Device-independent Randomness Extraction for Arbitrarily Weak Min-entropy Source

Jan Bouda[1,2,3], Marcin Pawłowski[4,5], Matej Pivoluska[1], and Martin Plesch[1,6]

1 **Faculty of Informatics, Masaryk University**
Botanická 68a, 602 00 Brno, Czech Republic
bouda@fi.muni.cz
2 **Física Teórica: Informació i Fenómens Quántics Universitat Autónoma de Barcelona**
08193 Bellaterra (Barcelona), Spain
3 **LIQUID: Lepanto Institute for Quantum Information and Decoherence**
Carrer de Lepant 307, 08025 Barcelona, Spain
4 **Instytut Fizyki Teoretycznej i Astrofizyki, Uniwersytet Gdański**
PL-80-952 Gdańsk, Poland
dokmpa@univ.gda.pl
5 **School of Mathematics, University of Bristol**
Bristol BS8 1TW, United Kingdom
6 **Institute of Physics, Slovak Academy of Sciences**
Bratislava, Slovakia
martin.plesch@savba.sk

Abstract

In this paper we design a protocol to extract random bits with an arbitrarily low bias from a single arbitrarily weak min-entropy block source in a device independent setting. The protocol employs Mermin devices that exhibit super-classical correlations. Number of devices used scales polynomially in the length of the block n, containing entropy of at least two bits. Our protocol is robust, it can tolerate devices that malfunction with a probability dropping polynomially in n at the cost of constant increase of the number of devices used.

1998 ACM Subject Classification G.3 Random Number Generation

Keywords and phrases Randomness Extraction, Device Independence

Digital Object Identifier 10.4230/LIPIcs.TQC.2014.205

1 Introduction

High quality randomness is a very useful resource in many computation and cryptographic tasks. In fact it has been shown that many protocols, including quantum ones, vitally require perfect randomness for their security [1, 2].

Unfortunately, even though we cannot fully predict certain processes it is very difficult to argue that they produce *perfect randomness* – independent and unbiased bits. The problem of imperfect randomness has a long history in classical computer science and long line of research was devoted to randomness extraction – algorithms to transform imperfect sources of randomness into (close to) perfect ones [3].

The drawback of randomness extractors are twofold. Firstly, extractors typically require at least two independent sources of (imperfect) randomness. Worse still, even imperfect randomness of classical processes has to be assumed, because in principle classical physics is


9th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC'14).
Editors: Steven T. Flammia and Aram W. Harrow; pp. 205–211

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

deterministic. Quantum physics, with its intrinsic randomness allows us, in theory, to drop the second assumption. Preparation of a pure state and measurement in its complementary basis will yield a perfectly random result. In practice, however, we are replacing the assumption of randomness by yet another assumption – perfect control of quantum devices. This assumption is also very problematic, as we have learned in case of quantum key distribution [4].

Luckily enough, thanks to Bell-type experiments, it is possible to certify by classical procedures that quantum processes are being observed and therefore intrinsic randomness is being produced. This is the basic idea behind device independent randomness extraction. Effectively, we are exchanging the assumption of independent randomness of the second source by a much weaker assumption – validity of quantum mechanics. Alternatively, one can view device-independent randomness extraction as quantum protocol for extracting randomness from a single weak source – a task that is classically impossible.

In this paper we work with (n, k)block min-entropy random sources. These are sources with n-bit blocks of output with guaranteed min-entropy k. Such a source can be modeled as a sequence of n-bit random variables $X_1, X_2, \ldots$, such that

$$\begin{aligned} &\forall x_1, \ldots, x_{i-1} \in \{0,1\}^n, \forall e \in \mathcal{I}(E), \\ &H_\infty(X_i | X_{i-1} = x_{i-1}, \ldots, X_1 = x_1, E = e) \geq k, \end{aligned} \tag{1}$$

where E is a random variable describing all adversary's information about the source and $\mathcal{I}(E)$ is it's image. Therefore, each new block has high min-entropy, even conditioned on the previous ones and any information of the adversary. This is a generalization of Santha-Vazirani sources [5], which can be viewed as block sources with $n = 1$.

Note that the task of transforming a single block source into a fully random bit is known to be impossible [3]. Furthermore, it is impossible to turn a block source with $n > 1$ into Santha-Vazirani source, therefore we cannot use existing randomness extraction protocols [6, 7, 8, 9].

It is also worth to note that similar results were independently obtained by Chung, Shi and Wu [10]. The main difference between the two results is that we work with min-entropy block sources, while their results hold also for general min-entropy sources.

2 Device Independent Concept and Mermin Inequality

In this paper we use the three-partite Mermin inequality. Let's consider three non–communicating boxes, each of them having a single bit input and a single bit output. Let us denote the input bits of the respective boxes by X, Y and Z and the corresponding output bits A, B and C. Input bits are correlated and it holds that $XYZ \in \{111, 100, 010, 001\}$. The inputs are simultaneously passed to all boxes, so each box only knows it's input. The value v of the Mermin term is a function of the 4 conditional probabilities defined by the behavior of the device and of the probability distribution on inputs

$$\begin{aligned} v = &P(A \oplus B \oplus C = 1 | XYZ = 111) P(XYZ = 111) + \\ &+ P(A \oplus B \oplus C = 0 | XYZ = 100) P(XYZ = 100) + \\ &+ P(A \oplus B \oplus C = 0 | XYZ = 010) P(XYZ = 010) + \\ &+ P(A \oplus B \oplus C = 0 | XYZ = 001) P(XYZ = 001). \end{aligned} \tag{2}$$

In particular, for the uniform input distribution we set $P(XYZ = 111) = P(XYZ = 010) = P(XYZ = 001) = P(XYZ = 100) = \frac{1}{4}$ and denote the Mermin term by v_u.

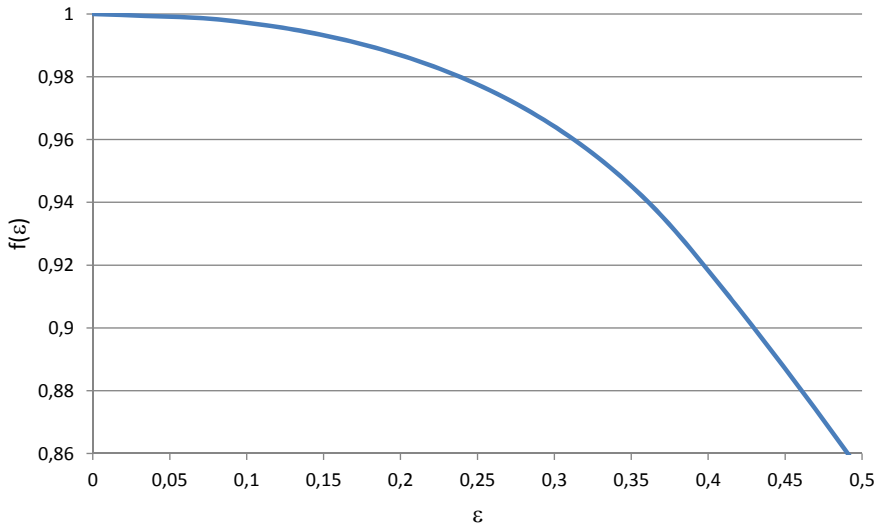

Figure 1 Depicted is the value of Mermin variable $v = f(\varepsilon)$ needed to certify the bias of the output bit to be at most ε.

Assuming the uniform distribution on all four inputs, the maximal value of v_u achievable by a classical device [11] is $\frac{3}{4}$ (thus the Mermin inequality reads $v_u \leq \frac{3}{4}$) and there exists a classical device that can make any 3 conditional probabilities simultaneously equal to 1. With the use of quantum mechanics we can achieve $v_u = 1$ and satisfy perfectly all 4 conditional probabilities using the tripartite GHZ state $\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$ and measuring σ_X (σ_Y) when receiving 0 (1) on input.

The beautiful property of the Mermin inequality is that the violation v gives us directly the probability that the device passes a specific test

$$A + B + C = X \cdot Y \cdot Z, \tag{3}$$

where addition and product are both taken modulo 2. The probability of failing the test is therefore $1 - v$.

Mironowicz and Pawlowski [9] showed the following result: Take a linearly ordered sequence of ℓ Mermin devices $D_1 \dots D_\ell$ (ℓ being arbitrary) that have uniform distribution on inputs, and each device knows inputs and outputs of its predecessors, but devices cannot signal to its predecessors. Let us assume that the inputs of devices are described by random variables $XYZ_1, \dots, XYZ_\ell$, and the outputs by $ABC_1, \dots, ABC_\ell$. Then there exists a function $f(\varepsilon)$ such that if the value of the Mermin term (2) using uniform inputs is at least $v_u \geq f(\varepsilon)$, then the output bit A_ℓ has a bias at most ε conditioned on the input and output of all its predecessors and the adversarial knowledge. This function can be lower bounded by a Semi-Definite Program (SDP) using any level of the hierarchy introduced in [12]. By using the second level of the hierarchy one can obtain the bound on $f(\varepsilon)$ as a function of ε shown in Fig. 1.

We can set $\ell = 1$ (having just a single device) and get the lower bound on the detection probability of producing a bit biased by more than ε, which is greater than $1 - f(\varepsilon)$. Our protocol uses many devices, which are forbidden to communicate at all, therefore they can be ordered arbitrarily and thus this limit holds for all of these devices simultaneously.

3 Single-round protocol

In the rest of our analysis we will be working with (n, k) block sources for an arbitrary n and $k \geq 2$. This is to simplify the explanation, since by taking $\lceil \frac{2}{k'} \rceil$ blocks of an arbitrary (n', k') source with $k' > 0$ we get a (n, k) source with $n = \lceil \frac{2}{k'} \rceil n'$ and $k = \lceil \frac{2}{k'} \rceil k' \geq 2$.

Let us start with a min-entropy $(n, 2)$ source (recall that (n, k) source with $k > 2$ is also an $(n, 2)$ source) and define $N = 2^n$. Let $H = \{h_1,$
$dots, h_m\}$ be a family of hash functions s.t. $h_i : \{0, \ldots, N-1\} \to \{0, 1, 2, 3\}$. Each hash-function h_i is used to provide input for a Mermin-type device D_i, where outputs of the function $0, 1, 2, 3$ identify $111, 100, 010, 001$ inputs for the device.

We want to construct H with the property that for every 4-element set $S \subseteq \{0, \ldots, N-1\}$ there exist at least one hash function $h \in H$ such that $h(S) = \{0, 1, 2, 3\}$. This is trivially satisfied for the set of all possible hashing functions $H_{full} = \{0, 1, 2, 3\}^N$, however, such a class of functions with its 4^N elements is impractically large. There exists a construction of such class of hash functions with logarithmic number of functions in N (see [13]), thus the number of devices needed scales polynomially with the length of the sequence n. We also stress that for large n one hash function covers as many as 9% of all four-tuples, independently on n. So the size of an optimal set of hash functions might not depend on n at all. Let us denote $m = |H|$. The protocol works as follows:

1. Obtain a (weakly) random n bit string r from the (n, k) block source.
2. Into each device D_i input the 3 bit string r_i chosen from set $\{111, 100, 010, 001\}$ – each one corresponding to one of the possible outputs of $h_i(r)$ – and obtain the outputs A_i, B_i and C_i.
3. Verify whether for each device D_i the condition $X_i + Y_i + Z_i = A_i \cdot B_i \cdot C_i$ holds. If this is not true, abort the protocol.
4. Output $b = \bigoplus_{i=1}^{m} A_i$.

The protocol is depicted in the Fig. 2.

Let us now examine the properties of the bit b_i. First consider only flat $(n, 2)$ distributions. Recall that these are exactly distributions that are uniform on 4-element subsets of the sample space. Our construction of the class H of hash functions assures that for any flat probability distribution there is a function $h_j \in H$ and the corresponding device D_j such that inputs of D_j (hashed by h_j) are uniform. Although output bits A_i are not independent in general, as most of them can be produced by fully deterministic strategies, (1) together with the arbitrary ordering we can impose on devices $\{D_i\}_{i=1}^m$ we have that if the adversary wants to bias (conditioned on the inputs and outputs of other devices) A_j by amount greater than ε, she must risk getting caught with probability at least $1 - f(\varepsilon)$. Therefore A_j is partially independent of other $\{A_i | i \neq j\}$, and the output of the round b is biased by at most ε with probability at least $1 - f(\varepsilon)$.

The set of all $(n, 2)$ distributions is convex and flat distributions are exactly all extremal points of this convex set [14]. Thus any $(n, 2)$ distribution d can be expressed as a convex combination of at most N $(n, 2)$ flat distributions d_i (Caratheodory theorem) as $d = \sum_{i=1}^{N} p_i d_i$ for some $p_i \geq 0$, $\sum_{i=1}^{N} p_i = 1$. The lower bound on probability that the adversary is not detected is given by the successful cheating probabilities when using flat distribution $d_i \in \{d_i\}_{i=1}^N$ averaged through the probability distribution on these flat distributions

$$v_u \leq \sum_{i=1}^{N} p_i P(\text{not detected} | d_i) \leq f(\varepsilon) \sum_{i=1}^{N} p_i \,. \tag{4}$$

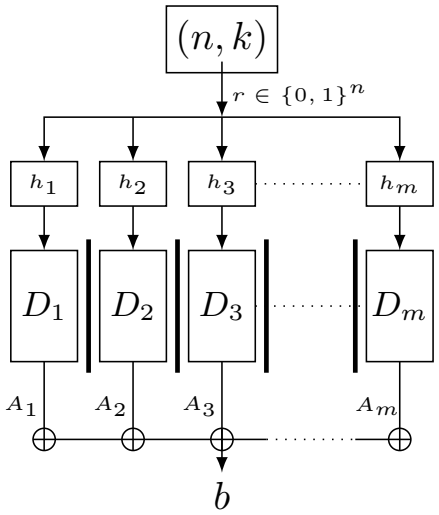

Figure 2 Depiction of a single round protocol. Bit string drawn from the flat random source is hashed into m inputs into Mermin devices so that at least one device receives all four inputs with non-zero probability. This guarantees at least one result almost perfectly random with high probability, which holds also for the product of individual results.

Thus the upper bound $v_u \leq f(\varepsilon)$ holds for non-flat distributions as well.

To summarize this part, having an (n, k) source with $k \geq 2$, with a single round of a protocol, we can produce a single bit that is biased at most by ε with a certainty of $1 - f(\varepsilon)$.

4 Multiple-round protocol

Let us state the most general task: we have an (n, k) block source with arbitrary n and $k \geq 2$. We would like to produce a bit that is biased by no more than ε with certainty of at least $1 - \delta$.

If the one-round version does not meet these parameters, we will repeat the whole protocol l times. By using new devices and new outputs of the block source, each of the runs j will produce a bit b_j that is biased by ε from perfectly random bit conditioned on all the previous bits $\{b_i | i < j\}$ up to a probability $f(\varepsilon)$. Thus, in order to achieve the bias of the output bit

$$b = \bigoplus_{j=1}^{l} b_j \tag{5}$$

of at least ε, all bits b_i has to have at least this bias. Therefore, after l rounds, the probability of the adversary not being detected will be upper bounded by $f(\varepsilon)^l$. Note that the product form does not come from the fact that the detection probabilities are independent (they are not). This is a product of a chain of conditional probabilities. Recall that the bound $f(\varepsilon)$ holds conditioned on any inputs and outputs of the previous devices (in an arbitrarily ordering that respects the causality). Thus choosing

$$l > \frac{\log \delta}{\log f(\varepsilon)} \tag{6}$$

will guarantee the fulfillment of the conditions for the parameters ε and δ.

Summing up, with an (n, k) block source and

$$O\left(\frac{\log \delta}{\log f(\varepsilon)} Poly\left[n\left\lceil\frac{2}{k}\right\rceil\right]\right) \tag{7}$$

Mermin devices we can produce a single random bit with bias smaller than ε with probability larger than $1-\delta$. For producing more bits we simply repeat the whole procedure: all the bits produced will have bias smaller than ε conditioned on the bits produced so far, with linear scaling of the resources.

5 Robustness

Aborting the protocol after even a single mistake of the devices is certainly highly impractical from the implementation point of view. Therefore we expand our analysis to a situation where we tolerate certain noise on the devices, which would manifest itself by occasional failing of the test condition even for honest devices. More specifically, we shall tolerate a certain fraction of the devices to malfunction without aborting the protocol.

In more technical version of this work [13] we show, that if we tolerate

$$\frac{(1-f(\varepsilon))}{2} l \tag{8}$$

devices to fail in the whole protocol and want to achieve security parameters ε, δ we can do so by increasing

$$l > \frac{8 \ln \delta}{f(\varepsilon) - 1}\,. \tag{9}$$

This translates into increasing the number of rounds of the protocol comparing to the case of ideal devices by a factor of $\frac{8\ln(f(\varepsilon))}{f(\varepsilon)-1}$. For small ε the parameter $f(\varepsilon)$ approaches 1 and the multiplication factor saturates by 8.

On the other hand we also show that for honest but faulty devices with individual failure probability bounded by

$$\frac{(1-f(\varepsilon))}{4m}\,, \tag{10}$$

the probability of aborting the protocol decreases exponentially with the number of protocol rounds l.

6 Conclusion

In this paper we have introduced a protocol that extracts weak randomness obtained from a min-entropy source in the device independent setting. The protocol works for arbitrarily weak block min-entropy sources with a reasonable scaling of the number of devices. Our protocol is also robust, as it allows tolerating some fraction of malfunctioning devices at the cost of a constant increase of the number of devices used.

Acknowledgements. Authors thank P. Horodecki, A. Winter, and S. Massar for insightful and stimulating discussions and Piotr Mironowicz for supplying the raw data for Fig. 1. JB, MP2 and MP3 acknowledge the support of the Czech Science Foundation GA CR project P202/12/1142 and support of the EU FP7 under grant agreement no 323970 (RAQUEL). MP3 acknowledges VEGA 2/0072/12. JB acknowledges support by the European Research Council through Advanced Grant "IRQUAT". MP1 acknowledges FNP TEAM, NCN grant 2013/08/M/ST2/00626 and ERC QOLAPS.

References

1 Jan Bouda, Matej Pivoluska, Martin Plesch, and Colin Wilmott. Weak randomness seriously limits the security of quantum key distribution. *Phys. Rev. A*, 86:062308, 2012.

2 Marcus Huber and Marcin Pawłowski. Weak randomness in device-independent quantum key distribution and the advantage of using high-dimensional entanglement. *Phys. Rev. A*, 88:032309, 2013.

3 Ronen Shaltiel. An introduction to randomness extractors. In *Automata, Languages and Programming*, volume 6756 of *Lecture Notes in Computer Science*, pages 21–41. 2011.

4 L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov. Hacking commercial quantum cryptography systems by tailored bright illumination. *Nature Photonics*, 4:686–689, 2010.

5 Miklos Santha and Umesh V. Vazirani. Generating quasi-random sequences from semi-random sources. *Journal of Computer and System Sciences*, 33(1):75 – 87, 1986.

6 F. G. S. L. Brandão, R. Ramanathan, A. Grudka, K. Horodecki, M. Horodecki, and P. Horodecki. Robust Device-Independent Randomness Amplification with Few Devices. 2013.

7 R. Colbeck and R. Renner. Free randomness can be amplified. *Nature Physics*, 8:450–454, 2012.

8 R. Gallego, L. Masanes, G. de la Torre, C. Dhara, L. Aolita, and A. Acín. Full randomness from arbitrarily deterministic events. *Nature Communications*, 4, 2013.

9 P. Mironowicz and M. Pawlowski. Amplification of arbitrarily weak randomness, arXiv: 1301.7722, 2013.

10 K.-M. Chung,Y. Shi, X. Wu, Physical Randomness Extractors, arXiv: 1402.4797 2014.

11 N. David Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65:1838–1840, 1990.

12 M. Navascués, S. Pironio, and A. Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, 10(7):073013, 2008.

13 J. Bouda, M. Pawlowski, M. Pivoluska, and M. Plesch. Device-independent randomness extraction for arbitrarily weak min-entropy source, arXiv: 1402.0974, 2014.

14 Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM J. Comput.*, 17(2):230–261, 1988.