Device-independent Randomness Amplification and Privatization

Max Kessler∗1 and Rotem Arnon-Friedman†1

1Institute for Theoretical Physics, ETH-Zürich, CH-8093, Zürich, Switzerland

Abstract
Randomness is an essential resource in computer science. In most applications perfect, and sometimes
private, randomness is needed, while it is not even clear that such a resource exists. It is well known
that the tools of classical computer science do not allow us to create perfect and secret randomness
from a single weak public source. Quantum physics, on the other hand, allows for such a process, even
in the most paranoid cryptographic sense termed “quantum device-independent cryptography”. In this
work we propose and prove the security of a new device-independent protocol that takes any single
public Santha-Vazirani source as input and creates a secret close to uniform string in the presence of a
quantum adversary.

Our work is the first to achieve randomness amplification with all the following properties: (1) ampli-
ification and “privatization” of a public Santha-Vazirani source with arbitrary bias (2) the use of a device
with only two components (compared to polynomial number of components) (3) non-vanishing extraction
rate and (4) maximal noise tolerance. In particular, this implies that our protocol is the first protocol
that can possibly be implemented with reachable parameters. We are able to achieve these by combining
three new tools: a particular family of Bell inequalities, a proof technique to lower bound entropy in the
device-independent setting, and a special framework for quantum-proof multi-source extractors.

1 Introduction
Randomness is widely used in computer science; it is essential for cryptography and (at the least) beneficial
for many other scenarios, e.g., when designing efficient algorithms or proving the existence of certain functions
and combinatorial objects of interest, via the probabilistic method [V+12].

Unfortunately, we cannot know for sure that randomness even exists; it might as well be that everything in
nature is completely deterministic and fixed in advance. Furthermore, even if we assume the existence of some
sources of randomness in nature, it is not clear at all that there are sources of perfect randomness. Physical
sources of randomness, such as radioactive decay or thermal noise, can be used to produce unpredictable
bit strings, but those are usually partially biased and correlated bits. Even worse, how unpredictable these
sources of randomness are depends on the knowledge of the observer regarding the physical system. For a
person who can keep track of all microscopic degrees of freedom the outcomes can be completely predictable.

The question addressed in this work is familiar — can we reduce the amount of perfect randomness
required for one’s task of interest? In particular, we are interested in the cryptographic point of view. That
is, when we say perfect randomness, for example, we mean that it should be uniform even with respect to
some prior knowledge or side information of a malicious party or an adversary.1 We then ask:

**Question 1.** Can perfect randomness be created from weak or short randomness?

**Question 2.** Can private, secret, randomness be created from public randomness?

By weak randomness we mean that the produced bits can be correlated and biased (though not completely
deterministic). One such source, investigated in many works and of relevance for the current one, is the so
called “Santha-Vazirani source”, or SV-source, [SV84] — a source that produces a sequence of bits, where
each bit has some randomness given all previous ones. This source is a special type of the more general

∗kesslerm@student.ethz.ch
†rotema@itp.phys.ethz.ch

1This is the most demanding context to consider randomness in. A positive answer to the these questions in the cryptographic
sense also implies a positive answer in applications where a malicious party is not of interest. The opposite direction is, of
course, not true.
“min-entropy source” [CG88] (both defined formally below). Public randomness means that anyone can see the random string once it is produced. This is the case, for example, for the random numbers produced by the NIST randomness beacon; they are publicly available over the internet.

“Classical” computer science addresses the first question by considering pseudorandom generators and randomness extractors. Pseudorandom generators take a short perfectly random seed and generate from it a longer string of bits that no efficient algorithm can distinguish from a uniformly random string (see, e.g., [Gol10] for a survey). Thus, for the existence of pseudorandom generators we must make some assumptions regarding the complexity of certain computational tasks [DH76, Sha83]. Hence, they cannot be used when considering an all-powerful adversary.

Randomness extractors are functions that take a weak random source as an input and return an almost-uniform string as the output (see [NTS99]). Extractors are “information-theoretically secure” in the sense that, in contrast to pseudorandom generators, they do not require the use of computational assumptions. However, as widely known, no function can take a single SV-source and create close to uniform randomness out of it [SV84]. We therefore ought to consider extractors which either take an additional independent (short) random seed as input or several independent weak sources of randomness. These are called seeded extractors and multi-source extractors, respectively. (See [DPVR12, KK10] for examples of extractors that work even in the presence of a quantum adversary).

The answer to the second question seems obvious and intuitive — if everything is known in public (i.e., the initial source of randomness and the procedure, or protocol, used to manipulate it) then there is no way to create some private, secret, information out of it.

Quantum physics allows us to tackle the above questions from another angle and derive different conclusions, without making assumptions regarding computational complexity or the number of independent sources [BAK+16, AM16]. By preparing certain quantum states, e.g., a photon in a particular configuration, and measuring them one can generate perfectly random bits which, according to the laws of physics, were not known to anybody in advance.

Taking such an approach to answer the above questions is “unfair” and unsatisfactory. Firstly, one can argue that allowing the use of a source of, say, photons is like allowing the use of private unbiased coins. (And allowing the use of entangled photons, distributed among several parties, is like allowing shared randomness). Secondly, and significantly more importantly, when trying to implement such a source of randomness we find that creating perfect quantum states and measurements is practically impossible. In the cryptographic setting, imperfections and noise in the implementation are being exploited to gain information on the generated randomness [GLLL+11].

To solve these issues (and many others) the quantum cryptography community took one step further [ER14]. In the so called device-independent approach we let the adversary prepare the quantum devices used to generate the desired randomness. The honest parties interact with the device prepared by the adversary to test it and abort the protocol if its behaviour does not fit some predefined requirements. Then, the entire procedure is known to the adversary and there are no “hidden private coins”. Furthermore, we can no longer assume anything about the inner-workings of the device. Hence, if we are able to prove that the produced outcomes are secure to use, then the statement is inherently independent of the physical device and therefore robust to imperfections in the implementation.

In the device-independent scenario it might be that the adversary programmed the device to output a certain fixed string which is completely known to her. Thus, at first sight, it seems impossible to prove that the outputs are random from the perspective of the adversary. As known for quite some time now, the solution is to base device-independent protocols on the violation of Bell inequalities [Eke91, MY98, BHK05, AM16].

A Bell inequality [Bel64] can be thought of as a game played by the honest parties using a device that includes two non-communicating components (the most famous one being the CHSH inequality [CHSH69] or CHSH game; see [BCP+14] for a review on Bell inequalities and non-locality). The game has a special

2http://www.nist.gov/itl/csd/ct/nist_beacon.cfm

3Note, however, that quantum physics (as well as any other physical theory) cannot exclude the possibility that there is no randomness in nature to begin with. To prove that the outcome of a measurement performed on a quantum state is random, for example, we must first assume that we have the ability to choose the different states and measurements we would like to perform.
property — some quantum, non-local, strategies can win the game with probability $\omega_q$ greater than any classical, local, strategy. Hence, if the honest parties observe that using their device they win the game with probability $\omega_q$, they conclude it must be quantum (further details are given in Section 2.5). Otherwise they abort the protocol. Experiments have verified the quantum advantage in such “Bell games” in a loophole-free way [HBD+15, SMSC+15, GVW+15] (in particular, this means that the experiments were executed without making assumptions that could otherwise be exploited by the adversary in the cryptographic setting). It is well established that the higher the winning probability in a game is, the higher the amount of secret randomness which was produced in the process. We show this in Section 3 for our scenario of interest.

In this work we suggest a new quantum device-independent cryptographic protocol that uses a single public SV source as input and produces secret close to uniform randomness, even with respect to a quantum adversary. We state the concrete result and compare it to previous works in the following.

1.1 Results and contributions

We focus in this work on the amplification of an SV-source. An SV-source with bias $\mu \in (0, 0.5)$ has the following property: for each bit produced by the source $b_i$, $\Pr[b_i = 0|b_1, \ldots, b_{i-1}] \in [\mu, 1-\mu]$, where $b_1, \ldots, b_{i-1}$ are all the previous bits produced by the source. Such sources describe physical processes in which the bits are produced one after the other. Hence, the bias of each bit can depend (adversarially) on the previous bits, but not on the bits that will be produced in the future. Many of the processes in nature produce a sequence of bits, one bit after the other; the chronological order then implies that each bit can only depend on the past and not on the future. Thus, an SV-source can be used to describe such process in a realistic way.

The first challenge when dealing with randomness amplification is to find an interesting (and relevant) setting to consider and devise a protocol that can be proven secure in that setting. Previous works considered different protocols and there is no “standard model”.\footnote{Though it is always the case that some Bell game is repeated many times, as in all device-independent protocols (e.g., device-independent quantum key distribution and randomness expansion).} We first describe the scenario that we focus on and its relevance. Then we state our result and explain the main steps and ideas of the proof.

The setting that we consider is illustrated in Figure 1. We start with an arbitrary, public, SV-source with bias $\mu \in (0, 0.5)$. $\lambda$ denotes all the bits produced before the adversary, Eve, prepares the device for the honest party, Alice. $\lambda$ can also include any other piece of classical information from the past that might be of relevance to Eve. Eve then creates the device, denoted by the black box in the figure, depending on $\lambda$. She can keep quantum side information $E = E(\lambda)$ correlated with the device for herself; this side information can later be used by Eve to gain information about the final random string. Once Alice holds the device she can use it together with additional bits produced by the source, $I$ and $Z$ in the figure, to create her final secret random string $K$.

The SV-source can be controlled by an untrusted party but we assume that every bit, when produced, has some randomness conditioned on all side information. Mathematically, for the first bit of $I$, $I_1$, for example, we have $\frac{1}{2} - \mu \leq \Pr[I_1 = 0|\lambda] \leq \frac{1}{2} + \mu$.

In particular, in the above explained scenario it holds that, given the history $\lambda$ and Eve’s knowledge $E$, the device $D$ and the sequence of bits $I \circ Z$ are independent. That is,\footnote{This should be understood on the intuitive level, as we did not define the device $D$ in a mathematical way. The exact setting is modelled formally in Section 4.1.}

$$I(D: I \circ Z|\lambda E) = 0 ,$$

where $I(\bullet: \bullet|\bullet)$ is the conditional mutual information.

We remark that the considered scenario is relevant for actual implementations of randomness amplification protocols: the chronological order of events is such that Eve can prepare the device depending only on past information (the history) but not on the bits which will be produced after delivering the device to Alice. This implies that all correlations between the following bits produced by the source and the device are due to past events and Eve’s side information. Thus, Equation (1) holds. Several previous works, e.g., [CR12, GMD+13, BRG+16], considered similar settings as well.
Figure 1: An illustration of the considered setting. We start with a public SV source and a device which was created by the adversary (the black box in the figure). The goal is to produce a secret, close to uniform, string $K$. The bits produced by the SV-source when running the protocol, $I$ and $Z$, and the device can be correlated via the previous bits of the source, $\lambda$, and the adversary $E$. Our protocol is such that the honest party first uses some of the bits, $I$, as input to the device. The output of the device is denoted by $O$. Then, a special type of randomness extractor is applied to $O$ and additional bits $Z$ from the source. The result is the output randomness $K$.

The main contribution of our work is a construction of a device-independent randomness amplification protocol that uses a single public SV-source to create secret and close to uniform randomness, with respect to all of the knowledge that the adversary has:

**Theorem 3 (Informal).** Given any public SV-source with bias $\mu \in (0, 0.5)$ there exists a protocol, requiring a two-component device, such that:

1. (Soundness) For any device $D$ used to implement the protocol such that Equation (1) holds, either the protocol aborts with overwhelming probability or an $\varepsilon$-close to uniform (given the adversary’s knowledge) string $K$ is produced.

2. (Completeness) There exists an honest implementation of the device such that the protocol aborts with negligible probability when using this device, even in the presence of noise.

The formal statement is given in Theorem 30. The soundness, or security, parameter $\varepsilon$ depends on the bias of the source, $\mu$, as well as the parameters of an extractor used in our protocol to create $K$. For certain choices of parameters the protocol can be made explicit.

Theorem 3 improves upon the prior state-of-the-art in several significant aspects (see Section 1.3 and Table 1 for comparison with previous works):

1. **Device requirement** – we only require that the device includes two components (the lowest possible), compared to a polynomial number in previous works that considered a public weak source of randomness.

This means that the black box in Figure 1 consists of two separated parts.\(^6\) Having two components is

\[^6\]One can imagine the two components as being two computers or, alternatively, two provers in a multi-prover interactive proof system.
a necessary requirement for protocols based on Bell inequalities. As we explain in Section 1.3, previous works that considered a public weak source had to use, at the least, polynomial number of components, which is not realistic. Other works that allowed a constant number of devices could not derive a result for an arbitrary bias $\mu$, a public SV-source, and/or quantum adversaries.

2. Extraction rate (efficiency) – for a large range of parameters we can extract a linear number of bits\(^7\) while maintaining cryptographic security level, compared to a vanishing extraction rate in previous works that considered a public weak source of randomness.

Using an extractor with sufficiently good parameters $\varepsilon$ can be made exponentially small in the number of bits taken from the SV-source while extracting a linear number of bits. Previous works could not achieve this, independently of the extractor used in the protocol.

3. Robustness – we are able to tolerate the maximal amount of noise, compared to low noise levels in previous works that considered a public weak source of randomness.

The completeness statement holds for any amount of noise in the implementation which still results in a violation of the Bell inequality.\(^8\) This is the maximal possible amount one can hope to tolerate.

Apart from randomness amplification, our protocol can also be used as a main building block for device-independent randomness expansion and key distribution using weak sources of randomness. More details are given in Section 5.

Theorem 3 cannot be derived by improving previously known techniques (as explained in Section 1.3). To prove it we present a completely new proof, which can be of independent interest. Our proof uses three different tools which were developed recently and were not used before in the context of randomness amplification. One particular example for an independent technical contribution is the proof given in Section 3, where we investigate a new type of Bell inequalities and show, for the first time, that they can also be used in a cryptographic setting. Another contribution is presenting a first application of a special type of extractors that were recently introduced in [AFPS16]. The existence of such extractors is what allows us to produce randomness, in the presence of a quantum adversary, when starting with a single public SV-source.

1.2 Main steps in the proof

Our protocol is stated as Protocol 2 in Section 4.3. The protocol is simple: the device is used sequentially with the inputs $I$ from the SV-source to create the outputs $O$. Once all the outputs are produced Alice calculates the average violation of a specific Bell inequality from the raw data and aborts if the violation is not sufficiently high. If she does not abort then a special type of extractor is applied to $O$ together with additional bits from the source $Z$.

Step 1: Choosing the “correct” Bell inequality

As all device-independent protocols, our protocol is based on the violation of a given Bell inequality above a certain threshold. This way Alice can make sure that the device implements a quantum non-local strategy. All previous protocols use the CHSH Bell inequality or other well known inequalities.

We use a recently developed family of Bell inequalities (with two parties, two inputs, and two outputs) which fits perfectly to the scenario of randomness amplification. As explained above, in our setting, the device

\(^7\)To be more precise – for a large range of parameters (the full details are given in Remark 37) there is an explicit extractor that can be used in our protocol to extract a linear number of bits. If one is interested in an explicit protocol for all parameters, there are two options: 1) A simple modification of our protocol, which requires the use of a device with 4 components, can be used to extract a sub-linear number of bits using a three-source extractor. (A similar thing was previously done in [BRG+16, Theorem 2] but the resulting protocol requires 8 components and the security proof uses an additional assumption of a private SV-source; see Section 1.3). 2) Using the current protocol (with only two components) one can extract a logarithmic number of bits. If, in the future, new (classical) two-source extractors with better parameters are developed, they can be used in our protocol to achieve better extraction rates without modifying the protocol or its security proof.

\(^8\)This can be seen, for example, from Figure 4 below which shows that non-zero entropy can be certified as long as there is a violation of the Bell inequality.
Figure 2: Correlations between the device and the inputs. The two components of the device are denoted by the black boxes. The inputs to the two components, $X$ and $Y$, come from the SV-source. The outputs are denoted by $A$ and $B$. The device and the inputs can be correlated via the history $\lambda$, as denoted by the dashed arrows. A violation of an MDL inequality certifies that the device cannot be classical in this setting.

and the inputs $I$ can be correlated via $\lambda$. The Bell inequalities developed in [PRB$^+$14], called “measurement dependent locality (MDL) inequalities”, are adapted to the situation illustrated in Figure 2 for any bias of the source. They therefore accommodate the dependency between the device and the side information. In contrast, the violation of the CHSH inequality cannot be used to “verify quantumness” above some threshold for the bias (see further details in Section 2.6). Other Bell inequalities which were used in the context of randomness amplification and allowed for an arbitrary bias of the SV-source require a device with more than two components [GMD$^+$13, BRG$^+$16, RBH$^+$15].

We note that, for the completeness of our protocol, it is crucial that for any bias of the source there is a quantum strategy (i.e., quantum state and measurements) that violate the inequality. This is indeed the case as shown in [PRB$^+$14]. When proving completeness we also explain how the maximal violation within quantum physics can be found numerically.

The rest of the steps in the proof deal with the soundness proof.

**Step 2: Certifying randomness from the MDL violation after a single use of the device**

The analysis done in [PRB$^+$14] for the MDL inequalities only ensures that a violation of the inequality implies that the device must be non-local, i.e., it cannot be implemented by a classical strategy. While this is important for the study of fundamental questions in physics, it is not sufficient in the cryptographic setting. A quantitative bound on how random the output of the device must look to an adversary was missing.

The first part of our proof is devoted to deriving a relation between the violation of the MDL inequality and the amount of knowledge Eve can gain regarding the output in a single use of the device. Specifically, we prove a lower-bound on the von Neumann entropy of the output given all side information:

$$H(O_i|I_i,E,\lambda) \geq t,$$

where $I_i$ and $O_i$ are the inputs and outputs when using the device for the $i$’th time and $t \geq 0$ depends on the bias of the source and the observed violation of the MDL inequality (see Lemma 27 for the exact bound and Figure 4 for a plot). The conditional von Neumann entropy is just one way of quantifying the amount of secret randomness, but as we will show below, this is the relevant quantity for us.

A bound similar to Equation (2), but for the CHSH inequality, was proven in [PAB$^+$09]. In the case of the CHSH inequality the inputs are assumed to be chosen uniformly and independently of the device and hence one cannot use the result of [PAB$^+$09] directly for randomness amplification. We find a way to connect the two scenarios and derive a bound as in Equation (2) for the MDL inequality from that of the CHSH inequality.

The resulting bound is non-trivial as long as the MDL inequality is violated (while if there is no violation the conditional entropy must be 0, since the device might be using a classical strategy). Combined with the
following step, this property allows us to tolerate maximal amount of noise in the honest implementation of the device used in the protocol.

**Step 3: Bounding the total amount of min-entropy after multiple uses of the device**

To bound the amount of extractable randomness from the outputs of the device $O$ we need to lower bound the total conditional smooth min-entropy $H_{\min}^{\varepsilon_s}(O|IE,\lambda)$, for $\varepsilon_s \in (0,1)$, given that our protocol did not abort.

If the different uses of the device in the protocol were independent and identical, getting a bound on $H_{\min}^{\varepsilon_s}(O|IE,\lambda)$ is rather easy. On the intuitive level, the total amount of entropy in that case is the sum of the entropies in each round of the protocol [TCR09, DW05]. However, as the adversary is the one preparing the device, there is no reason to believe that the device behaves in an independent and identical way. The analysis is therefore more delicate.

To overcome this difficulty we use a new information-theoretic tool, called the entropy accumulation theorem [DFR16], to bound the total amount of smooth min-entropy, in a sequential processes, using the von Neumann entropy of a single step of the process. More precisely, we use the framework developed in [AFRV16] for proving security of device-independent cryptographic protocols using the entropy accumulation theorem. In [AFRV16] the entropy accumulation theorem was used to prove security of device-independent key distribution and randomness expansion protocols. We adapt the different steps to our scenario of randomness amplification with the MDL inequalities.

To prove a lower bound on $H_{\min}^{\varepsilon_s}(O|IE,\lambda)$ we start by showing that for any SV-source and device, the sequential process defined by the rounds of our protocol and the actions of the device fulfill the prerequisites of the entropy accumulation theorem. Next, using Equation (2) we devise a “min-tradeoff function”. This function quantifies the “worst-case von Neumann entropy” that is accumulated in a single round of the protocol, while taking into account the observed violation of the MDL inequality. Once this function is constructed we can apply the techniques of [DFR16, AFRV16] to derive a bound on $H_{\min}^{\varepsilon_s}(O|IE,\lambda)$. The first order term of the lower bound on $H_{\min}^{\varepsilon_s}(O|IE,\lambda)$ is $nH(O_i|I_iE,\lambda)$, where $n$ is the number of rounds of the protocol. That is, $H_{\min}^{\varepsilon_s}(O|IE,\lambda) \in \Omega(n)$, which is optimal. For more details, see Section 4.5.

**Step 4: Extracting the randomness**

Once a bound on the conditional smooth min-entropy is derived we need to extract the randomness using an extractor. However, since only a single SV-source is available, there is no additional independent source of randomness. Thus, standard seeded or multi-source extractors cannot be used.

In the last step of our proof we show that the setting that we consider (as in Figure 1 above) implies that a newly developed model for quantum-proof multi-source extractors can be used [AFPS16]. The model presented in [AFPS16], termed the “Markov model”, deals with extraction from multiple weak sources which are independent only given some side information, possibly quantum. Each of the sources must have sufficient amount of entropy conditioned on that side information. It was proven in [AFPS16] that any (strong) multi-source extractor is also a (strong) quantum-proof multi-source extractor in the Markov model, with some loss in parameters (the exact statements which we use are presented in Section 2.8).

We show that the considered setting implies that

$$I(O : Z|IE,\lambda) = 0,$$

meaning that given $I$, $E$, and $\lambda$, $O$ and $Z$ are independent. Furthermore, the previous step of our proof ensures that $O$ has sufficient amount of entropy conditioned on $IE\lambda$. The same is true for $Z$ since it is taken directly from the SV-source. We can therefore use a strong quantum-proof two source extractor in the Markov model to create the final string $K = \text{Ext}(O,Z)$, which is close to uniform even given $ZIE\lambda$. This implies the security of our protocol.

---

9The smooth min-entropy is a standard quantity related to the, more commonly known, min-entropy; the formal definition is given in Section 2.3. The important thing to know at this stage is that it tightly determines how much randomness Alice can extract from $O$ in the presence of a quantum adversary [KRS09].
The use of this special type of extractors [AFPS16] is what allows us to start with nothing but a single public SV-source and consider quantum side-information. Previous models for quantum-proof multi-source extractors [KK10, CLW14] do not allow for the side information considered in the current setting. Moreover, a *strong* extractor is crucial here since the seed $Z$ is public (as it comes from the public SV-source).

We remark that $I$ and $Z$ cannot be used directly as the sources for the extractors, although they both have high min-entropy given $\lambda$ and $E$. The reason is that they are not independent given $\lambda E$. The use of the device is therefore necessary in order to create a string $O$ which is “decoupled” from $Z$.

The combination of all the steps above proves the soundness of our protocol.

### 1.3 Previous works

We now discuss the different works and assumptions and compare them to the current work. See also Table 1.

#### Public SV-source

Colbeck and Renner were the first to consider the task of randomness amplification [CR12] and give a “proof of concept”. There, the relation between the knowledge that an adversary has about a final single bit was bounded using the expected Bell violation. They showed that using a public SV-source with bounded bias ($\mu = 0.058$) and a two-component device a single close to uniform bit can be created in the presence of both quantum and non-signalling (super quantum) adversaries. The number of measurements, however, grew with their security parameter and only one bit was produced. Hence any protocol based on such approach would have resulted in a vanishing extraction rate.

Following that, [GMD+13] improved on the above result by considering a protocol that can accommodate arbitrary bias of the SV-source and tolerate some noise. Instead of restricting the analysis to quantum adversaries [GMD+13] focused on the stronger non-signalling adversaries. Unfortunately, the protocol required the use of many devices — polynomial in the number of bits used from the source. One can imagine this as requiring a polynomial number of laboratories separated in space, each of which runs a quantum experiment. This is of course unrealistic in any implementation.

To see why the proof technique of [GMD+13] could not be extended to get results similar to ours note the following. First, to deal with an arbitrary bias of the SV-source a 5-party Bell inequality was used. This implies that any protocol based on their Bell inequality would require, at the least, 5 devices (otherwise the violation is meaningless). Second, the final randomness is extracted using a deterministic process, which is only possible since their protocol requires a polynomial number of devices (for details see the discussion in [GMD+13, Supplementary information C]). To reduce the number of devices one would have to construct strong randomness extractors which are secure in the presence of non-signalling adversaries, but there are indications that such do not exist [AFTS12].

#### Private SV-source

In [BRG+16, RBH+15] a protocol using a constant number of devices was constructed, also when considering non-signalling adversaries. In addition, as in our work, the protocol is robust to noise and achieves a non-zero extraction rate. The crucial difference between [BRG+16, RBH+15] and the current work is that the security proof of [BRG+16, RBH+15] assumes that the SV-source must be private, i.e., no information about the bits produced by the source can leak to the adversary at any point (also after the end of the protocol).

One might argue that this is not such a strong requirement, especially since we anyhow assume that the final randomness created by the protocol is kept secret. However, there is one critical difference: it is implied by the security definition of randomness amplification protocols (sometimes termed composable; see Section 4.2) that if part of the produced randomness is leaked to the adversary the rest of the bits are still close to uniform. In contrast, when proving security with a private source it is not clear at all what happens when some information about the source is leaked to the adversary. It is nowhere proven (or conjectured) that if partial information about the used source (even a single bit) is leaked the entropy of the produced string remains somewhat high.
Table 1: Comparison of the different works. Q and NS stand for a quantum and non-signalling (super-quantum) adversary respectively. The number of devices is with respect to the number of bits used from the weak source of randomness. For a more detailed comparison of previous works see also [BRG+16, Supplementary Information] and [AM16, Table 1].

The proof of [BRG+16, RBH+15] cannot be used to get a protocol which can take a public SV-source as input. The reason is that the assumption regarding the privacy of the source is used in order to simplify the security criterion and argue that a classical multi-source extractor can be used to extract the randomness, although a non-signalling adversary is present. To allow for a public source one will need a strong multi-source extractor which is secure in the presence of a non-signalling adversary, but as mentioned above it is not clear that such exists.

We also remark that the simplification of the security definition to a classical one due to the use of private source enabled the analysis of the total amount of min-entropy in the outputs of the device. The same analysis cannot be used as is when considering the case of a public source or when trying to bound the smooth min-entropy as we do here. Moreover, in [BRG+16, RBH+15] as well, Bell inequalities with more than two parties are used. Thus, such protocols cannot lead to a protocol that requires only two components as ours.

Public min-entropy source

In two more recent works [CSW14, CSW] a protocol that can amplify a public min-entropy source was suggested and its security was proven. [CSW14] assumed a quantum adversary while [CSW] considered a non-signalling one. The first part of the protocol in these works takes the min-entropy source and extracts blocks of bits, some of them close to uniform with respect to the used devices, by enumerating all possible seeds. The different blocks are then used as inputs to a randomness expansion protocol [MS16]. This approach leads to a polynomial number of devices in [CSW14] and exponential in [CSW]. Furthermore, in both works the security parameter is inverse polynomial in the number of bits used from the source, the efficiency of the protocols vanishes, and the amount of tolerated noise is low.

A min-entropy source is of course more general than the SV-source considered in the current work. Our work cannot be extended as is to the case of a min-entropy source. On the other hand, it is also not clear how to take the work of [CSW14, CSW] and decrease the number of devices – to get close to uniform inputs for the randomness expansion protocol starting with a single weak source one must enumerate the seeds; each seed should then be used while running the protocol on a different set of devices. The number of devices (and hence also the zero extraction rate) is thus a fundamental part in the proof technique of [CSW14, CSW].

Source-device-adversary model

In [CSW14, CSW] the authors model the relation between the source, the adversary, and the device differently than what we do here. In particular, they allow for some quantum side information about the source, in contrast to our $\lambda$ which is classical. In all other mentioned works the assumptions regarding the relation between the three components are similar to the ones considered here (though not mentioned explicitly in the same way). In [WBG+16] a different scenario is considered, but the security analysis is not complete and only restricted SV-sources can be amplified.
Organisation of the paper. We start in Section 2 with some preliminaries. In particular, the necessary information regarding the MDL inequalities and two-source extractors in the Markov model is given. Section 3 is devoted to proving a relation between the observed violation of an MDL inequality and the knowledge that a quantum adversary can gain about the output of the device. In Section 4 we state our randomness amplification protocol and prove its security. We end in Section 5 with some open questions.

2 Preliminaries

2.1 Notation

In the following we will denote by

- capital letters classical registers (i.e., random variables) and quantum registers.
- a subscript register, e.g. $X_i$, a single register with label $i$ and a superscript register, e.g. $X^i$, the sequence of registers with labels up to $i$, i.e., $X^i = X_1...X_i$.
- the operator $\oplus$ addition modulo 2, sometimes also called the XOR operation.
- $P_A$ the set of probability distributions over an alphabet $A$.

2.2 Quantum mechanics

We introduce the concepts of quantum mechanics that we use throughout our work. For a more detailed view on quantum mechanics in quantum information theory we refer to Nielsen and Chuang [NC10].

A state of a quantum mechanical system can generally be described by density operators.

**Definition 4** (Density operator). A density operator $\rho$ on a Hilbert space $H$ is a normalized positive operator on $H$, i.e., $\rho \geq 0$ and $\text{Tr} \rho = 1$. A density operator is said to be pure if it has the form $\rho = |\psi\rangle\langle \psi|$, where, using Dirac notation, $|\psi\rangle \in H$.

A bipartite quantum state on two Hilbert spaces $H_A$ and $H_B$ is described by a density operator $\rho_{AB}$ on the Hilbert space $H_A \otimes H_B$. If we want to recover the state on $H_A$ alone we take the partial trace,

$$\rho_A = \text{Tr}_B(\rho_{AB}) = \sum_b (\text{id}_A \otimes \langle b|)\rho_{AB}(\text{id}_A \otimes |b\rangle),$$

where $\{|b\rangle\}$ is an orthonormal basis (ONB) on $H_B$.

Some special density operators are given in the following.

(i) The density operator is said to be fully mixed if $\rho = \frac{1}{d}\text{id}$, where $d = \text{dim}(H)$.

(ii) The density operator $\rho_{XA}$ is said to be a classical-quantum state (cq-state) if $\rho = \sum_{i=1}^d p_i |i\rangle\langle i| \otimes \rho_A^i$, where $\{|i\rangle\}$ is an ONB on a $d$-dimensional Hilbert space and $\sum_{i=1}^d p_i = 1$ with $p_i \geq 0 \forall i$. The notion can be extended to an arbitrary amount of classical registers.

We describe the evolution of a quantum state by completely positive trace preserving (CPTP) maps.

**Definition 5** (CPTP map). A linear map $E \in \text{Hom}(\text{End}(H_A), \text{End}(H_B))$ is said to be trace preserving if, for any $\rho \in \text{End}(H_A)$,

$$\text{Tr}(E(\rho)) = \text{Tr}(\rho).$$

The map $E$ is said to be completely positive if, for any $\rho_{AR} \in \text{End}(H_A \otimes H_R)$ and $\rho_{AR} \geq 0$,

$$(E \otimes \text{Id}_R)(\rho_{AR}) \geq 0,$$

where $H_R$ is any additional Hilbert space and $\text{Id}_R$ is the identity map on that Hilbert space.

When talking about the closeness of quantum states we quantify it by the trace distance which describes how well two states can be distinguished.

**Definition 6** (Trace distance). The trace distance between two density operators $\rho$ and $\sigma$ on a Hilbert space $H$ is defined as

$$\delta(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1 = \frac{1}{2} \text{Tr} \left( \sqrt{(\rho - \sigma)^\dagger(\rho - \sigma)} \right).$$
2.3 Entropies and Markov chains

**Entropies** We make use of the Shannon entropy for classical random variables [Sha48] and its quantum equivalent, the von Neumann entropy [Neu27]. The conditional Shannon entropy is defined as follows.

**Definition 7** (Shannon entropy). Let $X, Y$ be discrete random variables over the alphabets $\mathcal{X}, \mathcal{Y}$ distributed according to the probability distribution $P_{XY}$. Then the conditional Shannon entropy is defined as

$$H(X|Y) = -\sum_{x\in \mathcal{X}, y\in \mathcal{Y}} P_{XY}(x, y) \log_2 P_{X|Y=y}(x).$$

Its quantum equivalent, the von Neumann entropy, is defined for a quantum state $\rho_{AE}$.

**Definition 8** (von Neumann entropy). Let $\mathcal{H}_A$ and $\mathcal{H}_E$ be two Hilbert spaces and $\rho_{AE}$ a quantum state on $\mathcal{H}_A \otimes \mathcal{H}_E$. Then the von Neumann entropy is defined as

$$H(AE) = -\text{Tr}(\rho_{AE} \log \rho_{AE}).$$

Furthermore, the conditional von Neumann entropy is defined as

$$H(A|E) = H(AE) - H(E).$$

Let $\mathcal{H}_A$ and $\mathcal{H}_E$ be two Hilbert spaces and $\rho_{AE} = \sum_a |a\rangle\langle a| \otimes \rho_E^n$ a classical quantum state on $\mathcal{H}_A \otimes \mathcal{H}_E$. Then the conditional min-entropy is defined as

$$H_{\min}(A|E) = -\log p_{\text{guess}}(A|E),$$

where $p_{\text{guess}}(A|E)$ is the maximal probability of guessing $A$ given the quantum system $E$.

The maximization ranges over all sets of POVMs $\{M_a^E\}_a$ on $E$.

The smooth min-entropy is a smoothed version of the min-entropy, meaning it is the maximum of the smooth min-entropy in an $\varepsilon$-neighbourhood around the probability distribution or quantum state.

**Definition 10** (Smooth min-entropy). Let $\mathcal{H}_A$ and $\mathcal{H}_E$ be two Hilbert spaces and $\rho_{AE} = \sum_a |a\rangle\langle a| \otimes \rho_E^n$ a classical quantum state on $\mathcal{H}_A \otimes \mathcal{H}_E$. Then the conditional smooth min-entropy is defined as

$$H_{\min}^\varepsilon(A|E) = \max_{\sigma_{AE} \in \mathcal{B}^\varepsilon(\rho_{AE})} H_{\min}(A|E)_{\sigma_{AE}},$$

where $\mathcal{B}^\varepsilon(\rho_{AE})$ is the set of sub-normalised states that are at most $\varepsilon$ away from $\rho_{AE}$ in terms of purified distance (see [TCR10]).

When the quantum state is clear from the context we drop the subscript of the entropies and simply write $H(A|E)$ instead of $H(A|E)_{\rho_{AE}}$.

The mutual information $I(X : Y|Z)$ quantifies the common information of $X$ and $Y$, given $Z$ and can be described as a function of the entropies of the parts.

**Definition 11** (Mutual information). Let $X, Y$ and $Z$ be random variables. Then the Shannon mutual information is defined as

$$I(X : Y|Z) = H(X|Z) - H(X|YZ).$$

In the quantum case, let $\rho_{XYZ}$ be a quantum state. Then the quantum mutual information $s$ defined as

$$I(X : Y|Z) = H(X|Z) - H(X|YZ).$$

**Definition 12** (Markov chain). A set of random variables $X, Y, X$, or a tripartite quantum state $\rho_{XYZ}$, is said to form a (quantum) Markov chain, denoted by $X \leftrightarrow Y \leftrightarrow Z$, if the conditional mutual information $I(X : Z|Y)$ vanishes.
2.4 Weak sources of randomness

We consider two classes of weak random sources, an SV source and a min-entropy source. The SV source was first introduced by Santha and Vazirani [SV84]. Formally an SV source is defined as follows.

**Definition 13** ($\mu$-SV source, [SV84]). Let $S$ be any source producing a sequence of binary random variables $X_i$ that can depend on some side information $\lambda$. Then, for any $\mu \in (0, \frac{1}{2})$, $S$ is called an $\mu$-SV source if the random variables $X_i$ are distributed according to some probability distribution $P_{X_i|X^{i-1},\lambda}$ that depends on $\lambda$ and satisfies

$$1 - \frac{1}{2} \mu \leq P_{X_i|X^{i-1},\lambda}(0|x^{i-1}) \leq 1 + \mu \quad \forall i, x^{i-1}.$$  \hspace{1cm} (3)

We see that an SV source produces bits that are all, to some extent, random, even given the previous bits and some possible side information.

An MDL source produces two bits at a time and bounds the probability of each outcome in a similar way as the SV source.

**Definition 14** ($\mu$-MDL source, [PRB+14]). Let $S$ be any source producing binary random variables $X_i$ and $Y_i$ that can depend on some side information $\lambda$. Then, for any $\mu = \{\mu_{\text{min}}, \mu_{\text{max}}\} \in (0, \frac{1}{2}) \times (\frac{1}{2}, 1)$, $S$ is called a $\mu$-MDL source if the outputs are distributed according to some probability distribution $P_{X_i,Y_i|X^{i-1}Y^{i-1},\lambda}$ that depends on $\lambda$ and satisfies

$$\mu_{\text{min}} \leq P_{X_i,Y_i|X^{i-1}Y^{i-1},\lambda}(x,y|x^{i-1}y^{i-1}) \leq \mu_{\text{max}} \quad \forall x^i, y^i.$$  \hspace{1cm} (4)

In our work we use the notation of MDL sources. These are directly related to the SV sources as shown below.

**Lemma 15.** For all $0 \leq \mu \leq 1/2$ a $\mu$-SV source is a $\left\{ \left( \frac{1}{2} - \mu \right)^2, \left( \frac{1}{2} + \mu \right)^2 \right\}$-MDL source.

**Proof.** Employing the definition of conditional probabilities $P_{X_i|X^{i-1},\lambda} = \frac{P_{X_i\lambda}}{P_{X^{i-1}\lambda}}$ and $P_{X_{i+1}|X^i,\lambda} = \frac{P_{X_{i+1}\lambda}}{P_{X^i\lambda}}$ we find $P_{X_{i+1},X_{i}|X^{i-1},\lambda} = \frac{P_{X_{i+1},X_{i}\lambda}}{P_{X^{i-1}\lambda}} = P_{X_i|X^{i-1},\lambda}P_{X_{i+1}|X^i,\lambda}$. From that it follows immediately that the constraints for two consecutive outputs of the source are

$$\left( \frac{1}{2} - \mu \right)^2 \leq P_{X_{i+1},X_i|X^{i-1},\lambda}(x_{i+1}x_i|x^{i-1}) \leq \left( \frac{1}{2} + \mu \right)^2 \quad \forall i, x^{i+1}.$$  \hspace{1cm} (5)

Choosing $\mu_{\text{min}} = \left( \frac{1}{2} - \mu \right)^2$ and $\mu_{\text{max}} = \left( \frac{1}{2} + \mu \right)^2$ this satisfies Definition 14 of an MDL source. \hfill $\square$

Finally a min-entropy source is a source that produces a bit string that has a min-entropy which is lower bounded by some constant.

**Definition 16** ($k$-min-entropy source, [CG88]). Let $S$ be any source producing a sequence of binary random variables $X_i$ that can depend on some side information $\lambda$. Furthermore let $n$ be the arbitrary length of that sequence. Then $S$ is said to be a $k$-min-entropy source if the min-entropy of the bit string conditioned on the side information is lower bounded by $k$, i.e., $H_{\text{min}}(X^n|\lambda) \geq k$.

It is worthwhile noticing that any SV source can also be used as a min-entropy source. The reversed implication, however, is not true, since in an SV source each new bit must contain a minimal amount of randomness. In this sense the output of the SV source has more structure.

With regards to randomness amplification it has been shown by Santha and Vazirani [SV84] that, classically, a single SV-source, private or public, cannot be amplified. If one has access to two or more independent sources of which at least one is private, one can extract randomness from them using a randomness extractor. However, if all the sources are public this is still not possible.
2.5 Non-local games and Bell inequalities

Non-local games. During a non-local game two players, Alice and Bob, are given questions by a verifier and have to return answers. Both the questions and answers can be described simply as numbers. The questions, $x$ and $y$, are taken from alphabets (we restrict ourselves to binary alphabets) $\mathcal{X}$ and $\mathcal{Y}$, and distributed according to some probability distribution $P_{XY}$. Similarly, the answers, $a$ and $b$, can be chosen from (binary) alphabets $\mathcal{A}$ and $\mathcal{B}$. Alice and Bob win a round of the game if the questions and answers satisfy a previously defined condition. Formally we can think of a function $w : \mathcal{X} \times \mathcal{Y} \times \mathcal{A} \times \mathcal{B} \rightarrow \mathcal{W}$, where $\mathcal{W}$ is the set describing the outcome of the game.

In order to win the game with the highest probability Alice and Bob can, before the game starts, choose a strategy. After the game starts they are no longer allowed to communicate. The rules of the game are that the players are not allowed to communicate, one player does not know the other player’s question, and that the players cannot influence the questions they are asked.

In terms of strategy we distinguish two classes, the first one being classical local hidden variable (LHV)/shared randomness strategies. In an LHV strategy, Alice and Bob share some common information $\lambda$ and, according to the common information, choose their answers deterministically. The second class of strategies are quantum strategies. Using a quantum strategy, Alice and Bob can share a multipartite quantum state. They can then use the questions to choose measurements that are done on the quantum state. The results of these measurements can then be used to produce answers for the questions.

It can be shown that quantum strategies are strictly more powerful than LHV strategies. Namely, for some non-local games, there exist quantum strategies that achieve a winning probability that is higher than any LHV strategy can achieve. We call the probability distributions $P_{ABXY}$, of the questions and answers, that characterise these strategies non-local statistics.

Using this fact that strategies producing non-local statistics are more powerful than LHV strategies, we can certify quantumness using non-local games. We can do this by analysing the winning probability of the strategy in the game. If the winning probability is higher than the threshold for any LHV strategy we can conclude that Alice and Bob must have used a quantum strategy.

Bell inequalities. An equivalent description of non-local games are Bell inequalities. In this scenario we consider Bell experiments; i.e., experiments where we have two devices that take inputs (the questions) and produce outputs (the answers). The probability distributions over the inputs and outputs can, similar to the case of non-local games, be divided into LHV statistics and quantum statistics. However, the winning probability is replaced by a Bell parameter, a general function of the probability distribution, $f(P_{ABXY})$. The Bell inequality is then a constraint on the Bell parameter that is satisfied by all LHV statistics. A Bell inequality could for example look as follows

$$f(P_{ABXY}) \leq f_{\text{LHV}}.$$ 

In the Bell experiments we consider some hidden side information $\lambda$. The assumptions that we make about the setting are that firstly, given the inputs and the side information, the outputs do not depend on each other. Secondly, we assume that, given $x$ and $\lambda$, $a$ does not depend on $y$, and, given $y$ and $\lambda$, $b$ does not depend on $x$. Finally we require that the questions be independent of the side information. Given these assumptions we can, similar as with non-local games, certify quantumness by calculating the Bell parameter and comparing it to the local threshold. If the Bell parameter exceeds the local threshold we know that the statistics must be non-local. Statistics that are not non-local are called local. The set of local statistics is called the local polytope, $\mathcal{L}$. We can think of the facets of the local polytope as the Bell inequalities. If one Bell inequality is violated by $P_{AB|XY}$ the statistics lie outside of $\mathcal{L}$ and are thus non-local. The local polytope with its facets is schematically depicted in Figure 3.
2.5.1 The CHSH game

As an example of a non-local game we consider the CHSH game. The winning function for the game is

$$w : X \times Y \times A \times B \rightarrow \{0, 1\}$$

$$(x, y, a, b) \mapsto \begin{cases} 1 & \text{if } a \oplus b = x \land y, \\ 0 & \text{otherwise} \end{cases}$$

meaning the game is won, if and only if the questions and answers satisfy $a \oplus b = x \land y$. It can be shown that, if the questions are uniformly distributed, no classical strategy can achieve a winning probability higher than $p_{\text{win}} = \frac{3}{4}$. However, if Alice and Bob share a maximally entangled state and do measurements according to the questions and use the outputs as answers, they can achieve a winning probability $p_{\text{win}} = \frac{2 + \sqrt{2}}{4}$.

The CHSH game is the game corresponding to the CHSH inequality [CHSH69],

$$\beta \equiv \sum_{a,b,x,y \in \{0,1\}} (-1)^{a+b+xy} P_{AB|XY}(ab|xy) \leq 2.$$  \hspace{1cm} (6)

An equivalent version of the CHSH inequality (while enforcing non-signalling condition) is

$$\alpha \equiv P_{AB|XY}(00|00) - (P_{AB|XY}(01|01) + P_{AB|XY}(10|10) + P_{AB|XY}(00|11)) \leq 0.$$  \hspace{1cm} (7)

This inequality was first introduced by Eberhard [Ebe93]. Within quantum mechanics we can have non-local values $\beta \in [2, 2\sqrt{2}]$ and $\alpha \in [0, \frac{\sqrt{2}-1}{2}]$. Thus, given the affine relation between the two values we find the relation

$$\beta = 4\alpha + 2 \iff \alpha = \frac{\beta}{4} - \frac{1}{2}.$$  \hspace{1cm} (8)

2.6 Measurement dependent locality

In standard non-local games we usually assume that the questions are uniformly distributed and cannot be influenced by anyone. This assumption is called measurement independence. Pütz et. al [PRB+14] weakened the assumption of measurement independence to an assumption of limited measurement dependence, where Eve can influence the distribution of the questions to some extent, and studied Bell inequalities in this...
scenario. A schematic drawing of the setting in this scenario is shown in Figure 2. The inputs can now depend on some hidden information and need not be uniform anymore. The way Eve can influence the distribution of the questions is described by an MDL source (Definition 14). The main result of their work is that we can verify the usage of quantum strategies for any amount of measurement dependence, as long as \( \mu_{\text{min}} > 0 \).

In order to verify quantum strategies with an MDL source, we need a new Bell inequality, an MDL inequality \([PRB+14]\)

\[
S_\mu \equiv \mu_{\text{min}} P_{ABXY}(0000) - \mu_{\text{max}} (P_{ABXY}(0101) + P_{ABXY}(1010) + P_{ABXY}(0011)) \leq 0. \tag{9}
\]

Using this inequality we verify quantum strategies if \( S_\mu \geq 0 \). Furthermore we now call statistics that do not violate Equation (9) measurement dependent local (MDL). This MDL inequality translates into a game with winning function

\[
w(a, b, x, y) = \begin{cases} 
\mu_{\text{min}} & \text{if } (a, b, x, y) = (0, 0, 0, 0) \\
-\mu_{\text{max}} & \text{if } (a, b, x, y) \in \{(0, 1, 0, 1), (1, 0, 1, 0), (0, 0, 1, 1)\} \\
0 & \text{otherwise.}
\end{cases}
\]

### 2.7 Untrusted device

In our randomness amplification protocol we use two separated untrusted devices to play a non-local game. Untrusted in this context means that we assume that the adversary produces the devices and can produce them (almost) anyway she wants. However, we enforce the condition that we can use the device to play a two-player non-local game with binary inputs and outputs; i.e., upon receipt of a binary input, the device output does not depend on the inputs and the outputs can depend on the outcome of the measurement. If the devices do not produce outputs or produce outputs that are not binary we can simply abort the protocol.

Moreover, we assume that quantum mechanics is complete. Thus we can model the inner workings of the device as doing measurements on an unknown quantum state. The measurements can depend on the inputs and the outputs can depend on the outcome of the measurement. If the devices are used sequentially in a number of rounds like in our protocol, the measurements can be different in each round. In addition the new quantum state on which the measurements are done can depend on previous rounds.

In a device-independent adversarial scenario we play the non-local game to verify the quantumness of the inner workings of the devices. Hence we can think of the adversary implementing a strategy, i.e., a specific set of states and measurements, in the device such that she gains a maximal amount of knowledge of the outputs. This strategy also includes her attempt to trick us into thinking that the devices produce non-local statistics whereas they are not. Since the adversary is also assumed to be the manufacturer of the devices she can build a third device that contains a purification of the quantum states in the two other devices.

### 2.8 Quantum-proof randomness extractors in the Markov model

A (classical) two-source extractor is defined as follows.

**Definition 17** (Two-source extractor, [Raz05]). A function Ext : \( \{0,1\}^{n_1} \times \{0,1\}^{n_2} \rightarrow \{0,1\}^m \) is called a \((k_1, k_2, \varepsilon)\) two-source extractor if for any two independent sources \( X_1, X_2 \) with \( H_{\text{min}}(X_1) \geq k_1 \) and \( H_{\text{min}}(X_2) \geq k_2 \), we have

\[
\frac{1}{2} \left\| \rho_{\text{Ext}(X_1, X_2)} - \rho_U^m \right\| \leq \varepsilon,
\]

where \( \rho_U^m \) is the fully mixed state on a system of dimension \( 2^m \). Ext is said to be strong in the \( i \)'th input if

\[
\frac{1}{2} \left\| \rho_{\text{Ext}(X_1, X_2), X_i} - \rho_U^m \otimes \rho_{X_i} \right\| \leq \varepsilon.
\]

If Ext is not strong in any of its inputs it is said to be weak.
In our work we use extractors that work in the presence of quantum side information described by the Markov model introduced in [AFPS16]. In the Markov model we assume that the two sources of a two-source extractor together with the side information form a Markov chain: \( I(X_1 : X_2 | C) = 0 \) (where \( X_1 \) and \( X_2 \) are classical registers, while \( C \) can hold a quantum state). This can be interpreted as the requirement that, given the side information, the two sources are independent. Formally the quantum Markov model and a quantum-proof two-source extractor in the Markov model are defined as follows.

**Definition 18** (Quantum Markov model, [AFPS16]). A ccq-state \( \rho_{X_1,X_2,C} \) is said to belong to the Markov model if \( X_1 \leftrightarrow C \leftrightarrow X_2 \) is a Markov chain (i.e., \( I(X_1 : X_2 | C) = 0 \)).

**Definition 19** (Strong quantum-proof two-source extractor in the Markov model, [AFPS16]). A function \( \text{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m \) is a \((k_1, k_2, \varepsilon)\) quantum-proof two-source extractor in the Markov model, strong in the second source, if for all sources \( X_1, X_2 \), and quantum side information \( C \), where \( X_1 \leftrightarrow C \leftrightarrow X_2 \) form a Markov chain, and with min-entropy \( H_{\min}(X_1|C) \geq k_1 \) and \( H_{\min}(X_2|C) \geq k_2 \), we have

\[
\frac{1}{2} \| \rho_{\text{Ext}(X_1,X_2),X_2} - \rho_{U_m} \otimes \rho_{X_2|C} \| \leq \varepsilon.
\]

where \( \rho_{\text{Ext}(X_1,X_2)C} = \text{Ext} \otimes I_C \rho_{X_1,X_2,C} \) and \( \rho_{U_m} \) is the fully mixed state on a system of dimension \( 2^m \).

The main result of [AFPS16] is that any (classical) two-source extractor is also quantum-proof in the Markov model:

**Lemma 20.** Any \((k_1, k_2, \varepsilon)\)-[strong] two-source extractor is a \( (k_1 + \log \frac{1}{\varepsilon}, k_2 + \log \frac{1}{\varepsilon}, \sqrt{3\varepsilon} \cdot 2^{(m-2)})\)-[strong] quantum-proof two-source extractor in the Markov model, where \( m \) is the output length of the extractor.

In this work we use such an extractor, but for a source with a lower bound on the smooth min-entropy rather than the min-entropy itself. The effect of this on the parameters of the extractor was also investigated in [AFPS16]. We use the following form of the statement:

**Lemma 21.** Let \( \text{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m \) be a \((k_1, k_2, \varepsilon)\) quantum-proof two-source extractor in the Markov model, strong in the source \( X_1 \). Then for any Markov state \( \rho_{X_1,X_2,C} \) with \( H_{\min}(X_1|C) \rho \geq k_1 + \log(1/\varepsilon) + 1 \) and \( H_{\min}(X_2|C) \rho \geq k_2 + \log(1/\varepsilon) + 1 \),

\[
\frac{1}{2} \| \rho_{\text{Ext}(X_1,X_2),X_2} - \rho_{U_m} \otimes \rho_{X_2|C} \| \leq 6(\varepsilon_s + \varepsilon).
\]

### 2.9 The entropy accumulation theorem

The entropy accumulation theorem (EAT), introduced in [DFR16], gives a straightforward way of bounding the smooth min-entropy of a system consisting of \( n \) random variables that possibly depend on each other. For our work the simplified versions of the definitions and theorems of [DFR16], as presented in [AFRV16], suffice. In the following we introduce the definitions and theorems which are crucial to working with the EAT.

**Definition 22** (EAT channels). EAT channels \( N_i : R_{i-1} \to R_i A_i B_i I_i C_i \), for \( i \in [n] \), are CPTP maps such that for all \( i \in [n] \):

1. \( A_i, B_i, I_i \) and \( C_i \) are finite-dimensional classical systems (RV). \( A_i \) and \( B_i \) are of dimension \( d_{A_i} \) and \( d_{B_i} \) respectively. \( R_i \) is an arbitrary quantum register.

2. For any input state \( \sigma_{R_{i-1}R'} \), where \( R' \) is a register isomorphic to \( R_{i-1} \), the output state \( \sigma_{R_i A_i B_i I_i C_i R'} = (N_i \otimes I_{R'}) (\sigma_{R_{i-1}R'}) \) has the property that the classical value \( C_i \) can be measured from the marginal \( \sigma_{A_i B_i I_i} \) without changing the state.

3. For any initial state \( \rho_{R_0 E}^0 \), the final state \( \rho_{A^n B^n I^n C^n E} = (\text{Tr}_{R_n} \circ N_n \circ \cdots \circ N_1) \otimes I_E \rho_{R_0 E}^0 \) satisfies the Markov chain condition \( A_i^{-1} B_i^{-1} \leftrightarrow I_i^{-1} E \leftrightarrow I_i \) for each \( i \in [n] \).
Definition 23 (Min-tradeoff function). Let $\mathcal{N}_1, \ldots, \mathcal{N}_N$ be a family of EAT channels. Let $\mathcal{C}$ denote the common alphabet of $C_1, \ldots, C_n$. A function $f_{\min}$ from $\mathcal{F}_\mathcal{C}$ to the real numbers is called a min-tradeoff function for $\{\mathcal{N}_i\}$ if it satisfies

$$f_{\min}(p) \leq \inf_{\sigma \in R_{n-1}^+ \mathcal{N}_i(\sigma)_{C_i} = p} H(A_i|B_i|R_i')_{\mathcal{N}_i(\sigma)}$$

for all $i \in [n]$, where the infimum is taken over all input states of $\mathcal{N}_i$ for which the marginal on $C_i$ of the output state is the probability distribution $p$, and the infimum over the empty set is defined as plus infinity.

Definition 24. Let $C^n$ be a set of random variables over the alphabet $\mathcal{C}$. Then $\text{freq}_{C^n}$ defines the probability distribution over $\mathcal{C}$ defined by $\text{freq}_{C^n}(x) = \frac{1}{n} \left| \{ i \in \{1, \ldots, n\} : C_i = x \} \right|$.

Definition 25 (Infinity norm). Let $f : \Omega \rightarrow \mathbb{R}$ be a function over some set $\Omega \subset \mathbb{R}^m$. Then the infinity norm of the gradient of $f$ is defined as

$$\|\nabla f\|_\infty = \sup \left\{ \frac{\partial}{\partial x_i} f(x) : x \in \Omega, i \in \{1, \ldots, m\} \right\}.$$ 

Theorem 26 (EAT, [DFR16]). Let $\mathcal{N}_i : R_{i-1} \rightarrow R_i A_i B_i I_i C_i$ for $i \in [n]$ be EAT channels as in Definition 22, $\rho^{A^n B^n I^n C^n} = (\text{Tr}_{R_n} \circ N_n \circ \cdots \circ N_1) \otimes I_E \rho_{R_0 E}$ be the final state, $\Omega$ an event defined over $C^n$, $\rho_{\Omega}$ the probability of $\Omega$ in $\rho$, and $\rho_{\Omega \epsilon}$ the final state conditioned on $\Omega$. Let $\epsilon \in (0, 1)$.

For $f_{\min}$ a min-tradeoff function for $\{\mathcal{N}_i\}$, as in Definition 23, and any $t \in \mathbb{R}$ such that $f_{\min}(\text{freq}_{C^n}) \geq t$ for any $C^n \in C^n$ for which $\text{Pr}(C^n|\rho_{\Omega \epsilon}) > 0$,

$$H_{\min}^{\epsilon \min} (A^n B^n | I^n E)_{\rho_{\Omega \epsilon}} > nt - v \sqrt{\eta},$$

where $v = 2 \log(1 + 2d_{A_i B_i}) + \left\| \nabla f_{\min} \right\|_\infty \sqrt{1 - 2 \log(\epsilon \cdot \rho_{\Omega})}$ and $d_{A_i B_i}$ denotes the dimension of $A_i B_i$.

To gain some intuition regarding the EAT we now give a short explanation of how it is used below. The concrete and formal details are given in the following sections. Our EAT-channels are chosen to be the channels describing the actions in each step of the protocol (both of the honest parties and the uncharacterised quantum device). The event $\Omega$ is the event of not aborting the protocol. $\rho_{\Omega \epsilon}$ is hence the state in the end of the protocol conditioned on not aborting. The goal is then to lower-bound the conditional smooth min-entropy of this state and this is exactly what Theorem 26 gives us. The first order term in the given bound is $nt$ where $n$ is the number of rounds of the protocol and $t$ is the minimal amount of entropy accumulated in each step, quantified using the min-tradeoff function. In Section 3 we make the relevant analysis to find the value of $t$.

3 Secret randomness from a single round

In this section we quantify the randomness of the outputs of an MDL experiment. With that achieved we can carry on in Section 4 to quantify the randomness of the outputs in a sequence of MDL experiments. Hence quantifying the randomness in a single MDL experiment is crucial in our process of producing an arbitrary amount of randomness.

In our single MDL experiment we consider a device consisting of two separated components, such that one can enforce a situation in which the “non-signalling conditions” between the components hold. (i.e., the two components cannot signal, nor communicate, with one another). During the execution of the experiment the two operators of the device, Alice and Bob, draw inputs, $X$ and $Y$, from the MDL source. They then feed the inputs to their component and record the output, $A$ and $B$, that it generates. As noted in Section 2.7, we consider a third party that can hold a purification, $E$, of the quantum state in the device. We thus want to quantify the randomness of $A$ and $B$ given $X$, $Y$, and $E$. An algorithmic description of the MDL experiment is given in Protocol 1. In Step 3 we use a uniform and independent random bit $F$ to symmetrise
Protocol 1 Execution of an MDL experiment

Arguments:

\( M(\mu) \) – \( \mu \)-MDL source

\( D \) – untrusted device of two components

1. Alice and Bob choose inputs from the MDL source with parameters \( \mu \).
2. Alice and Bob use \( D \) with \( X,Y \) and record their outputs as \( A,B \).
3. [optional] Alice and Bob choose a uniform and independent binary random variable \( F \) and update their outputs as \( \tilde{A} = A \oplus F \) and \( \tilde{B} = B \oplus F \).

the outputs. Of course, in the context of randomness amplification we cannot do this. Nevertheless, we use this just as a step in the proof and later argue that the symmetrisation step can be dropped in practice.

Formally we choose to quantify the randomness by the von Neumann entropy, \( H(AB|XYE) \). The remaining part of this section is dedicated to proving the following bound on this entropy.

**Lemma 27.** Consider the MDL experiment described in Protocol 1 where both the inputs and the outputs are binary. Then, for a state and a set of measurements (i.e., strategy of the adversary) yielding a violation \( S_\mu > 0 \) of Inequality (9), the bound

\[
H(AB|XYE) \geq 1 - h \left( \frac{1}{2} + \frac{1}{\mu^*} \sqrt{S_\mu(S_\mu + \mu^*)} \right)
\]

(10)

on the von Neumann entropy of the outputs holds, where \( \mu^* = \mu_{\text{min}} \cdot \mu_{\text{max}} \).

We prove the lemma by employing the bound on the Holevo quantity (introduced later) that Pironio et al derived in [PAB+09] for the CHSH game. We adapt the bound to the MDL game with biased inputs.

To prove Lemma 27 we first express the entropy in terms of the Holevo quantity, \( \chi(\tilde{A}:FE|X = x) \), similarly to what was done as in [AFRV16]. The expression is given in the following lemma. The proof is given in Appendix B.

**Lemma 28.** In an MDL experiment with binary inputs and outputs, as described in Protocol 1, with two devices, between which the non-signaling condition holds, the entropy of the outputs can be lower bounded as

\[
H(\tilde{A}\tilde{B}|XYFE) \geq \sum_x \Pr[X = x] \cdot \left( 1 - \chi(\tilde{A}:FE|X = x) \right)
\]

(11)

where \( \chi(\tilde{A}:FE|X = x) = H(FE|X = x) - H(FE|\tilde{A},X = x) \) is the Holevo quantity.

We now proceed to prove Lemma 27.

**Proof of Lemma 27.** In the proof of our claim we first prove an upper bound on the Holevo quantity of the symmetrized outputs as a function of the MDL violation. Once the upper bound on the Holevo quantity is derived we make use of Lemma 28 and derive the lower bound on the von Neumann entropy of the symmetrized outputs. Finally we argue why the entropy bound for the symmetrized outputs is also an entropy bound for the unsymmetrized outputs.

In order to upper bound the Holevo quantity we start with the bound that was derived in [PAB+09, Equation (11)] for the standard CHSH scenario. Together with the relation in Equation (8) we find

\[
\chi(\tilde{A}:EF|X = x) \leq h \left( \frac{1}{2} + \frac{1}{2} \sqrt{\frac{\beta^2}{4} - 1} \right) = h \left( \frac{1}{2} + \sqrt{\alpha(\alpha + 1)} \right)
\]

(12)

where \( \beta \) is the violation of the CHSH inequality and \( \alpha \) is the violation of the Eberhard inequality.

Continuing we relate this bound to our scenario where the inputs for the Bell measurements are not uniform and not independent. To that end we consider two processes producing different states. In the
first process we consider an MDL source that produces biased bits that might be correlated with some side information \( \lambda \). In the second process we consider an input source that produces uniform and independent bits. We want to relate the Holevo quantity of the outputs in both processes.

In both processes we can describe the generation of the outputs as doing a measurement on an unknown quantum state

\[
\rho_{QAQB E, \lambda},
\]

where \( Q_A \) and \( Q_B \) are the quantum registers in Alice’s and Bob’s device respectively and \( E \) is the quantum side information that the adversary holds. The state can also depend on the classical side information \( \lambda \). We also include the uniform and independent random variable \( F \) in the measurements. This variable then determines whether the outputs are being flipped or not, as described in Step 3. More precisely we describe the measurement, implemented with a strategy \( \lambda \) for specific \( x, y, f \), as a CPTP map evolving the unknown quantum state

\[
\mathcal{E}_{xf, \lambda} : QAQB E \to QAQB \tilde{A} \tilde{B} E
\]

\[
\rho_{QAQB E, \lambda} \mapsto \mathcal{E}_{xf, \lambda}(\rho_{QAQB E, \lambda}).
\]

Note that the two parts of the device and the adversary are spatially separated and thus the CPTP map factors into three parts,

\[
\mathcal{E}_{xf, \lambda} = \mathcal{E}^A_{xf, \lambda} \otimes \mathcal{E}^B_{yf, \lambda} \otimes \mathcal{I}_E
\]

\[
\mathcal{E}^A_{xf, \lambda} : Q_A \to QA \tilde{A}
\]

\[
\mathcal{E}^B_{yf, \lambda} : Q_B \to QB \tilde{B},
\]

where \( \mathcal{I}_E \) is the identity map on the adversary’s quantum register.

**Process 1** is associated to the measurements in our MDL scenario. First we choose inputs \( X, Y \in \{0, 1\} \) according to a distribution \( P_{XY|\lambda} \) satisfying Definition 14. Furthermore we also choose an independent and uniform random variable \( F \in \{0, 1\} \) for the symmetrisation of the outputs. For a specific strategy \( \lambda \) of the adversary the post measurement state is

\[
\rho^1_{QAQB \tilde{A} \tilde{B} X YE, \lambda} = \sum_{x,y} P_{XY|\lambda}(x, y) \sum_f \frac{1}{2} (\mathcal{E}^A_{xf, \lambda} \otimes \mathcal{E}^B_{yf, \lambda} \otimes \mathcal{I}_E)(\rho_{QAQB E, \lambda}) \otimes |xyf \rangle \langle xyf|,
\]

where \( \{|x\rangle\}_x \), \( \{|y\rangle\}_y \), and \( \{|f\rangle\}_f \) each form an orthonormal basis of a two dimensional Hilbert space. After tracing out the systems \( Q_A, Q_B, B \) and \( Y \), which are irrelevant for the calculation of \( \chi \), we are left with

\[
\rho^1_{AXYE, \lambda} = \sum_x P_{X|\lambda}(x) \left( \sum_f \frac{1}{2} (\text{Tr}_{Q_A} \circ \mathcal{E}^A_{xf, \lambda} \otimes \mathcal{I}_E)(\rho_{QA E, \lambda}) \otimes |f \rangle \langle f| \right) \otimes |x \rangle \langle x|.
\]

We denote

\[
\rho^1_{AYE|X=x, \lambda} = \left( \sum_f \frac{1}{2} (\text{Tr}_{Q_A} \circ \mathcal{E}^A_{xf, \lambda} \otimes \mathcal{I}_E)(\rho_{QA E, \lambda}) \otimes |f \rangle \langle f| \right).
\]

**Process 2** is associated to the standard CHSH scenario. First we choose the inputs \( X, Y \in \{0, 1\} \) independent of everything else and uniformly at random. Then we choose an independent and uniform random variable \( F \in \{0, 1\} \) to symmetrise the outputs. Similar to Process 1, for a specific strategy \( \lambda \) of the adversary, the post measurement state is

\[
\rho^2_{QAQB \tilde{A} \tilde{B} X YE, \lambda} = \sum_{x,y} \frac{1}{4} \sum_f \frac{1}{2} (\mathcal{E}^A_{xf, \lambda} \otimes \mathcal{E}^B_{yf, \lambda} \otimes \mathcal{I}_E)(\rho_{QAQB E}) \otimes |xyf \rangle \langle xyf|.
\]
After tracing out the systems $Q_A, Q_B, B$ and $Y$ we are left with
\[
\rho^2_{\tilde{A}XFE,\lambda} = \sum_x \frac{1}{2} \left( \sum_f \frac{1}{2} \left( \text{Tr} Q_A \circ E^A_{x,\lambda} \otimes \mathcal{I}_E \right) \left( \rho_{Q_A} \otimes |f\rangle \langle f| \right) \right) \otimes |x\rangle \langle x|.
\]
We denote
\[
\rho^2_{\tilde{A}FE|X=x,\lambda} = \left( \sum_f \frac{1}{2} \left( \text{Tr} Q_A \circ E^A_{x,\lambda} \otimes \mathcal{I}_E \right) \left( \rho_{Q_A} \otimes |f\rangle \langle f| \right) \right).
\]
We observe that the states $\rho^1_{\tilde{A}FE|X=x,\lambda}$ and $\rho^2_{\tilde{A}FE|X=x,\lambda}$ are equal and consequently we find
\[
\chi(\tilde{A} : EF|X = x) \rho^1_{\tilde{A}FE|X=x,\lambda} = \chi(\tilde{A} : EF|X = x) \rho^2_{\tilde{A}FE|X=x,\lambda}.
\]
This concludes our prove that the Holevo quantity is the same in Process 1, with biased inputs, and Process 2, with uniform inputs.

In the next step we want to express the bound on the Holevo quantity as a function of the violation $S_\mu$ of our MDL inequality. Starting with Equation (12) we know that
\[
\chi(\tilde{A} : EF|X = x) \rho^1_{\tilde{A}FE|X=x,\lambda} = \chi(\tilde{A} : EF|X = x) \rho^2_{\tilde{A}FE|X=x,\lambda} \leq h \left( \frac{1}{2} + \sqrt{\alpha} \right).
\]
Now we can relate $S_\mu$ to a minimal Bell violation $\alpha$ that would have been observed with the given state and measurements. For the relation between the two violations we find
\[
S_\mu = \mu_{\text{min}} \frac{P_{\tilde{A}\tilde{B}|XY}(00) \cdot P_{XY}(00)}{\geq \mu_{\text{max}}} - \mu_{\text{max}} \left( \frac{P_{\tilde{A}\tilde{B}|XY}(01) \cdot P_{XY}(01) + P_{\tilde{A}\tilde{B}|XY}(10) \cdot P_{XY}(10) + P_{\tilde{A}\tilde{B}|XY}(00) \cdot P_{XY}(11)}{\geq \mu_{\text{min}}} \right)
\leq \mu^* \left( P_{\tilde{A}\tilde{B}|XY}(00) - (P_{\tilde{A}\tilde{B}|XY}(01) + P_{\tilde{A}\tilde{B}|XY}(10) + P_{\tilde{A}\tilde{B}|XY}(00)11)) \right)
\leq \mu^* \cdot \alpha
\]
and hence
\[
\alpha \geq \frac{1}{\mu^*} S_\mu. \tag{15}
\]
We find the final bound on the Holevo quantity by plugging this relation into Equation 12,
\[
\chi(\tilde{A} : EF|X = x) \leq h \left( \frac{1}{2} + \sqrt{\alpha} \right) \leq h \left( \frac{1}{2} + \frac{1}{\mu^*} \sqrt{S_\mu(S_\mu + \mu^*)} \right),
\]
where the last inequality holds because $h(x)$ is monotonically decreasing for $x \in [\frac{1}{2}, 1]$. A bound on the entropy can be found by employing Lemma 28,
\[
H(\tilde{A}\tilde{B}|XYFE) \geq 1 - h \left( \frac{1}{2} + \frac{1}{\mu^*} \sqrt{S_\mu(S_\mu + \mu^*)} \right).
\]
We conclude the proof by showing that the bound on the entropy of the symmetrized outputs is the same as the bound on the entropy of the unsymmetrized outputs. Namely we have

\[
H(\tilde{A}\tilde{B}|XYFE) \leq H(AB|XYFE) = H(AB|XYE),
\]

where the first step follows because, for fixed \(F\), the symmetrisation step is a deterministic operation, and the second step follows because \(F\) is independent of everything else.

A plot of the bound from Lemma 27 is shown in Figure 4. Once it shows the entropy bound as a function of the MDL violation \(S_\mu\) for different \(\mu\), and once a lower bound on the maximal achievable entropy as a function of \(\mu_{\min}\). The reason that we can only plot a lower bound on the maximal entropy is due to the dependence of the maximal entropy on the specific source. The exact reasoning and how we obtained the lower bound is explained in Appendix A.

In the plots we clearly see that the entropy of the outputs increases with increasing MDL violation. Furthermore we can observe that the maximal achievable entropy bound decreases with increasing source bias. Intuitively this makes sense since we expect to get a lower amount of randomness in the outputs if we start with less random inputs.

We also see in the above plot that the entropy is non-zero once there is a violation of the relevant inequality. As will be clear from the next Section, this implies that, asymptotically, we can tolerate maximal amount of noise — as long as there is a violation of the MDL inequality some randomness can be extracted.

4 Randomness amplification protocol

In the following sections we first introduce the setting of our randomness amplification protocol and explicitly state the assumptions that we are taking. After that we introduce the protocol and proceed to prove the completeness and the soundness of the protocol.

4.1 Setting and assumptions

We consider a setting where we have an MDL source and an untrusted device with at least two components. All the components in our setting are spatially separated and possibly manufactured by an adversary. Since the source and the components of the device are separated the non-signalling condition holds pairwise between them. Both the device and the source can be correlated with some classical side information \(\lambda\) that the adversary holds. Furthermore the adversary can have access to a quantum state \(\rho_E\) that can be correlated with the device and the source.

During the protocol the source produces the inputs \(X^nY^n\) for the device which then, upon receiving the inputs, produces the outputs \(A^nB^n\). After the device produced the outputs the source produces another string of binary random variables \(Z^d\). The extractor then produces the final output \(K^m\) using \(A^nB^n\) and \(Z^d\) as inputs. The whole setting is schematically depicted in Figure 5.

We summarise the general assumptions of the analysis of our protocol in the following list.

1. Quantum mechanics is correct.
2. The adversary is limited by quantum mechanics and without loss of generality we can assume that the adversary only holds a purification of Alice and Bob’s initial quantum state.
3. The untrusted device has at least two separated components.
4. The adversary only has classical side information, \(\lambda\), about the source of randomness.
5. The source of randomness is a public \(\mu\)-MDL source.
Figure 4: In Figure 4a we plot the entropy bound given in Equation (10) as a function of the MDL violation $S_\mu$ for several choices of $\mu$. The black dots indicate a lower bound on the maximal possible $S_\mu$ within quantum mechanics. In Figure 4b we show a lower bound on the maximal achievable entropy as a function of $\mu_{\min}$ (recall that $\mu_{\min} = 1/4$ corresponds to a uniform source). The exact way we lower bound the maximal achievable entropy is explained in Appendix A. The plot of the maximal achievable entropy shows one curve where we fixed $\mu_{\min}$ and chose $\mu_{\max} = 1 - 3 \cdot \mu_{\min}$, and one plot where we chose $\mu_{\max} = (1 - \mu_{\min})/3$. These two choices correspond to the two extreme cases for fixed $\mu_{\min}$. A source with $\mu_{\max} = 1 - 3 \cdot \mu_{\min}$ is, for our purpose, the worst case since it produces outputs such that one output pair is considerably favoured while all other pairs appear with low probability $\mu_{\min}$. A source where $\mu_{\max} = (1 - \mu_{\min})/3$ is, for our purpose, the best case since only one output pair appears with low probability while all the other pairs appear with equally (high) probability.
Assumptions 4 and 5 imply that the guessing probability (Definition 9) for the outputs of the source is bounded as follows $\mu_{\text{min}} \leq p_{\text{guess}}(X_iY_i|X_{i-1}Y_{i-1}E,\lambda) \leq \mu_{\text{max}} \forall \lambda, i$.

6. While the device produces outputs, it holds that

$$I(A^{i-1}B^{i-1}: X_iY_i|X_{i-1}Y_{i-1}E,\lambda) = 0$$

and, after the device is done, it holds that

$$I(Z^d: A^nB^n|X^nY^nE,\lambda) = 0.$$ 

The first two assumptions amount to the assumption that quantum mechanics is correct and complete. Since no experimental evidence has been found that this is not the case, these are reasonable. Furthermore, the fact that the device consists of at least two components can easily be verified by inspecting it before executing the protocol and the non-signalling condition can be reliably enforced by shielding the two parts of the device. Without having any restrictions on the source we could not do anything. One therefore must use some assumptions about the source. Here we use Assumptions 4 and 5 which are necessary for our proof technique. Assumption 6 can be understood as assuming that, given the adversary’s side information, the device and the source are independent, which again can be seen as the restriction that the adversary does not have access to the device and the source after they were produced. This condition can easily be enforced by securing the devices from being tampered with.

### 4.2 Security definition

We define the security via the secrecy of its outputs, similar as was done for the security definition of the DIQKD protocol in [AFRV16].

**Definition 29** (Secrecy). A randomness amplification protocol is said to be $\varepsilon_{\text{RA}}$-secret, when implemented using a device $D$, if for an output of length $m$,

$$\left(1 - \Pr[\text{abort}]\right) \| \rho_{K^m} - \rho_{U^m} \otimes \rho_{\Sigma} \| \leq \varepsilon_{\text{RA}},$$
Protocol 2 Randomness Amplification Protocol

Arguments:
- \( M(\mu) \) – \( \mu \)-MDL source
- \( D \) – untrusted device of (at least) two components
- \( n \in \mathbb{N}_+ \) – number of rounds
- \( S_{\text{exp}} \) – lower bound on the expected violation of the MDL inequality for an honest (possibly noisy) implementation
- \( \delta_{\text{est}} \in (0, S_{\text{exp}}) \) – width of statistical confidence interval for the estimation test
- Ext : \( \{0,1\}^{2n} \times \{0,1\}^d \rightarrow \{0,1\}^m \) – (\( k_1, k_2, \varepsilon_{\text{ext}} \)) quantum-proof randomness extractor in the Markov model which is strong in the second input.

Entropy Accumulation:
1: For every round \( i \in \{1, \ldots, n\} \) do steps 2-4.
2: Alice and Bob choose inputs \( X_i \) and \( Y_i \) from the MDL source.
3: Alice and Bob use \( D \) with \( X_i, Y_i \) and record their outputs as \( A_i, B_i \).
4: Alice and Bob set \( C_i = w(A_i, B_i, X_i, Y_i) \) for \( w \) as defined in Equation (16).
5: Alice and Bob abort the protocol if \( \bar{C} \equiv \frac{1}{n} \sum_j C_j < (S_{\text{exp}} - \delta_{\text{est}}) \).

Randomness Extraction:
6: Draw a bit string \( Z^d \) from \( M(\mu) \).
7: Use Ext to create \( K^m = \text{Ext}(A^n B^n, Z^d) \).

where \( K^m \) is the output of the RAP, \( \Sigma \) is the adversary’s side information that can be correlated with \( D \), and \( U^m \) is a uniform random variable of \( m \) bits.

The protocol is thus said to be secure if either the protocol aborts with high probability or the outputs are close to uniform.

4.3 The protocol

The protocol is given in Protocol 2.

Our proposed RAP consists of two parts. In the first part we accumulate entropy. For that matter we perform a series of MDL experiments, similar to the one described in Section 3. In these MDL experiments we draw inputs from an MDL source and feed them to a device that produces outputs. Ideally these outputs will be generated by doing measurements on a quantum state such that an MDL inequality is violated. In the second part we draw another string from the MDL source and use this string, as well as the output from the entropy accumulation part, as inputs for the extractor. The extractor then produces the final output.

During the entropy accumulation part of the protocol the variable \( C_i \) is set in each round. This variable is set to help evaluate whether the protocol should abort or not. In each round \( C_i \) is set according to the winning function

\[
 w(A_i, B_i, X_i, Y_i) = \begin{cases} 
 \mu_{\text{min}} & \text{if } (A_i, B_i, X_i, Y_i) = (0, 0, 0, 0) \\
 -\mu_{\text{max}} & \text{if } (A_i, B_i, X_i, Y_i) \in \{(0, 1, 0, 1), (1, 0, 1, 0), (0, 0, 1, 1)\} \\
 0 & \text{otherwise}.
\end{cases}
\]

After the \( n \) rounds of the entropy accumulation routine we decide whether to abort or not by comparing \( \bar{C} \equiv \frac{1}{n} \sum_j C_j \) with \( (S_{\text{exp}} - \delta_{\text{est}}) \). Note that we almost always abort for sources where \( S^*_\mu \leq \delta_{\text{est}} \) (\( S^*_\mu \) is the maximal MDL value in quantum mechanics, see Appendix A). Thus we cannot amplify randomness for sources for which \( S^*_\mu \leq \delta_{\text{est}} \). However, we need a positive \( \delta_{\text{est}} \) in order to get a low probability for aborting in an honest implementation (see Section 4.4). To remedy this problem we can decrease \( \delta_{\text{est}} \) at the cost of increasing \( n \). Hence, it is possible to have a reasonable probability of aborting in an honest implementation and still be able to amplify arbitrary SV sources.
We state the following theorem that quantifies the quality of the protocol’s output. It is a formal version of Theorem 3 which was given in the introduction. The proof of the theorem is given in the end of the section as it combines our separate proofs of soundness and completeness.

**Theorem 30.** Given any public SV-source with bias \( \mu \in (0, 0.5) \) and any two-component device \( D \) that fulfills the assumptions described in Section 4.1, let \( n \) be the number of rounds in Protocol 2, \( \varepsilon_s, \varepsilon_{EA} \in (0, 1) \), \( S_{exp} \leq S_{\mu}^* \), \( \delta_{est} \in (0, S_{exp}) \) and \( m, \varepsilon_{ext} \) the parameters of the \((k_1, k_2, \varepsilon_{ext})\)-extractor used in Protocol 2, with \( k_1, k_2 \) fulfilling Equation (34). Then:

1. (Secrecy) Protocol 2 produces a string \( K^m \) of length \( m \) such that:
   \[
   (1 - \Pr[\text{abort}]) \| \rho_{K^n \Sigma} - \rho_{U^n} \otimes \rho_{\Sigma} \| \leq 12 (\varepsilon_s + \varepsilon_{ext}) + \varepsilon_{EA} ,
   \]
   where \( \Sigma = EX^nY^nZ^d \lambda \) is the adversary’s side information.

2. (Completeness) There exists an honest implementation of the device such that Protocol 2 aborts with probability \( \varepsilon_c \leq \exp\left(-\frac{2n\delta_{est}^2}{(\mu_{\min} + \mu_{\max})^2}\right) \) when using this device.

### 4.4 Completeness

In order for our RAP to be useful we do not only need a protocol that, in theory, produces uniform outputs but also one that can be implemented. We call this criterion the completeness of the protocol.

**Lemma 31** (Completeness). Let \( M(\mu) \) be any \( \mu \)-MDL source and let \( S_{\exp} \leq S_{\mu}^* \), where \( S_{\mu}^* \) is the maximal possible value for \( S_{\mu} \) in quantum theory for the given MDL source. Then Protocol 2 is complete with completeness parameter \( \varepsilon_c \leq \exp\left(-\frac{2n\delta_{est}^2}{(\mu_{\min} + \mu_{\max})^2}\right) \); i.e., the probability to abort in an honest implementation is upper bounded by \( \varepsilon_c \).

**Proof.** We want to show that there exists a device such that the protocol aborts with probability less than \( \varepsilon_c \). If we implement our device to perform \( n \) independent, identical measurements on the product state \( \rho_{Q_AQ_B} \), where \( \rho_{Q_AQ_B} \) together with the chosen measurements achieves an MDL value \( S_{\exp} \) of the MDL inequality, the expectation value of \( \bar{C} = \frac{1}{n} \sum_j C_j \) is given by \( E[C] = S_{\exp} \). We can then use Hoeffding’s inequality to get an upper bound on the probability that the protocol aborts. We have

\[
\Pr[\text{aborting}] = \Pr[\bar{C} < (S_{\exp} - \delta_{est})] \\
= \Pr[S_{\exp} - \bar{C} > \delta_{est}]
\leq \exp\left(-\frac{2n\delta_{est}^2}{(\mu_{\min} + \mu_{\max})^2}\right).
\]

In Lemma 31 we showed that, as long as \( S_{\exp} \) is less than the maximal quantum value of \( S_{\mu} \), there is an honest implementation of the protocol such that it aborts with low probability. In order for our protocol to be useful in reality it is important to notice that, as shown in [PRB+14], for all \( \mu_{\min} > 0 \) the maximal quantum value of \( S_{\mu} \) is greater than zero. Moreover, in Appendix A we explain how to obtain a state that achieves a violation of the MDL inequality. Thus for all MDL sources with \( \mu_{\min} \neq 0 \) the entropy bound that we derive later on is non-trivial.

### 4.5 Soundness

In the previous part we showed that our proposed protocol is complete. Besides that we also want that the protocol does what it is supposed to do, i.e., if it does not abort the outputs should be secret with high probability. This property, which is sometimes called soundness, is quantified in Definition 29.

In the following we prove that Protocol 2 is secret and determine the secrecy parameter \( \varepsilon_{RA} \). In a first step we derive a lower bound on the smooth min-entropy of the MDL experiments’ outputs. In the second step we show that in our protocol we can make use of the quantum-proof randomness extractors introduced in Section 2.8, and then proceed to determine the exact value of \( \varepsilon_{RA} \).
4.5.1 Lower-bounding the smooth min-entropy

In the first part of the RAP we have the entropy accumulation routine where we perform a number of MDL experiments with our device. The goal now is to lower bound the smooth min-entropy of the outputs of these experiments. To achieve this we employ the EAT (introduced in Section 2.9) together with the entropy bound for a single MDL experiment that was derived in Section 3.

In order to apply the entropy accumulation theorem we need a protocol that evolves the states using EAT channels. In our proposed protocol we have in each round two quantum registers $Q_{A,i}$ and $Q_{B,i}$ holding the quantum state of either of the device’s two parts. Furthermore we have the classical registers $X_i$, $Y_i$ for the inputs, $A_i$, $B_i$ for the outputs of the device, and $C_i$ evaluating the outcome of the MDL experiment. Comparing our registers to Definition 22, we can identify $R_i = Q_{A,i}Q_{B,i}$ and $I_i = X_iY_i$, and denote the channels evolving the states in our protocol as

$$N_i : Q_{A,i-1}Q_{B,i-1} \rightarrow Q_{A,i}Q_{B,i}A_iB_iX_iY_iC_i$$

$$\rho_{Q_{A,i-1}Q_{B,i-1}} \mapsto \rho_{Q_{A,i}Q_{B,i}A_iB_iX_iY_iC_i}. \quad (17)$$

The state after the $n$ rounds of the entropy accumulation part, just before step 5 is denoted by

$$\rho_{A^nB^nX^nY^nC^nE} = (\text{Tr}_{Q_{A,n}Q_{B,n}} \circ N_n \circ \cdots \circ N_1) \otimes I_E \rho_{Q_AQ_BE} \quad (18)$$

In step 5 Alice and Bob decide whether to abort the protocol or not. We denote by $\Omega$ the event of not aborting,

$$\Omega = \{ \bar{C} \geq (S_{\text{exp}} - \delta_{\text{est}}) \}.$$  \hspace{1cm} (19)

Combined we denote by $\rho_{A^nB^nX^nY^nC^nE|\Omega}$, or short $\rho_|\Omega$, the state after the protocol conditioned on not aborting the protocol.

We need to prove that these channels are indeed EAT channels.

**Lemma 32.** The channels $N_i$ that evolve the unknown quantum state of Protocol 2, are EAT channels, i.e., they satisfy Definition 22.

**Proof.** 1. Condition 1. is satisfied because $A_i$, $B_i$, $I_i$ represent the (classical, discrete) inputs and outputs of the device that is employed, and $Q_{A,i}, Q_{B,i}$ are quantum registers.

2. Condition 2. is satisfied because $A_i$, $B_i$, $I_i$ are classical registers and $C_i$ is a classical function of those registers.

3. As is stated in Section 4.1 it holds that $I(A^{i-1}B^{i-1} : I_i|I^{i-1}Em\lambda) = 0$. Thus, the Markov chain condition is satisfied. \hfill $\Box$

Now that we have the necessary preconditions, we can prove a bound on the smooth min-entropy of $A^nB^n$ given the inputs and the side information. More precisely, in Theorem 33, we lower bound $H^{\varepsilon_n}_{\text{min}}(A^nB^n|X^nY^nF^nE)_{\rho_{\Omega}}$ for any $\varepsilon_n \in \{0,1\}$. In our proof we combine [AFRV16, Lemma 9 and Theorem 10] and adapt the proofs to our setting.
The bound is described with the help of the following functions, where \( p \in \mathbb{P}_C \):

\[
S_\mu(p) = \mu_{\min} \cdot p(1) - \mu_{\max} \cdot p(-1),
\]

\[
g_\mu(p) = \begin{cases} 
1 - \left( \frac{1}{2} + \frac{1}{\mu^*} \sqrt{S_\mu(p) (S_\mu(p) + \mu^*)} \right) & \text{for } \frac{S_\mu(p)}{\mu^*} \in \left[ 0, \frac{\sqrt{2}-1}{2} \right), \\
1 & \text{for } \frac{S_\mu(p)}{\mu^*} \in \left[ \frac{\sqrt{2}-1}{2}, 1 \right],
\end{cases}
\]

\[
a(s_t) = \frac{d}{ds_\mu(p)} g_\mu(p) \bigg|_{s_\mu(p) = s_t} \quad \text{and} \quad b(s_t) = g(s_t) - a(s_t) \cdot s_t.
\]

\[
f_{\min}(p, s_t) = \begin{cases} 
g_\mu(p) & \text{for } S_\mu(p) \leq s_t \\
a(s_t) \cdot S_\mu(p) + b(s_t) & \text{for } S_\mu(p) > s_t,
\end{cases}
\]

\[
\zeta_\mu(s_t, \varepsilon_s, \varepsilon_{\text{EA}}) = 2(\log(9) + a(s_t) \cdot \mu_{\max}) \sqrt{1 - 2 \log(\varepsilon_s \varepsilon_{\text{EA}})}.
\]

\[
\eta_{\text{opt}}(\varepsilon_s, \varepsilon_{\text{EA}}, S_{\exp} - \delta_{\text{est}}, n, \mu) = \max_{0 < s_t < \mu^* \cdot \frac{\sqrt{2}-1}{2}} \left[ f_{\min}(S_{\exp} - \delta_{\text{est}}, s_t) - \frac{1}{\sqrt{n}} \zeta_\mu(s_t, \varepsilon_s, \varepsilon_{\text{EA}}) \right].
\]

**Theorem 33** (Main). Let \( D \) be any device, \( \rho \) the state (as defined in Equation (18)) generated using Protocol 2, \( \Omega \) (as defined in Equation (19)) the event that the protocol does not abort, and \( \rho_{\Omega} \) the state conditioned on not aborting. Then, for any \( \varepsilon_{\text{EA}}, \varepsilon_s \in (0, 1) \), either the protocol aborts with probability greater than \( 1 - \varepsilon_{\text{EA}} \) or

\[
H_{\min}^{\varepsilon_s}(A^n B^n | X^n Y^n E)_{\rho_{\Omega}} > n \cdot \eta_{\text{opt}}(\varepsilon_s, \varepsilon_{\text{EA}}, S_{\exp} - \delta_{\text{est}}, n, \mu)
\]

where \( \eta_{\text{opt}} \) is defined in Equation (25).

The entropy bound from Theorem 33 is plotted in Figure 6.

**Proof of Theorem 33.** We begin the proof by devising a min-tradeoff function for the EAT channels. We then proceed to lower bound the smooth min-entropy by employing the EAT with the given min-tradeoff function.

**Claim 1.** Let \( \{N_i\} \) be the set of EAT channels implemented in Protocol 2. Then, for any \( 0 < s_t < \mu^* \cdot \frac{\sqrt{2}-1}{2} \), where \( \mu^* = \mu_{\min} \cdot \mu_{\max} \), the function (23) is a min-tradeoff function for the set \( \{N_i\} \).

**Proof of Claim 1.** Note that, due to Assumption 5, each \( N_i \) describes a single MDL experiment (as described in Chapter 3). Thus, employing the bound from Equation (10), it follows directly that

\[
\inf_{\sigma_{B_{r-1}B_t} : N_i(\sigma)_{C_i} = p} H(A_i B_t | X_i Y_i R_i)_{N_i(\sigma)} \geq 1 - h \left( \frac{1}{2} + \frac{1}{\mu^*} \sqrt{S_\mu(p) (S_\mu(p) + \mu^*)} \right).
\]

Let \( C = \{-\mu_{\max}, 0, \mu_{\min}\} \) and define the function \( g_\mu \) on \( \mathbb{P}_C \) as

\[
g_\mu(p) = \begin{cases} 
1 - \left( \frac{1}{2} + \frac{1}{\mu^*} \sqrt{S_\mu(p) (S_\mu(p) + \mu^*)} \right) & \text{for } \frac{S_\mu(p)}{\mu^*} \in \left[ 0, \frac{\sqrt{2}-1}{2} \right), \\
1 & \text{for } \frac{S_\mu(p)}{\mu^*} \in \left[ \frac{\sqrt{2}-1}{2}, 1 \right].
\end{cases}
\]

Then any function \( f_{\min}(p) \), that is differentiable and satisfies \( f_{\min}(p) \leq g_\mu(p) \) for all \( p \), is a min-tradeoff function for the set \( \{N_i\} \). Unfortunately, as \( \frac{S_\mu(p)}{\mu^*} \) approaches \( \frac{\sqrt{2}-1}{2} \), the derivative of \( g_\mu \) diverges. Since the bound that we derive later depends on the derivative, we want to avoid this. Therefore we linearize \( g_\mu \) starting at some point \( p_t \) with \( \frac{S_\mu(p_t)}{\mu^*} < \frac{\sqrt{2}-1}{2} \) and thus avoid the problem of a diverging derivative.

Consider the change of variables

\[
s = S_\mu(p) \\
t = \mu_{\max} \cdot p(1) + \mu_{\min} \cdot p(-1).
\]

27
In this orthogonal coordinate system we clearly see that \( g_\mu(s, t) \) is independent of \( t \) and only changes with \( s \). Thus we can restrict our attention in analysing \( g_\mu(s, t) \) to a slice where \( t \) is constant. The divergence of the derivative now happens as \( \frac{\mu}{\mu^*} \) approaches \( \sqrt{\frac{3}{2}} \). Hence we linearize \( g_\mu \) at some point \( s_t < \mu^* \cdot \sqrt{\frac{3}{2}} \).

For the linearization we define
\[
a(s_t) = \frac{d}{ds} g_\mu(s) \bigg|_{s_t} \quad \text{and} \quad b(s_t) = g(s_t) - a(s_t) \cdot s_t.
\]

Given these constants we can define the function \( f_{\min} \) as
\[
f_{\min}(s, s_t) = \begin{cases} 
g_\mu(s) & \text{for } s \leq s_t \\
a(s_t) \cdot s + b(s_t) & \text{for } s > s_t.
\end{cases}
\]

Note that this is technically not yet a min-tradeoff function, since it is a function taking arguments in \( \mathbb{R} \) instead of \( P_C \). However, expressing the new variables as a function of \( p \) we can get the final min-tradeoff function as
\[
f_{\min}(p, s_t) = \begin{cases} 
g_\mu(p) & \text{for } s(p) \leq s_t \\
a(s_t) \cdot s(p) + b(s_t) & \text{for } s(p) > s_t.
\end{cases}
\]

Note also that this is a min-tradeoff function for all \( 0 < s_t < \mu^* \cdot \sqrt{\frac{3}{2}} \). Hence, when we derive the entropy bound for Protocol 2 we can optimize over the parameter \( s_t \) to get the best possible bound.

Now that we have a min-tradeoff function we can continue to derive a lower bound on the smooth min-entropy. As stated in Lemma 32 the channels in the protocol are EAT channels and we can employ the EAT. Furthermore we realize that the event \( \Omega \) of the protocol not aborting implies that the estimation for the MDL violation is at least \( S_{\exp} - \delta_{\est} \), i.e.,
\[
S_\mu(f_{\text{EAT}}) \geq S_{\exp} - \delta_{\est}
\]
for any \( C^n \) for which \( \Pr[C^n]_{p, \Omega} > 0 \). Thus, employing the EAT, we find that either the protocol aborts with probability \( 1 - \Pr(\Omega) \geq 1 - \varepsilon_{\text{EA}} \) or, the lower bound
\[
H_{\min}^\varepsilon(A^n B^n | X^n Y^n E)_{p, \Omega} > n f_{\min}(S_{\exp} - \delta_{\est}, s_t) - \sqrt{n} \zeta_\mu(s_t, \varepsilon_\mu, \varepsilon_{\text{EA}}),
\]
holds. Here we introduced \( \zeta_\mu(s_t, \varepsilon_\mu, \varepsilon_{\text{EA}}) = 2(\log(9) + a(s_t) \cdot \mu_{\max}) \sqrt{1 - 2\log(\varepsilon_\mu \varepsilon_{\text{EA}})} \), where we used that \( \|f_{\min}\|_\infty = a(s_t) \cdot \mu_{\max} \) due to the linearization in the direction of the steepest slope. Additionally, in the description of the lower bound we used the argument \( S_{\exp} - \delta_{\est} \) as shorthand to denote any probability distribution with \( S_\mu(p) = S_{\exp} - \delta_{\est} \). We can use this abbreviated notation because for all \( p \) with fixed \( S_\mu(p) \), the value of the min-tradeoff function is the same. Furthermore, the fact that \( f_{\min}(p) \) is constant as long as \( S_\mu(p) \) is constant is also the reason that, in the EAT, we can set \( t = f_{\min}(S_{\exp} - \delta_{\est}, s_t) \) in our lower bound.

Since \( s_t \) is chosen arbitrarily, we can optimize over it. For the final entropy bound define
\[
\eta_{\text{opt}}(\varepsilon_\mu, \varepsilon_{\text{EA}}, S_{\exp} - \delta_{\est}, n, \mu) = \max_{0 < s_t < \mu^* \cdot \sqrt{\frac{3}{2}}} \left[ f_{\min}(S_{\exp} - \delta_{\est}, s_t) - \frac{1}{\sqrt{n}} \zeta_\mu(s_t, \varepsilon_\mu, \varepsilon_{\text{EA}}) \right].
\]
Thus, the entropy bound reduces to
\[
H_{\min}^\varepsilon(A^n B^n | X^n Y^n E)_{p, \Omega} > n \cdot \eta_{\text{opt}}(\varepsilon_\mu, \varepsilon_{\text{EA}}, S_{\exp} - \delta_{\est}, n, \mu).
\]

As stated before, in Figure 6 the entropy rate, \( \eta_{\text{opt}}(\varepsilon_\mu, \varepsilon_{\text{EA}}, S_{\exp} - \delta_{\est}, n, \mu) \), plotted for several different parameters of the RAP. In Figure 6a the asymptotic rates are equal to the single round entropy bounds of Figure 4a with corresponding \( \mu \). Furthermore, we observe that, as expected, the entropy rate decreases for a decreasing
number of rounds, $n$. If the number of rounds decreases below a certain threshold, we do not get a non-trivial (positive) entropy bound anymore.

In Figure 6b we see that, as was the case for a single MDL experiment, the maximal entropy (rate) decreases as the bias of the source increases. Moreover, we observe that, as $n$ decreases, the minimal MDL violation to achieve a non-trivial entropy bound increases. Therefore, we can compensate imperfections in the implementation (which lead to a decreasing MDL violation) with an increasing number of rounds.

4.5.2 Applying the extractor

So far we gave an explicit lower bound on the smooth min-entropy of the entropy accumulation routine’s output. The last part in our RAP, that produces the final bits, is the application of a randomness extractor (cf. Section 2.8). More precisely we are using a quantum-proof two-source randomness extractor in the Markov model.

Using a two-source extractor we need, as the name indicates, two inputs. The first input that we use is the outputs generated by the entropy accumulation routine, $A^n B^n$. As the second input to the extractor we use additional raw bits from the source. $^{10}$ Thus we first use the source to produce the inputs to the MDL experiment and then to draw inputs, $Z^d$, for the extractor directly. The exact setup that we use for that is described in Section 4.1.

Since we are using extractors that work in the Markov model we require that the two inputs to the extractor are independent conditioned on the adversary’s side information, i.e., $I(Z^d : A^n B^n | X^n Y^n E, \lambda) = 0$. The fact that this is indeed the case in our setting is explicitly stated in Section 4.1. Hence we can use the extractor to quantify the secrecy of the outputs.

Remark 34. When using $I(Z^d : A^n B^n | X^n Y^n E, \lambda) = 0$ we assume that the adversary has full access to $E$ and $\lambda$. However in a realistic setting this might not be the case, thus leading to the Markov condition not being satisfied. Nevertheless, as stated in Section 5.2 in [AFPS16], the deletion of a part of the side information cannot decrease the security of the extractor. Consequently, if the adversary is less powerful and does not have access to all of $E$ and $\lambda$, and thus the Markov condition is not satisfied, the extractor is still secure.

The quality of the extractor’s output depends on the (smooth) min-entropy of the two sources. Thus, in addition to the mutual information vanishing, we also need to know what the min-entropy of the random variables $Z^d$ is.

Lemma 35. Let $Z^d$ be the output of a $\mu$-MDL source. Then, the lower bound

$$H_{\min}(Z^d | X^n Y^n E, \lambda) \geq -d \cdot \log(\mu_{\max})$$

on the min-entropy holds.

Proof. For a $\mu$-MDL source we require that the guessing probability of the outputs is bounded (recall form Section 4.1),

$$\mu_{\min} \leq p_{\text{guess}}(Z_{2i} | Z_{2i+1} Z_{2i-1} E, \lambda) \leq \mu_{\max} \forall \lambda, i.$$

Thus the maximal probability of any particular string appearing is at most $\mu_{\max}^{d/2}$. Finally, since the min-entropy is the negative logarithm of the maximal guessing probability, the lemma follows.

Using the results from [AFPS16], Theorem 33, and Lemma 35 we can determine how close to uniform the output of our RAP is.

---

$^{10}$The first part of the source’s output is used as input for the entropy accumulation routine and the second part as second input for the extractor.
Figure 6: Plots of the entropy rate, $\eta_{opt}$, as a function of the expected MDL violation, $S_{exp}$, for different parameters of the EAP. In addition to the noted values of $n$, $\mu$, and $\delta_{est} = 10^{-5}$, we used the parameters $\varepsilon_s = \varepsilon_{EA} = 10^{-7}$. As in Figure 4a the black dots indicate the maximal possible MDL violation for given $\mu$. Note that in Figure 6a the asymptotic rates are equal to the curves in Figure 4a with corresponding $\mu$. Furthermore the curves in Figure 6b converge to the curves in Figure 4a. Both these facts are a consequence of the asymptotic equipartition property. For convenience we choose $\mu$ such that $\mu_{max} = 1 - 3 \cdot \mu_{min}$.

(a) Entropy rate, $\eta_{opt}$, as function of $S_{exp}$ for different $n$ and, in each plot, fixed $\mu$.

$$n = 10^{11}, \delta_{est} = 10^{-4}$$

(b) Entropy rate, $\eta_{opt}$, as function of $S_{exp}$ for different $\mu$ and, in each plot, fixed $n$.

$$n = 5 \cdot 10^{8}, \delta_{est} = 10^{-4}$$
Lemma 36. Let Ext : \{0, 1\}^{2n} \times \{0, 1\}^d \rightarrow \{0, 1\}^m be a \(k_1, k_2, \varepsilon_{\text{ext}}\) be a two-source quantum-proof extractor in the Markov model, strong in the second input, such that
\[
\begin{align*}
k_1 &\leq n \cdot \eta_{\text{opt}}(\varepsilon_{\text{EA}}, \varepsilon_s, S_{\text{exp}} - \delta_{\text{est}}, n, \mu) - \log(1/\varepsilon_{\text{ext}}) - 1 \\
k_2 &\leq -\frac{d}{2} \cdot \log(\mu_{\text{max}}) - \log(1/\varepsilon_{\text{ext}}) - 1
\end{align*}
\]

Consider the RAP (Protocol 2) using Ext and any \(\varepsilon_{\text{EA}}, \varepsilon_s \in (0, 1)\). Then, either the protocol aborts with probability greater than \(1 - \varepsilon_{\text{EA}}\), or for the output \(K^n\) together with the whole information the adversary possibly has access to, \(\Sigma = Z^d X^n Y^n E\lambda\), it holds that
\[
\frac{1}{2} \|\rho_{K^m \Sigma} - \rho_{U_m} \otimes \rho_{\Sigma}\| \leq 6(\varepsilon_s + \varepsilon_{\text{ext}})
\]

Proof. Starting with Theorem 33, we know that, either the protocol aborts with probability greater than \(1 - \varepsilon_{\text{EA}}\), or the smooth min-entropy of the entropy accumulation routine’s output is lower bounded by \(n \cdot \eta_{\text{opt}}(\varepsilon_{\text{EA}}, \varepsilon_s, S_{\text{exp}} - \delta_{\text{est}}, n, \mu)\). For the second input of the extractor, using Lemma 35, we know that the min-entropy of the string \(Z^d\) is lower bounded by \(\frac{d}{2} \cdot \log(1/\mu_{\text{max}})\). Furthermore by assumption the state that is generated in the protocol is a Markov state. Thus we can employ Lemma 21 to get an upper bound on the distance between the output \(K^n\) and a uniform string, and proof the claim.

Remark 37. As stated in Lemma 20, one can construct two-source quantum-proof extractors in the Markov model from classical ones. The parameters of the chosen extractor affect the parameters of our protocol directly. In particular, the security parameter (given below) and the efficiency of the protocol (the extraction rate) depend on the extractor. It is important to note that there are explicit extractors with good parameters for our purpose. In particular:

1. If one of the two sources (either the device’s outputs \(A^n B^n\) or the seed \(Z^d\) for \(d = 2n\)) has (smooth) min-entropy of more than \(n\) one can use the explicit construction of an extractor given in [AFPS16, Corollary 25] to extract a linear number of bits. Focusing on the the seed, the min-entropy is sufficiently high when \(\mu_{\text{max}} \leq 1/2\).\(^\text{11}\)

2. Otherwise, one can use the explicit construction of an extractor given in [AFPS16, Corollary 30] to extract a logarithmic number of bits.

3. To extract a sub-linear number of bits using an explicit extractor one can also consider a simple modification of our protocol, similarly to what was done in [BRG+ 16, Theorem 2] - given another device with two components one can use the inputs to run the same protocol with the additional device and by this create another source of randomness. Combined with what we had before, we now have three sources of randomness (the outputs of the two devices and the seed) in the Markov model (see [AFPS16, Definition 7]) . Thus, the three-source extractor given in [AFPS16, Corollary 28] can be used to extract a sub-linear number of bits.

After putting everything together we can determine the secrecy parameter for our RAP corresponding to the secrecy definition (Definition 29). In the final theorem we state \(\varepsilon_{\text{RA}}\) in terms of the RAP’s parameters.

Theorem 38 (Secrecy). For any \(\varepsilon_{\text{EA}}, \varepsilon_s \in (0, 1)\) the RAP (Protocol 2) with the given parameters is \(\varepsilon_{\text{RA}}\)-secret (according to Definition 29), with \(\varepsilon_{\text{RA}} = 12(\varepsilon_s + \varepsilon_{\text{ext}}) + \varepsilon_{\text{EA}}\).

Proof. In the following let \(\Sigma = Z^d X^n Y^n E\lambda\) be the whole information the adversary has access to. Starting with Lemma 36 we can distinguish two cases.

Case 1. The protocol aborts with probability greater than \(1 - \varepsilon_{\text{EA}}\).

\(^{11}\)In terms of an SV-source, the source should be such that, roughly, \(0.3 \leq p(0) \leq 0.7\); recall Lemma 15.
In that case, we find

\[(1 - \Pr[\text{abort}]) \|\rho_{K^m \Sigma} - \rho_{U^m} \otimes \rho_{\Sigma}\| \leq \varepsilon_{\text{EA}} \|\rho_{K^m \Sigma} - \rho_{U^m} \otimes \rho_{\Sigma}\| \leq \varepsilon_{\text{EA}},\]

since the trace distance is always less than one.

**Case 2.** The protocol aborts with probability less than \(1 - \varepsilon_{\text{EA}}\) (hence the entropy is sufficiently high).

In that case, using the bound from Lemma 36, we find

\[(1 - \Pr[\text{abort}]) \|\rho_{K^m \Sigma} - \rho_{U^m} \otimes \rho_{\Sigma}\| \leq \|\rho_{K^m \Sigma} - \rho_{U^m} \otimes \rho_{\Sigma}\| \leq 12 (\varepsilon_s + \varepsilon_{\text{ext}}).\]

We can now continue to prove Theorem 30.

**Proof of Theorem 30.** Part 1 follows directly from the proof of Theorem 38. Part 2 follows directly from Lemma 31.

## 5 Open questions

We end with some open questions:

1. Is the amount of extractable randomness given in our work tight? There are few things that one can consider when trying to improve the extraction rate:
   i. While the bound given in Lemma 27 is non-trivial for any violation of the MDL inequality, it might not be tight.
   ii. We used the MDL inequality derived in [PRB+14]. They derived their inequality with the motivation of detecting quantumness for an arbitrary MDL source. Thus it might be possible that there are other MDL inequalities that are more suitable for quantifying randomness.
   iii. The final length of the extracted randomness depends on the parameters of the extractor used. Finding quantum-proof multi-source extractors for the Markov model which have good parameters is therefore of interest. This can be achieved by considering better specific (classical) two-source extractors and then applying the technique of [AFPS16], or by improving over the parameters of [AFPS16] for general constructions.

2. Can the analysis be extended such that the adversary is allowed to hold some quantum side information about the source? Currently we only allow the adversary to know \(\lambda\) in advance (while \(E\) is the quantum side information about the device itself). In particular, this is a realistic assumptions in scenarios where the device and the producer of the weak source are different parties. Nevertheless, it will be interesting to see whether holding quantum side information about the source before producing the device is beneficial for the adversary and what the consequences for the security of our protocol are.

3. Is it possible to amplify min-entropy sources while maintaining similar parameters? In particular, can it be done with a constant number of devices? (in contrast to what was done in [CSW14]). The technique presented here does not work if the SV (MDL) source is replaced with a min-entropy sources (while it might be possible to extend them to block-sources). Thus, another approach has to be taken.

4. Similarly, is it possible to amplify randomness against a non-signalling adversary while maintaining similar parameters? Our RAP works only against an adversary that is bound by quantum mechanics and an extension to the non-signalling case is not possible using the techniques that we employed. In particular, the proofs of both [DFR16] and [AFPS16] use the assumption that everything can be described with the formalism of quantum physics. We remark that, while it might be possible to extend one of these results to the non-signalling case, an extension of both of them will result in a contradiction with [AFTS12]. Previous works that focused on non-signalling adversaries cannot be used to achieve similar statements as we derived in this work.
5. What is the effect of using a weak source of randomness in device-independent protocols that assume perfect randomness, e.g., device-independent quantum key distribution protocol or randomness expansion? In such protocols random bits are used not only for choosing the inputs for the devices, but also for choosing the rounds in which a “test” is carried out. To analyse the effect of replacing perfect randomness with weak randomness one can use our RAP. One trivial possibility to include our RAP would be to just use it separately to generate uniform bits, before starting with the other protocols. Another option is to use our protocol as the main building block for the test rounds. The test rounds themselves can then be chosen with the SV-source, by using techniques such as enumeration.

Acknowledgments

We thank Gilles Pütz for helpful discussions about the MDL inequalities and for letting us use his code to evaluate numerically the optimal violation of the inequalities within quantum physics. We also thank Jean-Daniel Bancal, Roger Colbeck, Christopher Portman, and Thomas Vidick for helpful comments. RAF was supported by the Swiss National Science Foundation via the National Center for Competence in Research, QSIT, and by the Air Force Office of Scientific Research (AFOSR) via grant FA9550-16-1-0245.

Appendices

A Finding the maximal quantum violation of an MDL inequality

It is not possible to find the maximal MDL value (\(S_\mu^*\)) in quantum mechanics for an MDL source with fixed \(\mu\) since this value depends on the specific probability distribution of the source. However, we can find a lower bound on \(S_\mu^*\) by taking the worst case probability distribution for a fixed \(\mu\). What we get is the value \(\tilde{S}_\mu\) defined as:

\[
\tilde{S}_\mu \equiv \mu_{\text{min}}^2 P_{AB|XY}(00|00) - \mu_{\text{max}}^2 \left( P_{AB|XY}(01|01) + P_{AB|XY}(10|10) + P_{AB|XY}(00|11) \right).
\] (35)

The value \(\tilde{S}_\mu\) is a lower bound on \(S_\mu\) that is independent of the source as long as \(\mu\) is fixed. Therefore, when we find the maximum of \(\tilde{S}_\mu\) in quantum mechanics (\(\tilde{S}_\mu^*\)) we also get lower bound on \(S_\mu^*\).

Lemma 39. For fixed state and measurements \(\tilde{S}_\mu\) is a lower bound for \(S_\mu\) (as defined in Equation 9).

Proof. First note that with \(\mu_{\text{min}} \leq P_{XY|\Sigma}(xy|\sigma) \leq \mu_{\text{max}} \forall x, y, \sigma\) and \(P_{XY} = \sum_{\sigma} P_{\Sigma}(\sigma) P_{XY|\Sigma}(xy|\sigma)\) it also holds that

\[
\mu_{\text{min}} \leq P_{XY}(xy) \leq \mu_{\text{max}} \forall x, y.
\] (36)

Employing these bounds we find

\[
S_\mu = \mu_{\text{min}} P_{ABXY}(0000) - \mu_{\text{max}} \left( P_{ABXY}(0101) + P_{ABXY}(1010) + P_{ABXY}(0011) \right)
\]

\[
= \mu_{\text{min}} P_{XY}(00) P_{AB|XY}(00|00) - \\
\geq \mu_{\text{min}} P_{XY}(01) P_{AB|XY}(01|01) + P_{XY}(10) P_{AB|XY}(10|10) + P_{XY}(11) P_{AB|XY}(00|11)\]

\[
\geq \mu_{\text{min}}^2 P_{AB|XY}(00|00) - \mu_{\text{max}}^2 \left( P_{AB|XY}(01|01) + P_{AB|XY}(10|10) + P_{AB|XY}(00|11) \right) = \tilde{S}_\mu
\]

\[\Box\]

For fixed \(\mu\) the probability distribution for the source’s outputs is not necessarily fixed.
We found $\hat{S}^*_{\mu}$ by maximising the eigenvalue of the Bell operator as a function of the measurement parameters in Matlab. For a Bell inequality $\sum_{a,b,x,y} \alpha_{abxy} P_{AB|XY}(ab|xy) \leq p_{\text{local}}$ with parameters $\alpha_{abxy}$ and measurement operators $\{M^x_a\}_a$ and $\{M^y_b\}_b$ the Bell operator is defined as

$$B = \sum_{a,b,x,y} \alpha_{abxy} M^x_a \otimes M^y_b.$$ 

## B Additional proofs

### Proof of Lemma 28

First of all we have

$$H(AB|XYFE) \geq H(A|XYFE) = H(A|XFE).$$

Here the first step follows because $A$ and $B$ are classical registers. The second step follows because the non-signalling condition holds between the two components of the device. Thus the dependence of $A$ on $Y$ can only be through $X$; i.e., $A$, $X$, and $Y$ form a Markov chain, $A \leftrightarrow X \leftrightarrow Y$. Furthermore, it holds that

$$H(A|XFE) = \sum_x \Pr[X = x] \cdot H(A|FE, X = x).$$

Finally we can rewrite

$$H(A|FE, X = x) = H(AFE|X = x) - H(FE|X = x)$$

$$= H(A|X = x) + H(FE|A, X = x) - H(FE|X = x)$$

$$= H(A|X = x) - \chi(A:F|X=x)$$

$$= 1 - \chi(A:F|X=x),$$

where we used the fact that the outputs are symmetrized (Step 3) and we introduced the Holevo quantity

$$\chi(A:F|X=x) = H(FE|X=x) - H(FE|A, X = x).$$

Combining everything, the result follows. $\square$

### References

- **[AFPS16]** R. Arnon-Friedman, C. Portmann, and V. B. Scholz. Quantum-Proof Multi-Source Randomness Extractors in the Markov Model. In *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 fuer Informatik. [arXiv:1510.06743](http://arxiv.org/abs/1510.06743), [doi:10.4230/LIPIcs.TQC.2016.2](http://dx.doi.org/10.4230/LIPIcs.TQC.2016.2).

- **[AFRV16]** R. Arnon-Friedman, R. Renner, and T. Vidick. Simple and tight device-independent security proofs. *ArXiv e-prints*, Jul 2016. [arXiv:1607.01797](http://arxiv.org/abs/1607.01797).

- **[AFTS12]** R. Arnon-Friedman and A. Ta-Shma. Limits of privacy amplification against nonsignaling memory attacks. *Physical Review A*, 86(6):062333, 2012.

- **[AM16]** A. Acín and L. Masanes. Certified randomness in quantum physics. *Nature*, 540(7632):213–219, 2016.

- **[BAK+16]** M. N. Bera, A. Acín, M. Kuś, M. Mitchell, and M. Lewenstein. Randomness in quantum mechanics: Philosophy, physics and technology. *ArXiv e-prints*, 2016. [arXiv:1611.02176](http://arxiv.org/abs/1611.02176).

- **[BCP+14]** N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. *Reviews of Modern Physics*, 86(2):419, 2014.
[Bel64] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1(3):195–200, 1964.

[BHK05] J. Barrett, L. Hardy, and A. Kent. No signaling and quantum key distribution. *Physical Review Letters*, 95(1):010503, 2005.

[BRG+16] F. G. S. L. Brandão, R. Ramanathan, A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, T. Szarek, and H. Wojewódka. Realistic noise-tolerant randomness amplification using finite number of devices. *Nature communications*, 7, 2016.

[CG88] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, 1988. doi:10.1137/0217015.

[CHSH69] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* , 23:880–884, Oct 1969. doi:10.1103/PhysRevLett.23.880.

[CLW14] K.M. Chung, X. Li, and X. Wu. Multi-source randomness extractors against quantum side information, and their applications. *arXiv preprint arXiv:1411.2315*, 2014.

[CR12] R. Colbeck and R. Renner. Free randomness can be amplified. *Nat Phys*, 8(6):450–453, Jun 2012. arXiv:1105.3195, doi:10.1038/nphys2300.

[CSW] K.-M. Chung, Y. Shi, and X. Wu. General randomness amplification with non-signaling security.

[CSW14] K.-M. Chung, Y. Shi, and X. Wu. Physical Randomness Extractors: Generating Random Numbers with Minimal Assumptions. *ArXiv e-prints*, Feb 2014. arXiv:1402.4797.

[DFR16] F. Dupuis, O. Fawzi, and R. Renner. Entropy accumulation. *ArXiv e-prints*, Jul 2016. arXiv:1607.01796.

[DH76] W. Diffie and M. Hellman. New directions in cryptography. *IEEE transactions on Information Theory*, 22(6):644–654, 1976.

[DPVR12] A. De, C. Portmann, T. Vidick, and R. Renner. Trevisan’s extractor in the presence of quantum side information. *SIAM Journal on Computing*, 41(4):915–940, 2012.

[DW05] I. Devetak and A. Winter. Distillation of secret key and entanglement from quantum states. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 461, pages 207–235. The Royal Society, 2005.

[Ebe93] P. H. Eberhard. Background level and counter efficiencies required for a loophole-free einstein-podolsky-rosen experiment. *Phys. Rev. A*, 47:R747–R750, Feb 1993. doi:10.1103/PhysRevA.47.R747.

[Eke91] Artur K. Ekert. Quantum cryptography based on Bell’s theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991. doi:10.1103/PhysRevLett.67.661.

[ER14] A. Ekert and R. Renner. The ultimate physical limits of privacy. *Nature*, 507(7493):443–447, 2014.

[GLLL+11] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov. Full-field implementation of a perfect eavesdropper on a quantum cryptography system. *Nature communications*, 2:349, 2011.

[GMD+13] R. Gallego, L. Masanes, G. De La Torre, C. Dhara, L. Aolita, and A. Acín. Full randomness from arbitrarily deterministic events. *Nature Communications*, 4:2654 EP –, Oct 2013. Article.
[Gol10] O. Goldreich. *A primer on pseudorandom generators*, volume 55. American Mathematical Society, 2010.

[GVW+15] M. Giustina, Marijn A. M. V., S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J. Larsson, C. Abellán, et al. Significant-loophole-free test of bells theorem with entangled photons. *Physical review letters*, 115(25):250401, 2015.

[HBD+15] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, 2015.

[KK10] R. Kasher and J. Kempe. Two-source extractors secure against quantum adversaries. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 656–669. Springer, 2010.

[KRS09] R. Konig, R. Renner, and C. Schaffner. The operational meaning of min-and max-entropy. *IEEE Transactions on Information theory*, 55(9):4337–4347, 2009.

[MS16] C. A. Miller and Y. Shi. Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices. *Journal of the ACM (JACM)*, 63(4):33, 2016.

[MY98] D. Mayers and A. Yao. Quantum cryptography with imperfect apparatus. In *Foundations of Computer Science, 1998. Proceedings. 39th Annual Symposium on*, pages 503–509. IEEE, 1998. doi:10.1109/SFCS.1998.743501.

[NC10] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 1 edition, January 2010. doi:10.1119/1.1463744.

[Neu27] J. von Neumann. Thermodynamik quantenmechanischer gesamtheiten. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1927:273–291, 1927. URL: http://eudml.org/doc/59231.

[NTS99] N. Nisan and A. Ta-Shma. Extracting randomness: A survey and new constructions. *Journal of Computer and System Sciences*, 58(1):148–173, 1999.

[PAB+09] S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani. Device-independent quantum key distribution secure against collective attacks. *New Journal of Physics*, 11(4):045021, Apr 2009. arXiv:0903.4460, doi:10.1088/1367-2630/11/4/045021.

[PRB+14] G. Pütz, D. Rosset, T. J. Barnea, Y. C. Liang, and N. Gisin. Arbitrarily small amount of measurement independence is sufficient to manifest quantum nonlocality. *Phys. Rev. Lett.*, 113:190402, Nov 2014. arXiv:1407.5634, doi:10.1103/PhysRevLett.113.190402.

[Raz05] R. Raz. Extractors with weak random seeds. In *Proceedings of the 37th Symposium on Theory of Computing, STOC ’05*, pages 11–20. ACM, 2005. doi:10.1145/1060590.1060593.

[RBH+15] R. Ramanathan, F. G. S. L. Brandão, K Horodecki, M. Horodecki, P. Horodecki, and H. Wojewódka. Randomness amplification against no-signaling adversaries using two devices. *ArXiv e-prints*, 2015. arXiv:1504.06313.

[Ren05] R. Renner. *Security of Quantum Key Distribution*. PhD thesis, December 2005. arXiv: quant-ph/0512258.

[Sha48] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, July 1948. doi:10.1002/j.1538-7305.1948.tb01338.x.

[Sha83] A. Shamir. On the generation of cryptographically strong pseudorandom sequences. *ACM Transactions on Computer Systems (TOCS)*, 1(1):38–44, 1983.
[SMSC+15] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, et al. Strong loophole-free test of local realism. *Physical review letters*, 115(25):250402, 2015.

[SV84] M. Santha and U. V. Vazirani. Generating quasi-random sequences from slightly-random sources. In *25th Annual Symposium on Foundations of Computer Science, 1984.*, pages 434–440, Oct 1984. doi:10.1109/SFCS.1984.715945.

[TCR09] M. Tomamichel, R. Colbeck, and R. Renner. A fully quantum asymptotic equipartition property. *Information Theory, IEEE Transactions on*, 55(12):5840–5847, 2009.

[TCR10] M. Tomamichel, R. Colbeck, and R Renner. Duality between smooth min- and max-entropies. *IEEE Transactions on Information Theory*, 56(9):4674–4681, 2010. arXiv:0907.5238, doi:10.1109/TIT.2010.2054130.

[V+12] S. P. Vadhan et al. Pseudorandomness. *Foundations and Trends® in Theoretical Computer Science*, 7(1–3):1–336, 2012.

[WBG+16] H. Wojewodka, F. G. S. L. Brandao, A. Grudka, M. Horodecki, K. Horodecki, P. Horodecki, M. Pawlowski, and R. Ramanathan. Amplifying the randomness of weak sources correlated with devices. *ArXiv e-prints*, 2016. arXiv:1601.06455.