Quantum Cryptography and Security

Quantum Random Number Generation

Paolo Lapo Cerni
paololapo.cerni/at/studenti.unipd.it

University of Padua

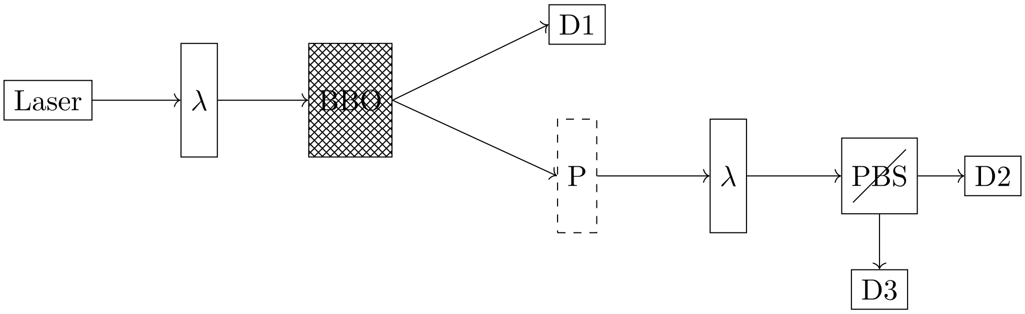
8th November 2023

1 Introduction

In the framework of quantum mechanics, the result of a measure is intrinsically probabilistic. This property of nature itself can be applied to generate the so-called *true random numbers*, namely uniform distributed and uncorrelated. These strings of random bits are nowadays crucial to several security applications.

The implementation of QRNG can be characterized by the degree of trust in the different elements of the protocol. The simplest case is the *trusted setup*, in which all the elements are supposed to be controlled and uncorrelated with the environment. We can relax this hypothesis and consider *semi-Device-Independent setups* with uncharacterized sources or measurements. In these scenarios, we must consider the possible side information of the eavesdropper in order to properly quantify the security of the protocol.

2 Experimental setup



In this experiment, we use the phenomenon known as *spontaneous parametric downconversion* to generate a two-photon entangled state and characterize it in terms of polarization.

- The light source is a 404 nm laser.
- The following half-wave place produces a $|D\rangle$ state.
- A nonlinear crystal (BBO) of Type 1 phase matching can convert a single photon into a pair of photons (SPDC or parametric scattering). If the incident photon is polarized parallel to a specific axis of the crystal, then the resulting photons will both have polarization perpendicular to the initial one (and they will be at 808 nm due to the conservation of energy).

Using two crystals with perpendicular axes and evaluating a specific plane within the two cones of light, we can lose the distinguishability of the two photons and then create an entangled state $|\Psi^+\rangle$.

- The first photon is measured by the first detector and it is used to capture the coincidences. In so doing we are able to limit the errors given by the dark counts of the detectors.
  The detectors are *single-photon avalanche diode*, characterized by a timing jitter of $\sim$ 100ps, $\sim$ 70% efficiency, and a dark count rate of about 500Hz.

- The second photon is the one we will actually characterize. We can collapse its wave function using a polarizer (projective operation) or change the measurement basis with a λ plate (unitary operation).

- We use a polarizer beam splitter and two single-photon detectors to perform the measures.

3 Methods and analysis

3.1 Theoretical framework

The QRNG works as follows [1]: A has access to the state ρ_A which might be correlated with another quantum system E. A obtains the random variable Z by recording the outcome z of a measurement in the $\mathbb{Z}$ basis. Each outcome z is drawn by some probability density function with a given probability P_z.
It is necessary to estimate the amount of side information on Z held by the adversary who has access to the system E. This can be done by sampling the state ρ_A, measuring the system also in the conjugate bases $\mathbb{X}$ and $\mathbb{Y}$.

The state ρ_A represents a single-qubit density matrix that can be uniquely written as [2], [3]:

$$\hat{\rho}_A = \frac{1}{2}\sum_{i=0}^{3} S_i \hat{\sigma}_i$$

where $\hat{\sigma}_i$ are the Pauli's matrices and the S_i are the Stokes parameters whose values are given by $S_i = \mathrm{Tr}\{\hat{\sigma}_i \hat{\rho}_A\}$. Physically, each of these parameters directly corresponds to the outcome of a specific pair of projective measurements:

$$\begin{aligned} S_0 &= P_{|H\rangle} + P_{|V\rangle} = 1 \\ S_1 &= P_{|D\rangle} - P_{|A\rangle} \\ S_2 &= P_{|L\rangle} - P_{|R\rangle} \\ S_3 &= P_{|H\rangle} - P_{|V\rangle} \end{aligned}$$

where $P_{|\psi\rangle}$ is the probability to measure the state $|\psi\rangle$. The measurement of the Stokes parameter can be considered equivalent to the tomography of the density matrix of an ensemble of single qubits.

3.2 Experimental setting

Experimentally, we will consider the frequentist probability. The detection events are accurately described by a Poissonian distribution (rare events), as we can see from the exponential trend in the measured Δt between contiguous detections (figure 1). As expected, most of the events are independent.

Once filtered out the independent events, the coincidences are normally distributed (figures 2, 3, 4) with a standard deviation of a few nanoseconds. Each channel has a specific delay given by differences in the fiber lengths or in the cables that connect the detectors to the time-to-digital converters. Before the analysis, one needs to get rid of the offset and align the time taggers.

Once removed the offset, we selected the coincidences in a time window of 2ns.

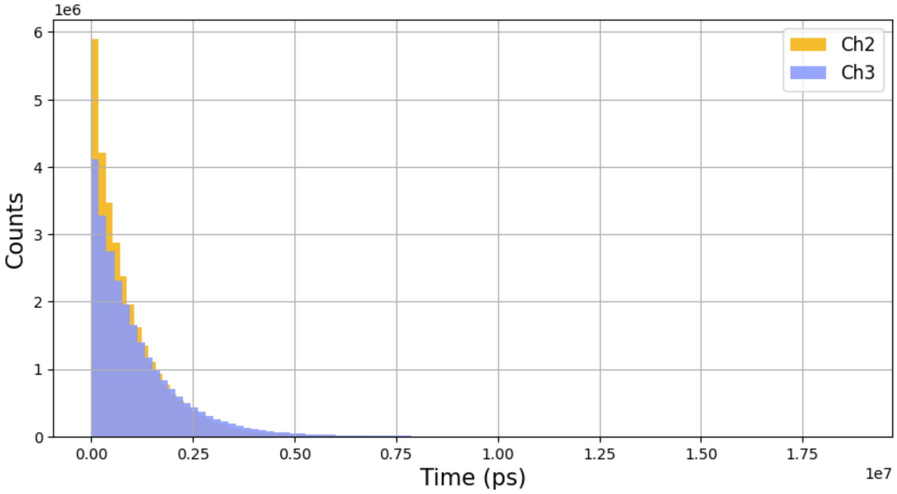


Figure 1. Exponential decay in detections

Note also that, in the more rigorous framework described by James et al. [3], the Stokes parameters are obtained up to a constant $\mathcal{N}$ which depends on the detector efficiency and the light intensity. Such a constant cancels out in the equation of the density matrix (which should be normalized by S_0) under the hypothesis of unbiased detectors.

3.3 Trusted device

If the device is trusted, i.e. we assume that ρ_A is pure, uncorrelated with the environment, we can quantify the number of true random bits that we can obtain via the classical min-entropy:

$$H_{\min}(Z|E) = H_{\min}(Z) = -\log_2(\max_z P_z)$$

3.4 Source-DI based on uncertainty principle

Relaxing the hypothesis of a trusted source, we can provide a bound to the conditional min-entropy $H_{\min}(Z|E)$ based on the entropy uncertainty principle, as described by Vallone et al. [1]:

$$H_{\min}(Z|E) \geq \log_2(\max_{x,z} ||\sqrt{\hat{\mathcal{M}}_z}\sqrt{\hat{\mathcal{N}}_x}||_\infty^2) - H_{\max}(X|B)$$

In so doing, we can bound the amount of information of the attacker using experimental data in the $\mathbb{X}$ basis, without any assumption on the source ρ_A.
In this specific case, we got $A = B$, and $\hat{\mathcal{M}}_z$ and $\hat{\mathcal{N}}_x$ are projective and mutually unbiased over a Hilbert space of dimension $d = 2$. Thus, we can simply this expression to:

$$H_{\min}(Z|E) \geq 1 - H_{\max}(X)$$

where $H_{\max}(X)$ is the Rényi entropy of order 1/2 for the discrete random variable X.

3.5 Tomographic method

As described by Fiorentino et al. [4], we can obtain the min-entropy of a system described by an arbitrary 2×2 density matrix $\hat{\rho}$ estimating the Stokes parameters. This approach upper bounds the amount of the information of the eavesdropper by taking the minimum value of the min-entropy over all the possible decomposition of ρ_A.
One can prove that such an approach leads to:

$$H_{\min}(\rho_A) = -\log_2\left(\frac{1+\sqrt{1-|S_1 - iS_2|^2}}{2}\right)$$

This formula requires the full tomography of the state, i.e. measurements both on $\mathbb{X}$ and on $\mathbb{Y}$. To deal with a partial tomography of ρ_A, we can assume $S_2 = 0$.

3.6 Security parameter

The QRNG procedure generates a string of raw data (size n) with a given min-entropy $H_{\min}(Z|E)$. We now want to *extract* the uniform randomness from that string and to estimate the security parameter ε. To address this problem, we can use Toeplitz matrices of dimension $n \times l$ for universal hashing function construction and implement the extractor [5]. These matrices are uniquely determined by the first row and the first column, so the total number of random bits required to construct them is $l + n - 1$.
The Leftover Hashing Lemma (LHL, [6]) tells us the relation between the information of the eavesdropper and the security of our protocol. In this case, we can bound the variational distance from the ideal counterpart of the protocol by the security parameter:

$$\varepsilon = \frac{1}{2}\sqrt{2^{l-H_{\min}(Z|E)}}$$

4 Results

4.1 Trusted setup

In the trusted scenario, we consider only measurements on the $\mathbb{Z}$ basis. It is insightful to compare a pure state and a mixed state to understand the differences in the results.
For what concern the **pure state** we analyzed $\sim 100M$ events and we found $\sim 2M$ coincidences, leading to:

$$P^{(p)}_{|H\rangle} = 0.488 \qquad P^{(p)}_{|V\rangle} = 0.512$$

On the other hand, the **mixed state** has been characterized by $\sim 2M$ events and $\sim 1M$ coincidences:

$$P^{(m)}_{|H\rangle} = 0.505 \qquad P^{(m)}_{|V\rangle} = 0.495$$

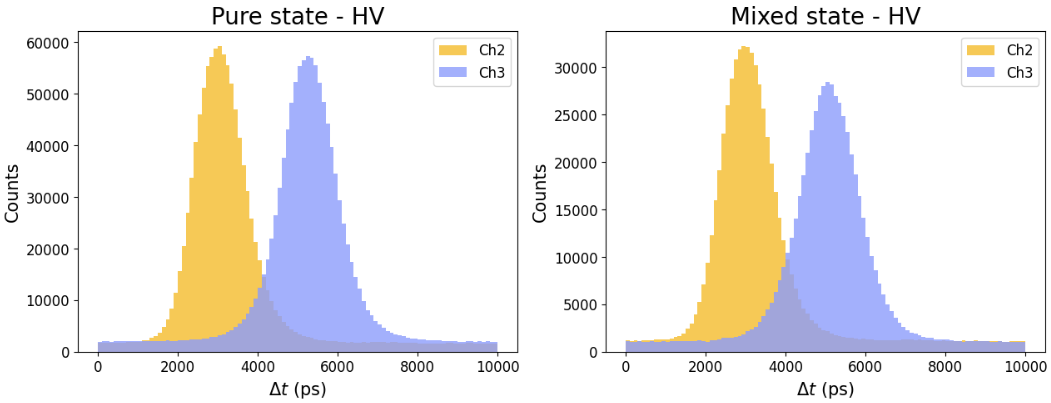


Figure 2. Coincidences in the $\mathbb{Z}$ basis, pure and mixed states.

The estimated min-entropy is:

$$H^{(p)}_{\min}(Z) = 0.966 \qquad H^{(m)}_{\min}(Z) = 0.987$$

With this setup, it's impossible to distinguish the two states. This is proof of the strength of the assumptions made on the system. In particular, the mixed system could be entangled with the eavesdropper's environment, so definitely not secure, and we have no method to evaluate the privacy of our random string.
We can also notice that the pure state has a lower coincidences/events ratio. This is reasonable because it's obtained with an additional optical element which increases the losses.

4.2 Partial tomography

We can increase the security of our protocol by measuring also on the $\mathbb{X}$ basis. In so doing, we can distinguish the pure state from the mixed one and better estimate the conditional min-entropy, using two different approaches, as described above (3).
The analysis of the **pure state** investigated $\sim 40M$ events and $\sim 800K$ total coincidences:

$$P^{(p)}_{|D\rangle} = 0.008 \qquad P^{(p)}_{|A\rangle} = 0.992$$

While, for the **mixed state**, we used $\sim 30M$ events and $\sim 1M$ coincidences:

$$P^{(m)}_{|D\rangle} = 0.509 \qquad P^{(m)}_{|A\rangle} = 0.491$$

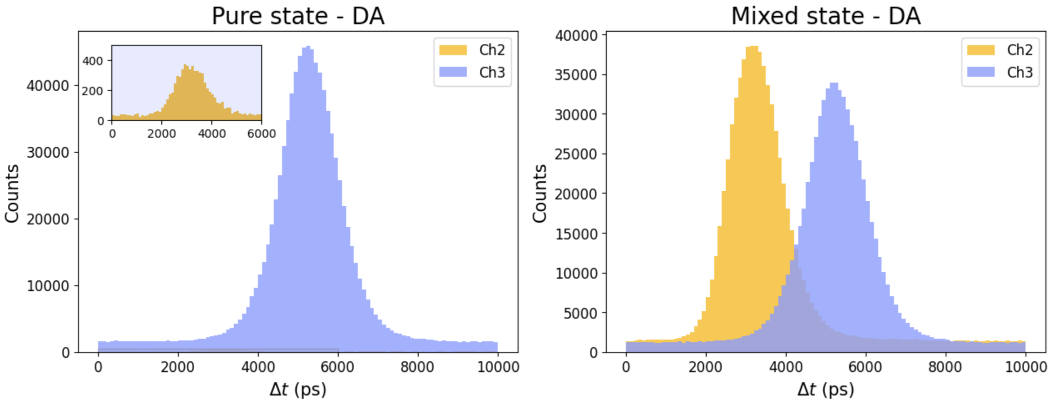


Figure 3. Coincidences in the $\mathbb{X}$ basis, pure and mixed states.

We can bound the conditional min-entropy either using the uncertainty principle or the tomographic method (setting $S_2 = 0$). The two approaches lead to a similar result up to the experimental uncertainties.

$$H_{\min}^{(p)}(Z|E) = 0.767 \qquad H_{\min}^{(m)}(Z|E) = 1 \cdot 10^{-4}$$

In this case we can distinguish the mixed state from the pure one and evaluate the degrees of freedom of the eavesdropper on the system.

4.3 Full tomography

In the third scenario, we measure our state on all the bases, using $\mathbb{Z}$ to produce the random string and both $\mathbb{X}$ and $\mathbb{Y}$ to establish the privacy of the random numbers.
Briefly, the analysis on the $\mathbb{Z}$ basis ($\sim 28M$ events, $\sim 600K$ coincidences) leads to:

$$P^{(l)}_{|H\rangle} = 0.497 \qquad P^{(l)}_{|V\rangle} = 0.503$$

The one on the $\mathbb{X}$ basis ($\sim 25M$ events, $500K$ coincidences) estimates:

$$P^{(l)}_{|D\rangle} = 0.497 \qquad P^{(l)}_{|A\rangle} = 0.503$$

while, for what concern the $\mathbb{Y}$ basis ($\sim 30M$ events, $550K$ coincidences):

$$P^{(l)}_{|L\rangle} = 0.988 \qquad P^{(l)}_{|R\rangle} = 0.012$$

As we can see, the state is $\rho = |L\rangle\langle L|$ up to small corrections.

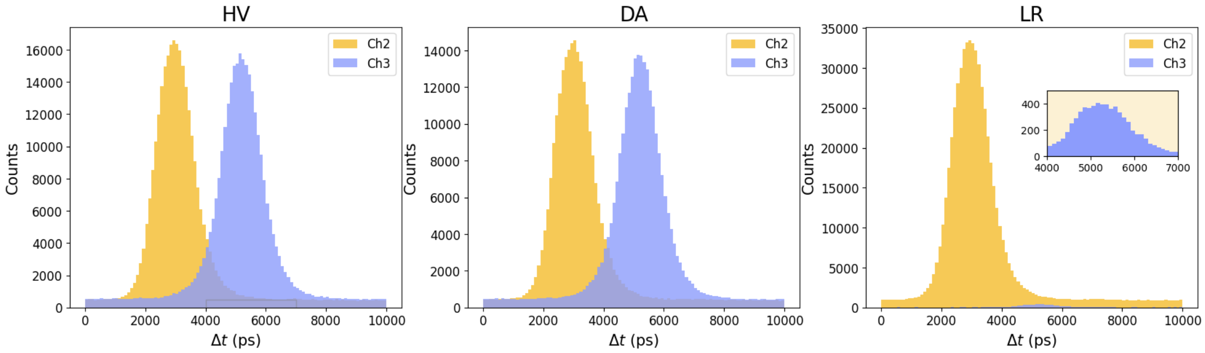


Figure 4. Complete tomography of the $|L\rangle$ state.

To estimate the information held by the adversary, we can again use the two approaches.
In this case, we have the full-state tomography so we can estimate also the second Stokes parameter S_2. For what concerns the method based on the uncertainty principle, in principle we could choose what basis to use between $\mathbb{X}$ and $\mathbb{Y}$. In practice, it's useless to use the first one because it leads to a weak bound since $P_{|D\rangle} \approx P_{|A\rangle} \approx 1/2$. Again, the two methods return the same outcome up to the experimental uncertainties.
Thus, we can estimate the conditional min-entropy as:

$$H_{\min}^{(p)}(Z|E) = 0.712$$

4.4 Density matrices

The tomographic method allows us to reconstruct the density matrices of the different states. In this subsection, we plot the results of such reconstruction.

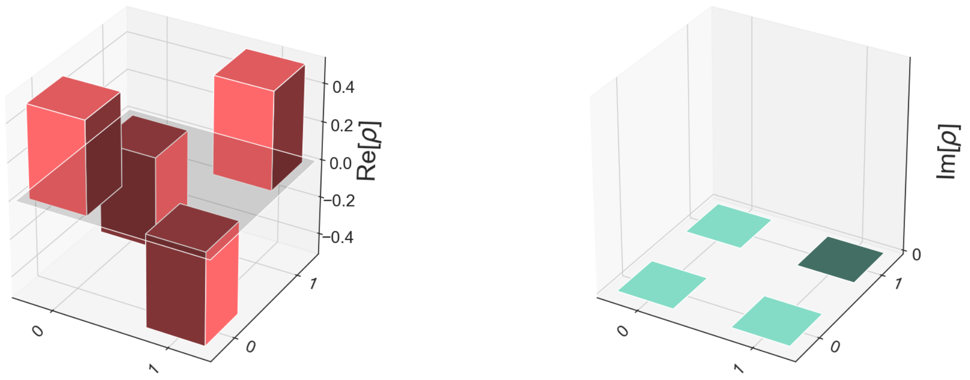


Figure 5. Partial tomography of the $|D\rangle$ state ($S_2 = 0$).

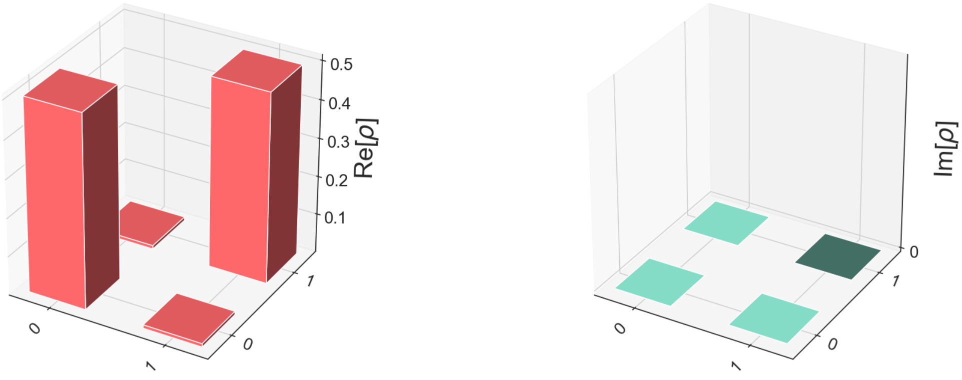


Figure 6. Partial tomography of the mixed state ($S_2 = 0$).

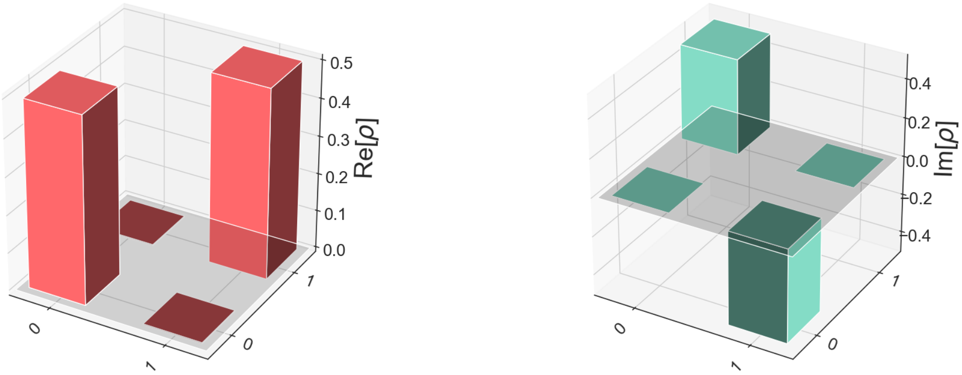


Figure 7. Complete tomography of the $|L\rangle$ state.

4.5 Security parameter

As briefly summarized above (3.6), we can extract the randomness from the raw string of n bits using Toeplitz matrixes of dimension $n \times l$. The security parameter ε scales as:

$$\varepsilon = \frac{1}{2}\sqrt{2^{l-H_{\min}(Z|E)}}$$

Here we plot ε as a function of the length of the extracted bits l, at fixed min-entropy.

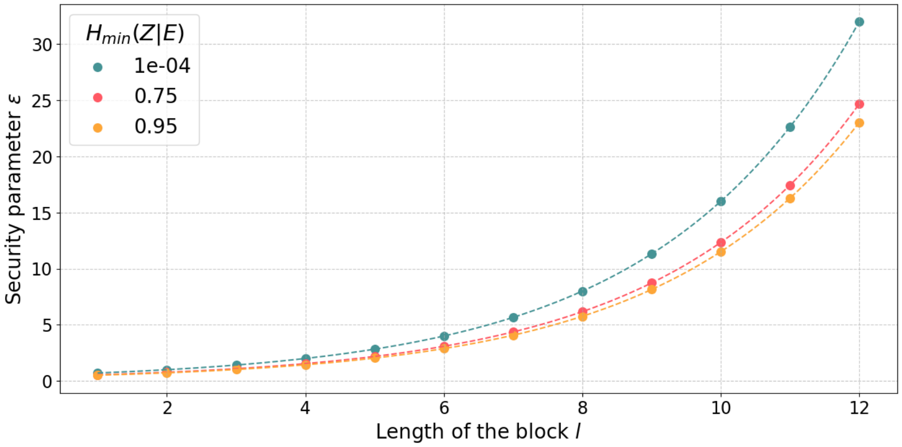


Figure 8. Scaling of the security parameter $\varepsilon(l)$.

5 Conclusions

In this report, we briefly introduced the problem of the implementation of a quantum random number generator. Using different setups, characterized by different degrees of trust in the sources, we established some bounds to the eavesdropper's side information and we quantified the security of the protocol.

References

[1] Giuseppe Vallone, Davide G Marangon, Marco Tomasin, and Paolo Villoresi. Quantum randomness certified by the uncertainty principle. *Physical Review A*, 90(5):052327, 2014.

[2] Joseph B Altepeter, Evan R Jeffrey, and Paul G Kwiat. Photonic state tomography. *Advances in atomic, molecular, and optical physics*, 52:105–159, 2005.

[3] Daniel FV James, Paul G Kwiat, William J Munro, and Andrew G White. Measurement of qubits. *Physical Review A*, 64(5):052312, 2001.

[4] M Fiorentino, C Santori, SM Spillane, RG Beausoleil, and WJ Munro. Secure self-calibrating quantum random-bit generator. *Physical Review A*, 75(3):032334, 2007.

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

[6] Marco Tomamichel, Christian Schaffner, Adam Smith, and Renato Renner. Leftover hashing against quantum side information. *IEEE Transactions on Information Theory*, 57(8):5524–5535, 2011.