbf{a}\cdot\mathbf{r}} 2^{-m}\right)| \leq C) \\ &\quad = 1 - \bigcup_{\mathbf{k},\mathbf{r}} \mathrm{P}(|\sum_{\mathbf{a}} \left(X_{\mathbf{a}}(\mathbf{k}, \mathbf{r}) - (-1)^{\mathbf{a}\cdot\mathbf{r}} 2^{-m}\right)| \geq C) \qquad (55) \\ &\quad \geq 1 - \sum_{\mathbf{k},\mathbf{r}} \mathrm{P}(|\sum_{\mathbf{a}} \left(X_{\mathbf{a}}(\mathbf{k}, \mathbf{r}) - (-1)^{\mathbf{a}\cdot\mathbf{r}} 2^{-m}\right)| \geq C) \\ &\quad = 1 - 2^{n_\mathsf{r}+m} \mathrm{P}(|\sum_{\mathbf{a}} \left(X_{\mathbf{a}}(\mathbf{k}^*, \mathbf{r}^*) - (-1)^{\mathbf{a}\cdot\mathbf{r}^*} 2^{-m}\right)| \geq C) \\ &\quad \geq 1 - 2^{n_\mathsf{r}+m} p^* . \qquad (56)\end{aligned}$$

Finally, to prove the existence of functions satisfying Equation (44) for all $\mathbf{k}$ and $\mathbf{r}$, it suffices to show that $1 - 2^{n_\mathsf{r}+m} p^* > 0$, i.e., $2^{n_\mathsf{r}+m} p^* < 1$. Given $n_\mathsf{r} - m > 0$ and $n_\mathsf{r} > 5$; we

have that

$$\begin{aligned} n_{\mathsf{r}}^4 2^{n_{\mathsf{r}}-m} &> 4n_{\mathsf{r}}2^{n_{\mathsf{r}}-m}(1+\frac{1}{3}n_{\mathsf{r}}^2) \\ &> 4n_{\mathsf{r}}(2^{n_{\mathsf{r}}-m}+\frac{1}{3}n_{\mathsf{r}}^2\sqrt{2}^{n_{\mathsf{r}}-m}) \\ &\geq \ln(2)(n_{\mathsf{r}}+m)2(2^{n_{\mathsf{r}}-m}+\frac{1}{3}n_{\mathsf{r}}^2\sqrt{2}^{n_{\mathsf{r}}-m})+1 \ , \end{aligned} \tag{57}$$

which ensures $2^{n_{\mathsf{r}}+m}p^* < 1$ and completes the proof. □

Lemma 6. Suppose that for every $i \in \mathcal{I}$ there are two Hermitian operators C_i, S_i acting on a Hilbert space $\mathcal{A}_i$ such that $\pm C_i \leq S_i$, then

$$\pm \bigotimes_{i\in\mathcal{I}} C_i \leq \bigotimes_{i\in\mathcal{I}} S_i \ . \tag{58}$$

Proof. For any assignment $\xi_i = \pm$ for all $i \in \mathcal{I}$ we have $\bigotimes_i (S_i + \xi_i C_i) \geq 0$. Therefore, if we average this product over all configurations $\{\xi_i\}$ such that $\prod_i \xi_i = 1$ we obtain the positive operator

$$0 \leq \mathop{\mathbb{E}}_{\{\xi_i\}} \bigotimes_i (S_i + \xi_i C_i) = \bigotimes_i S_i + \bigotimes_i C_i \ . \tag{59}$$

Similarly, if we average the product over all configurations $\{\xi_i\}$ such that $\prod_i \xi_i = -1$ then

$$0 \leq \mathop{\mathbb{E}}_{\{\xi_i\}} \bigotimes_i (S_i + \xi_i C_i) = \bigotimes_i S_i - \bigotimes_i C_i \tag{60}$$

is positive too. □

B Proof of Theorem 5

Theorem 5. The protocols of Section 3 generate a final output $\rho_{\mathcal{K}\mathcal{E}|\mathbf{t},\mathbf{z}}$ satisfying the following security condition

$$\sum_{\mathbf{t}\in\{0,1\}^n} \sum_{\mathbf{z}\in\{0,1\}^{n_{\mathsf{e}}}} \mathrm{P}(\mathbf{t},\mathbf{z}) \left\| \rho_{\mathcal{K}\mathcal{E}|\mathbf{t},\mathbf{z}} - u_{\mathcal{K}} \rho_{\mathcal{E}|\mathbf{t},\mathbf{z}} \right\|_1 \leq \epsilon \ . \tag{61}$$

Proof. The random variables $\mathbf{t} \in \{\texttt{estimation}, \texttt{rawbit}\}^n$ are independent and identically distributed according to $(p_{\mathsf{e}}, p_{\mathsf{r}})$. If in round l we have $t_l = \mathsf{estimation}$ then the systems $\mathcal{A}_l$ and $\mathcal{B}_l$ are included in $\mathcal{A}_{\mathsf{e}} = \bigotimes_{j=1}^{n_{\mathsf{e}}} \mathcal{A}_j$ and $\mathcal{B}_{\mathsf{e}} = \bigotimes_{j=1}^{n_{\mathsf{e}}} \mathcal{B}_j$, and used for estimation. If $t_l = \mathrm{rawbit}$ then the systems $\mathcal{A}_l$ and $\mathcal{B}_l$ are included in $\mathcal{A}_{\mathsf{r}} = \bigotimes_{i=1}^{n_{\mathsf{r}}} \mathcal{A}_i$ and $\mathcal{B}_{\mathsf{r}} = \bigotimes_{i=1}^{n_{\mathsf{r}}} \mathcal{B}_i$, and used for generating the extractor input. Without loss of generality we can assume that $\mathbf{t}$ is initially generated before any measurement, and right after, we can re-order the rounds and write the global state as $\rho_{\mathcal{A}_{\mathsf{r}}\mathcal{B}_{\mathsf{r}}\mathcal{A}_{\mathsf{e}}\mathcal{B}_{\mathsf{e}}\mathcal{E}|\mathbf{t}}$.

In each estimation round $j \in \{1, \ldots, n_{\mathsf{e}}\}$ the pair $\mathcal{A}_j \otimes B_j$ is measured with

$$Q_{z_j} = \sum_{a,b,x,y} \frac{1}{4} A_j(a|x) B_j(b|y) \, \delta^{z_j}_{a+b+xy \bmod 2} \ , \tag{62}$$

which has outcomes $z_j = 0, 1$. This produces the estimation data $\mathbf{z} = (z_1, \ldots, z_{n_{\mathsf{e}}})$ distributed according to

$$\mathrm{P}(\mathbf{z}|\mathbf{t}) = \mathrm{tr}\left[\rho_{\mathcal{A}_{\mathsf{e}}\mathcal{B}_{\mathsf{e}}|\mathbf{t}} \prod_{j=1}^{n_{\mathsf{e}}} Q_{z_j} \right] \ . \tag{63}$$

The global state conditioned on a particular value of the estimation data $\mathbf{z}$ is

$$\rho_{\mathcal{A}_\mathsf{r}\mathcal{B}_\mathsf{r}\mathcal{E}|\mathbf{t},\mathbf{z}} = \frac{1}{\mathrm{P}(\mathbf{z}|\mathbf{t})}\,\mathrm{tr}_{\mathcal{A}_\mathsf{e}\mathcal{B}_\mathsf{e}}\left[\rho_{\mathcal{A}_\mathsf{r}\mathcal{B}_\mathsf{r}\mathcal{A}_\mathsf{e}\mathcal{B}_\mathsf{e}\mathcal{E}|\mathbf{t}}\left(\prod_{j=1}^{n_\mathsf{e}} Q_{z_j}\right)\right] \ . \tag{64}$$

We start by proving the case when using the XOR seedless extractor, as described in Section 3.1. By using Theorem 2 and the function which defines the output length of the XOR extractor in our spot checking protocol, Equation (32), the left-hand side of Equation (61) can be upper bounded by

$$\begin{aligned}&\sum_{\mathbf{t}}\sum_{\mathbf{z}} \mathrm{P}(\mathbf{t},\mathbf{z})\left\|\rho_{\mathcal{K}\mathcal{E}|\mathbf{t},\mathbf{z}} - u_{\mathcal{K}|\mathbf{t},\mathbf{z}}\rho_{\mathcal{E}|\mathbf{t},\mathbf{z}}\right\|_1 \\ \leq &\sum_{\mathbf{t}}\sum_{\mathbf{z}} \mathrm{P}(\mathbf{t},\mathbf{z})\, m(\mathbf{t},\mathbf{z})\,\mathrm{tr}\left[\rho_{\mathcal{A}_\mathsf{r}\mathcal{B}_\mathsf{r}|\mathbf{t},\mathbf{z}}\prod_{i=1}^{n_\mathsf{r}} S_i\right] \ ,\end{aligned} \tag{65}$$

since, in the case $m(\mathbf{t},\mathbf{z}) = 0$, the term inside the trace norm is 0 and in the case $m(\mathbf{t},\mathbf{z}) = 1$, the error can be bounded by Theorem 2. Now, using the global state conditioned on a particular value of the estimation data $\mathbf{z}$ (64) and the facts that $\mathrm{P}(\mathbf{t},\mathbf{z}) = \mathrm{P}(\mathbf{t})\mathrm{P}(\mathbf{z}|\mathbf{t}) = p_\mathsf{e}^{n_\mathsf{e}}\, p_\mathsf{r}^{n_\mathsf{r}}\mathrm{P}(\mathbf{z}|\mathbf{t})$ and

$$m(\mathbf{t},\mathbf{z}) = \begin{cases} 1 & \text{if} \quad \max\limits_{\substack{s\in(2,2\sqrt{2})\\ \alpha_0,\alpha_1,\beta\in\mathbb{R}}} \sum_{j=1}^{n_\mathsf{e}} \alpha_{z_j} + (\beta-1)n_\mathsf{r} - 2\log_2(1/\epsilon) \geq 0 \\ 0 & \text{otherwise} \end{cases} \ , \tag{66}$$

is upper bounded by the exponential $\sqrt{2}^{\sum_{j=1}^{n_\mathsf{e}} \alpha_{z_j} + (\beta-1)n_\mathsf{r} - 2\log_2(1/\epsilon)}$, we bound Equation (65)

$$\begin{aligned}&\leq \sum_{\mathbf{t}}\sum_{\mathbf{z}} p_\mathsf{e}^{n_\mathsf{e}}\, p_\mathsf{r}^{n_\mathsf{r}}\sqrt{2}^{\sum_{j=1}^{n_\mathsf{e}} \alpha_{z_j} + (\beta-1)n_\mathsf{r}}\,\epsilon\,\mathrm{tr}\left[\rho_{\mathcal{A}_\mathsf{r}\mathcal{B}_\mathsf{r}\mathcal{A}_\mathsf{e}\mathcal{B}_\mathsf{e}|\mathbf{t}}\prod_{i=1}^{n_\mathsf{r}} S_i \prod_{j=1}^{n_\mathsf{e}} Q_{z_j}\right] \\ &= \epsilon\sum_{\mathbf{t}}\mathrm{tr}\left[\rho_{\mathcal{A}_\mathsf{r}\mathcal{B}_\mathsf{r}\mathcal{A}_\mathsf{e}\mathcal{B}_\mathsf{e}|\mathbf{t}}\prod_{i=1}^{n_\mathsf{r}}\left(p_\mathsf{r}\sqrt{2}^{\beta-1}S_i\right)\prod_{j=1}^{n_\mathsf{e}}\left(p_\mathsf{e}\left[\sqrt{2}^{\alpha_0}Q_{0_j} + \sqrt{2}^{\alpha_1}Q_{1_j}\right]\right)\right] \\ &= \epsilon\,\mathrm{tr}\left[\rho_{\mathcal{A}_\mathsf{g}\mathcal{B}_\mathsf{g}}\prod_{l=1}^{n}\left(p_\mathsf{r}\sqrt{2}^{\beta-1}S_l + p_\mathsf{e}\left[\sqrt{2}^{\alpha_0}Q_{0_l} + \sqrt{2}^{\alpha_1}Q_{1_l}\right]\right)\right] \ ,\end{aligned} \tag{67}$$

where we denote the global Hilbert spaces of Alice and Bob by $\mathcal{A}_\mathsf{g} = \mathcal{A}_\mathsf{r}\otimes\mathcal{A}_\mathsf{e}$ and $\mathcal{B}_\mathsf{g} = \mathcal{B}_\mathsf{r}\otimes\mathcal{B}_\mathsf{e}$. Finally, using the following identities

$$S = \mu_s\mathbb{1} - 4\nu_s\left(Q_0 - Q_1\right), \tag{68}$$

$$\mathbb{1} = Q_0 + Q_1 \ , \tag{69}$$

we can write each of the factors in Equation (67) as

$$\begin{aligned}&p_\mathsf{r}\sqrt{2}^{\beta-1}S + p_\mathsf{e}\left[\sqrt{2}^{\alpha_0}Q_0 + \sqrt{2}^{\alpha_1}Q_1\right] \\ &= p_\mathsf{r}\sqrt{2}^{\beta-1}\left[\mu_s(Q_0+Q_1) - 4\nu_s\left(Q_0 - Q_1\right)\right] + p_\mathsf{e}\left[\sqrt{2}^{\alpha_0}Q_0 + \sqrt{2}^{\alpha_1}Q_1\right] \\ &= \left[p_\mathsf{r}\sqrt{2}^{\beta-1}(\mu_s - 4\nu_s) + p_\mathsf{e}\sqrt{2}^{\alpha_0}\right]Q_0 + \left[p_\mathsf{r}\sqrt{2}^{\beta-1}(\mu_s + 4\nu_s) + p_\mathsf{e}\sqrt{2}^{\alpha_1}\right]Q_1 \\ &= Q_0 + Q_1 = \mathbb{1} \ ,\end{aligned} \tag{70}$$

where the penultimate equality follows form imposing conditions

$$p_{\mathsf{r}}\sqrt{2}^{\beta-1}\left(\mu_s - 4\nu_s\right) + p_{\mathsf{e}}\sqrt{2}^{\alpha_0} = 1 \ , \tag{71}$$

$$p_{\mathsf{r}}\sqrt{2}^{\beta-1}\left(\mu_s + 4\nu_s\right) + p_{\mathsf{e}}\sqrt{2}^{\alpha_1} = 1 \ , \tag{72}$$

expressed in Equation (33) and Equation (34). Substituting Equation (70) back in Equation (67) gives us the bound (61) and completes the proof.

Similarly, we make the same proof in the m-bit extractor function case, as described in Section 3.2. In this case, we introduce an indicator function

$$\mathrm{I}(m(\mathbf{t},\mathbf{z})) = \begin{cases} 1 & \text{if} \quad m(\mathbf{t},\mathbf{z}) > 0 \\ 0 & \text{otherwise} \ , \end{cases} \tag{73}$$

to encode the case when no output is produced and the error is 0. By using Theorem 4 and substituting the global state conditioned on the estimation data (64), the output length (39) and the indicator function (73), we can write the left-hand side of Equation (61) as follows

$$\begin{aligned}
&\sum_{\mathbf{t}}\sum_{\mathbf{z}} \mathrm{P}(\mathbf{t},\mathbf{z}) \left\| \rho_{\mathcal{K}\mathcal{E}|\mathbf{t},\mathbf{z}} - u_{\mathcal{K}|\mathbf{t},\mathbf{z}}\rho_{\mathcal{E}|\mathbf{t},\mathbf{z}} \right\|_1 \\
&\leq \sum_{\mathbf{t}}\sum_{\mathbf{z}} \mathrm{P}(\mathbf{t},\mathbf{z})\, \mathrm{I}(m(\mathbf{t},\mathbf{z}))\sqrt{2}^{m(\mathbf{t},\mathbf{z})-n_{\mathsf{r}}+4\log_2(n_{\mathsf{r}})} \operatorname{tr}\left[\rho_{\mathcal{A}_{\mathsf{r}}\mathcal{B}_{\mathsf{r}}|\mathbf{t},\mathbf{z}} \prod_{i=1}^{n_{\mathsf{r}}} (\mathbb{1} + S_i)\right] \\
&\leq \sum_{\mathbf{t}}\sum_{\mathbf{z}} \mathrm{P}(\mathbf{t},\mathbf{z})\, \sqrt{2}^{m(\mathbf{t},\mathbf{z})-n_{\mathsf{r}}+4\log_2(n_{\mathsf{r}})} \operatorname{tr}\left[\rho_{\mathcal{A}_{\mathsf{r}}\mathcal{B}_{\mathsf{r}}|\mathbf{t},\mathbf{z}} \prod_{i=1}^{n_{\mathsf{r}}} (\mathbb{1} + S_i)\right] \\
&= \sum_{\mathbf{t}}\sum_{\mathbf{z}} p_{\mathsf{e}}^{n_{\mathsf{e}}}\, p_{\mathsf{r}}^{n_{\mathsf{r}}} \sqrt{2}^{\sum_{j=1}^{n_{\mathsf{e}}}\alpha_{z_j}+(\beta-1)n_{\mathsf{r}}}\, \epsilon \operatorname{tr}\left[\rho_{\mathcal{A}_{\mathsf{r}}\mathcal{B}_{\mathsf{r}}\mathcal{A}_{\mathsf{e}}\mathcal{B}_{\mathsf{e}}|\mathbf{t}} \prod_{i=1}^{n_{\mathsf{r}}} (\mathbb{1} + S_i) \prod_{j=1}^{n_{\mathsf{e}}} Q_{z_j}\right] \\
&= \epsilon \sum_{\mathbf{t}} \operatorname{tr}\left[\rho_{\mathcal{A}_{\mathsf{r}}\mathcal{B}_{\mathsf{r}}\mathcal{A}_{\mathsf{e}}\mathcal{B}_{\mathsf{e}}|\mathbf{t}} \prod_{i=1}^{n_{\mathsf{r}}} \left(p_{\mathsf{r}}\sqrt{2}^{\beta-1}[\mathbb{1} + S_i]\right) \prod_{j=1}^{n_{\mathsf{e}}} \left(p_{\mathsf{e}}\left[\sqrt{2}^{\alpha_0}Q_{0_j} + \sqrt{2}^{\alpha_1}Q_{1_j}\right]\right)\right] \\
&= \epsilon \operatorname{tr}\left[\rho_{\mathcal{A}_{\mathsf{g}}\mathcal{B}_{\mathsf{g}}} \prod_{l=1}^{n} \left(p_{\mathsf{r}}\sqrt{2}^{\beta-1}\left[\mathbb{1} + S_l\right] + p_{\mathsf{e}}\left[\sqrt{2}^{\alpha_0}Q_{0_l} + \sqrt{2}^{\alpha_1}Q_{1_l}\right]\right)\right] \ ,
\end{aligned} \tag{74}$$

where we again denote the global Hilbert spaces of Alice and Bob by $\mathcal{A}_{\mathsf{g}} = \mathcal{A}_{\mathsf{r}} \otimes \mathcal{A}_{\mathsf{e}}$ and $\mathcal{B}_{\mathsf{g}} = \mathcal{B}_{\mathsf{r}} \otimes \mathcal{B}_{\mathsf{e}}$. Using the identities for S and $\mathbb{1}$ from Equation (68) and Equation (69), we can write each of the factors in Equation (74) as

$$\begin{aligned}
&p_{\mathsf{r}}\sqrt{2}^{\beta-1}\left[\mathbb{1} + S\right] + p_{\mathsf{e}}\left[\sqrt{2}^{\alpha_0}Q_0 + \sqrt{2}^{\alpha_1}Q_1\right] \\
&= p_{\mathsf{r}}\sqrt{2}^{\beta-1}\left[(1+\mu_s)(Q_0+Q_1) - 4\nu_s\left(Q_0 - Q_1\right)\right] + p_{\mathsf{e}}\left[\sqrt{2}^{\alpha_0}Q_0 + \sqrt{2}^{\alpha_1}Q_1\right] \\
&= \left[p_{\mathsf{r}}\sqrt{2}^{\beta-1}(1+\mu_s-4\nu_s) + p_{\mathsf{e}}\sqrt{2}^{\alpha_0}\right]Q_0 + \left[p_{\mathsf{r}}\sqrt{2}^{\beta-1}(1+\mu_s+4\nu_s) + p_{\mathsf{e}}\sqrt{2}^{\alpha_1}\right]Q_1 \\
&= Q_0 + Q_1 = \mathbb{1} \ ,
\end{aligned} \tag{75}$$

where the penultimate equality follows from the conditions expressed in Equation (40) and Equation (41). Substituting this back into Equation (74) gives us the bound Equation (61) and completes the proof. □