Experimental private quantum randomness generation invulnerable to the detection loophole

Gustavo Cañas,[1,2,3] Jaime Cariñe,[2,4] Esteban S. Gómez,[1,2,3] Johanna F. Barra,[1,2,3]
Adán Cabello,[5] Guilherme B. Xavier,[2,3,4] Gustavo Lima,[1,2,3] and Marcin Pawłowski[6]

[1]*Departamento de Física, Universidad de Concepción, 160-C Concepción, Chile*
[2]*Center for Optics and Photonics, Universidad de Concepción, 160-C Concepción, Chile*
[3]*MSI-Nucleus for Advanced Optics, Universidad de Concepción, 160-C Concepción, Chile*
[4]*Departamento de Ingeniería Eléctrica, Universidad de Concepción, 160-C Concepción, Chile*
[5]*Departamento de Física Aplicada II, Universidad de Sevilla, E-41012 Sevilla, Spain*
[6]*Instytut Fizyki Teoretycznej i Astrofizyki, Uniwersytet Gdański, PL-80-952 Gdańsk, Poland*

The generation of random numbers is an important task in many fields. Unfortunately, when such numbers are created from classical processes, they may be predictable and therefore nonprivate. One solution is to use random number generators exploiting the intrinsic uncertainty of quantum phenomena. Nevertheless, imperfections in these devices can leave undetected patterns that compromise the privacy of the random string. Recently, device independent quantum randomness generation was shown to be possible, however requiring not yet practical schemes. Here we introduce a protocol for generating private randomness in a prepare-and-measure semi-device independent scenario, and demonstrate its practicability. Remarkably, the privacy in our protocol cannot be compromised by cheating strategies that exploit detectors efficiencies.

Introduction—Private random numbers strings are essential for multiple applications, including, but not limited to, cryptography and digital rights management. However, random numbers produced from classical processes may be predictable and therefore nonprivate. One solution is to use quantum random number generators (QRNGs) based on the intrinsic uncertainty of quantum measurement outcomes [1–5]. Unfortunately, imperfections in their components can leave correlations between some random bits of the generated string. These correlations may be undetected by standard randomness tests [6], and then be exploited by an adversary [7]. Device-independent (DI) QRNGs can certify private randomness generation [7], as it does not rely on a detailed model of the components comprising the device, nor in a knowledge of the employed quantum systems. However, DI-QRNGs are not yet practical as they require a detection efficiency that, so far, has only been achieved with trapped ions [8] and with photons detected with transition-edge superconducting sensors [9, 10].

Here we introduce a protocol for the quantum generation of private random bit sequences. It is based on the semi-device independent (SDI) approach where, again, no detailed description on the internal components of the QRNG is needed, while the dimension of the employed quantum system is assumed to be upper-bounded [11]. Our protocol works in a prepare-and-measure scenario and, unlike device-independent QRNGs, it is able to certify private randomness even with a very low detection efficiency. The central idea is to introduce a blocker that randomly adds its own failures in the communication between the QRNG preparation and measurement stages. This counter-measure uncorrelates any pre-existent shared randomness between both stages, and foils detector-loophole cheating strategies. Using single-photon qubit states, we demonstrate the practicability of this protocol. We obtain a high certified private random bit generation rate of 0.28 Hz, while adopting standard avalanche photo-detectors (APD) providing an overall detection efficiency of only 6 %.

Protocol description.—Figure 1 shows schematically the scenario of our protocol. In each experimental round, the preparation stage P uses an input $x \in \{00,01,10,11\}$ and produces a qubit state ρ_x that goes into a measurement apparatus M. M uses an input $z \in \{0,1\}$ and produces an outcome $b \in \{0,1,\emptyset\}$. In the middle, there is a blocker B that blocks ρ_x depending on input $y \in [0,1]$. It is blocked if and only if $y \leq \lambda$, where λ is a setting parameter of B. The outcome $b = \emptyset$ corresponds to the case in which no result is obtained either because the blocker B was acting or due to imperfect detection efficiency. In the protocol we consider that: (i) P and M are black boxes built by the adversary and their detailed internal functioning is unknown to the user who employs them for private randomness generation. Thus, these boxes may even contain an agent of the adversary. (ii) P and M are shielded, that is, P only receives x and only outputs ρ_x, while M only receives as inputs z and ρ_x (or nothing if the system is lost) and outputs b. (iii) x, y, z are independently produced numbers that pass standard tests of randomness [6], but are not proved to be private. Thus, they may be produced by imperfect QRNGs. (iv) The laboratory containing P, B, and M is shielded: no signals can enter or exit. Under these assumptions, the protocol generates private random numbers from untrusted devices.

Each round produces an event $(b|x,y,z)$. To generate private random bits the user takes the following steps. Step 1. The user estimates the probabilities $p(b|x,y > \lambda, z)$ and the probabilities $p'(b|x,y > \lambda, z)$. These last ones are obtained after postselecting rounds where $b = \emptyset$. The overall detection efficiency η is estimated from the detection statistics. Conventionally, η is defined as $\sum_{x,z}\sum_{b\in\{0,1\}} p(b|x,y > \lambda, z) / \sum_{x,z}\sum_{b\in\{0,1,\emptyset\}} p(b|x,y > \lambda, z)$. Step 2. Here, the user checks whether it is possible for stages P and M to have correlated shared variables (shared randomness). That is, whether both generators are fed with the same seed, which would compromise the privacy of the final string. For this, he estimates the average probability that M correctly guesses the value of one of the two input bits of P. This probability is $p'_{\text{av}} = \frac{1}{8}\sum_{x,z} p'(b = x_z|x,y > \lambda, z)$, where x_z is the first or sec-

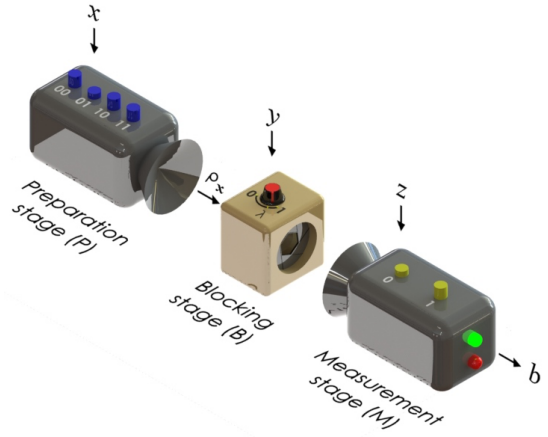

FIG. 1: (Color online) Prepare-and-measure scenario of our SDI protocol for private randomness generation. See text for details.

ond bit of each input x to be guessed. If p'_{av} is above a certain threshold that depends on λ and η, then the user can conclude that there is no shared randomness and the random sequence generated is private (see details below). If it is beneath this threshold, the user aborts. The private random sequence $\mathscr{S}$ is defined as follows: b if $y > \lambda$ and $b \in \{0,1\}$, 0 if $y > \lambda$ and $b = \emptyset$. Hereafter we will denote these events as $(b|x,z)$. Step 3. The user checks the amount of randomness of the private random sequence obtained $\mathscr{S}$. This should be done considering as the randomness indicator a list $\vec{P}$ of the 8 probabilities $p(0|x,z)$ measured in the experiment, and that an adversary can still attack the system even though shared randomness was discarded in step 2 (see details below).

Blocking to counteract shared randomness.—An adversary who built P and M may learn the generated sequence by exploiting shared randomness between these stages. For this he will try to feed their corresponding random generators with the same seed. Since the generators are inside a shielded laboratory, the seed can not be sent to them from outside. This leaves only two options for the adversary: (i) the seed can be stored inside the generators when they were built, or (ii) a seed may be communicated from the adversary's agent in P to the adversary's agent in M through the qubit sent in the protocol. The user can easily check option (i) by employing several P and M stages paired randomly. Since the adversary cannot know in advance how the devices will be paired, they must all have the same seed. Therefore correlations between inputs and outputs will be observed. Note that this user's strategy also verifies the possibility that the adversary's stages are using the current time as the seed. The blocking stage B is employed to counteract any shared randomness between P and M, if strategy (ii) is followed by the adversary. The random blocking of some rounds forces the adversary's agents to use more rounds to communicate the seed and demands synchronisation. Synchronisation is required because the agents must agree on which round of the protocol they are. In this process, P sends a message about the experiment's current round to M using the qubit states. Here we consider this message to be only 1 bit long. This favours the adversary because in practice more information is needed for synchronisation. If the probability of the signal being blocked is R and the adversary wants to have confidence c that M receives the message, then $n_1 = \log_R(1-c)$ rounds of the protocol have to be used for synchronisation. When synchronisation is achieved, the adversary's agents can use shared randomness but only for a limited time. They use the seed to generate a random number for each round, but as soon as another qubit is blocked, P will count more rounds than M and synchronisation will again be required. The average number of rounds before this happens is $n_2 = \frac{1-R}{R}$.

In the round used for synchronisation, the success probability is $p'_{\text{av}} = \frac{1}{2}$ because there is no correlation between the output b of M and the input x of P. Therefore, in principle, by monitoring the value of p'_{av} the user can check whether there is shared randomness or not. This is possible because the measured value of p'_{av} will become smaller than the 85% predicted by quantum mechanics. Nevertheless, there is a strategy that the agents can adopt to cheat the user: In the rounds where there is shared randomness, they can fake $p'_{\text{av}} = 1$ to compensate the reduction that will be observed during the synchronisation. Moreover, they can exploit the inefficiency of the detectors, since to estimate p'_{av} the user considers only the rounds when the detector registered a particle. The detectors can be controlled by them and always register the qubits in the rounds where $p'_{\text{av}} = 1$, and only "fire" in some rounds where there is synchronisation (i.e., when $p'_{\text{av}} = \frac{1}{2}$). If the total observed detection efficiency is η, the probability of firing during a round with $p'_{\text{av}} = \frac{1}{2}$ is $\gamma = \max\{0, \frac{\eta(n_1+n_2)-n_2}{n_1}\}$. The observed success probability is then

$$p'_{\text{av}} = \frac{\frac{1}{2}\gamma n_1 + n_2}{\gamma n_1 + n_2}. \tag{1}$$

Thus, the user can conclude that there is no shared randomness if the measured p'_{av} is larger than (1). Note that p'_{av} is a function of R, c and the observed efficiency η. In Fig. 2 we plot this dependence of p'_{av} with varying blocking probability R and η, while considering $c = 0.99$. The black dot is the observed p'_{av} in our implementation of the protocol. Remarkably, Fig. 2 shows that, for $\eta > 0$, there is always a blocking probability smaller than 1 such that observing $p'_{\text{av}} > 0.5$ allows the user to counteract the use of shared randomness.

Computing the amount of randomness.—Assuming no shared randomness, the most general attack to our SDI protocol consists of performing von Neumann measurements at the M stage. Let us consider two arbitrary measurement directions corresponding to the case when the input is $z = 0$ and $z = 1$. For each x, the adversary's agent in P prepares a pure qubit state $|x\rangle$ in the equatorial plane of the Bloch sphere. For a given z, the adversary's agent in M: (I) with frequency p_z, performs a measurement that projects the qubit onto the eigenvectors $|m_z = 0\rangle$ or $|m_z = 1\rangle$ and outputs $b = 0$ or $b = 1$, respectively. The eigenvectors lie on the equatorial plane of the

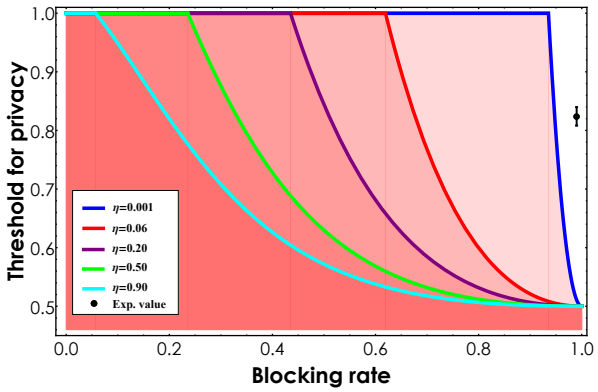

FIG. 2: (Color online) Threshold value for p'_{av} that guarantees no shared randomness, as a function of R and for different values of η. The plots were done considering a 1 bit-long seed and $c = 0.99$. The black dot indicates the p'_{av} observed in our experiment in which $R = 0.99$ and $\eta = 0.06$. This blocking probability was chosen to show that in the protocol shared randomness can be discarded even when $\eta = 0.001$.

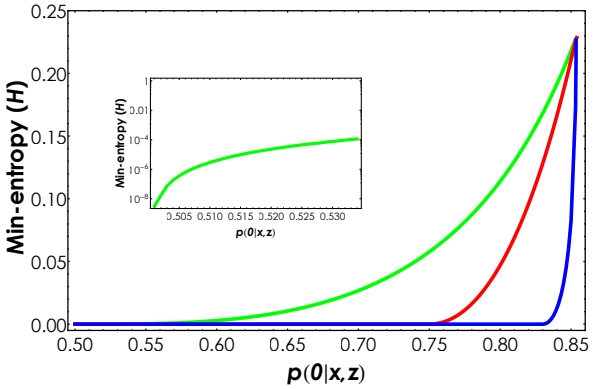

FIG. 3: (Color online) Certified randomness, measured by the min-entropy, using $\vec{P}$ (green curve), p_{wc} (red curve), and p_{av} (blue curve) [17]. It is assumed a precision such that $p(0|x,z)^{\text{min}} = \alpha - \frac{\delta}{2}$ and $p(0|x,z)^{\text{max}} = \alpha + \frac{\delta}{2}$, with $\delta = 10^{-4}$. The threshold probability for certifying randomness when using $\vec{P}$ is 0.5, while it is 0.75 for p_{wc}, and 0.829 for p_{av}. Small diagram: Randomness certified using $\vec{P}$ for small α and $\delta = 10^{-8}$. p_{wc} and p_{av} cannot certify randomness in this case.

Bloch sphere. (II) With frequency q_z^0, he makes no measurement on the qubit and generates a fake output $b = 0$. (III) With frequency q_z^1, he again does not perform a measurement and outputs $b = 1$. Since in those cases in which no measurement is performed, the choice of output $b = 0$ or $b = 1$ can be done arbitrarily, we will assume that there is no randomness with frequency $q_z^0 + q_z^1$. By the normalisation condition, one has that $p_z + q_z^0 + q_z^1 = 1$. Thus, the amount of randomness in the private sequence $\mathscr{S}$ of the events $(b|x,z)$, which is measured by the min-entropy, is given by

$$H_\infty(b|x,z) = -p_z \log_2 |\langle m_z = b|x\rangle|^2, \tag{2}$$

where $|\langle m_z = b|x\rangle|^2$ is the probability of projecting into $|m_z = b\rangle$ when the state prepared is $|x\rangle$. Since P and M are black boxes, the user does not know p_z, $|x\rangle$ and $|m_z = b\rangle$. The user only has access to the probabilities $p(b|x,z)$ that can be estimated from the experiment. The relation between them is $p(b|x,z) = q_z^b + p_z |\langle m_z = b|x\rangle|^2$. The randomness indicator is a list $\vec{P}$ of the 8 probabilities $p(0|x,z)$. After the experiment is complete, the user has each $p(0|x,z)$ with a confidence $p(0|x,z)^{\text{min}} \leq p(0|x,z) \leq p(0|x,z)^{\text{max}}$. Then, a numerical minimization of the min-entropies of the 8 events $(0|x,z)$, under the constraint above, is performed to obtain the amount of randomness of $\mathscr{S}$. One interesting feature is that $\vec{P}$ allows the user to certify randomness more efficiently than with any indicator used in previous works, e.g., the average success probability $p_{\text{av}} = \frac{1}{8}\sum_{x,z} p(0|x,z)$ [11, 12], or the worst case probability $p_{\text{wc}} = \min_{x,z} p(0|x,z)$ [13, 14]. The benefit of using $\vec{P}$ over p_{wc} or p_{av} is illustrated in Fig. 3. This contrasts with the marginal benefit when certifying randomness in Bell-inequality scenarios using the whole probability distribution [15, 16].

Experiment.—Our implementation of the protocol is shown schematically in Fig. 4 (a). As the sources of x, y, and z we use three commercial QRNGs QUANTIS [18]. They have passed standard tests of randomness [18], but no assumptions are made by us regarding their privacy. As previously mentioned, imperfections in these QRNGs can compromise the string's privacy, even though the device may have passed standard tests of randomness [6]. A field programmable gate array (FPGA) in P produces an electrical synchronisation signal, which also drives an acousto-optical modulator (AOM) producing attenuated optical pulses from a continuous laser. These pulses are then sent through a sequence of four spatial light modulators (SLMs) [19]. Sets of lenses are employed to project the image of one SLM onto the next one. We use the linear transverse momentum of the single photons transmitted by the SLMs as the degree of freedom for codifying qubit states [20]. This is done by projecting masks with only two paths available for the photon transmission in the liquid crystal displays of the SLMs [21]. The qubit state preparation in P (the projections in M) is implemented using SLM 1 and SLM 2 [SLM 3, SLM 4 (and an APD)] working with amplitude-only and phase-only modulation, respectively [22**?**]. The real and imaginary parts of the generated and measured states are set by adjusting the grey level of the pixels on the SLMs. In our demonstration we set these states to maximise the success probability p'_{av}, which is the figure of merit of the protocol. Nevertheless, in accordance with the SDI scenario considered in the protocol, this last step is not required to certify the generation of private random bits. Whenever the observed p'_{av} is higher than the discussed threshold value (see Fig. 2), the generation of privacy randomness is certified.

The repetition rate of the attenuated optical pulses is set to 30 Hz, which is the limit of the employed SLMs. The applied modulation in each SLM is triggered by the sync signal. An internal delay in respect to the AOM in the FPGAs

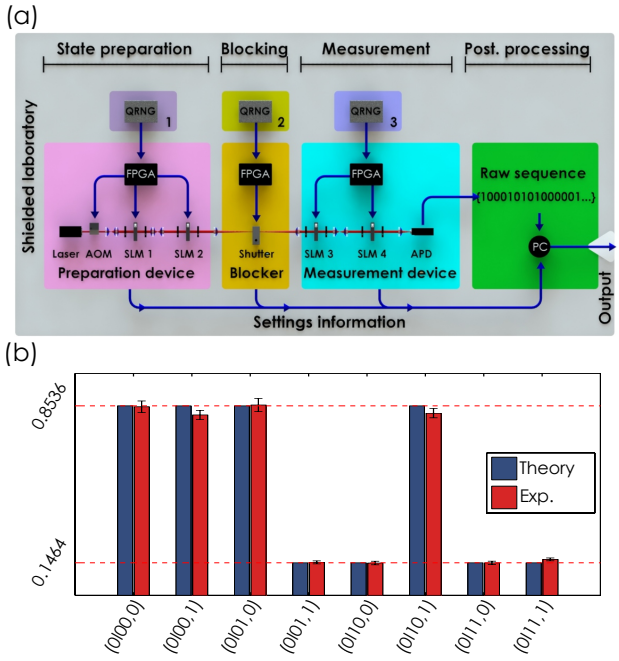

FIG. 4: (Color online) (a). Experimental setup (see details in the main text). (b) Predicted and the observed experimental probabilities for the events $(0|x,z)$.

is used to ensure that the SLMs in P and M are properly set by the time each pulse is sent. In each round, pre-determined modulations are applied to the SLMs by their corresponding FPGAs based on the numbers produced by each QRNG. The optical blocker placed between them is a commercial shutter and is controlled by a third FPGA unit, fed by another QRNG. The blocker's electronics also receive the clock signal from P. For each round of the experiment, the blocker's FPGA unit randomly blocks the pulse, with an adjustable probability. In our implementation, with $\eta = 0.06$, a blocking probability of 99% was employed. The recorded probabilities are in very good agreement with the predictions of quantum mechanics, as illustrated in Fig. 4 (b). The corresponding experimental p'_{av} is shown in Fig. 2, and one can clearly see that the random numbers generated are certified to be private. In each round of the experiment 0.0093 true private random bits are certified, yielding a private random bit generation rate of 0.28 Hz. The final bit string $\mathscr{S}$ generated in our experiment is 10^5 bits long. We emphasize that our rate could be increased to the Mbit/s range with current technology of SLMs based on integrated silicon photonics [23].

Conclusions.—In this work we have introduced a SDI protocol for generating certified private random numbers. Its relevance lies in the following features: (I) Unlike DI protocols, our protocol works with inneficient detectors, which is required for real-world applications. (II) Still, like DI protocols, no detailed knowledge of the internal functioning of the QRNG is needed; the only assumption is that the dimension of the transmitted quantum system is upper bounded. (III) The protocol works even if the adversary has introduced shared randomness between the preparation and measurement devices. We demonstrate the protocol with single-photon qubit states and obtain a high random bit rate generation, which is certifiably private, of 0.28 Hz with a detection efficiency of only 6 %. Our results pave the way towards a new generation of QRNGs and SDI protocols.

[1] J. G. Rarity, P. C. M. Owens, and P. R. Tapster, J. Mod. Opt. **41** 2435 (1994).
[2] A. Stefanov, N. Gisin, O. Guinnard, L. Guinnard, and H. Zbinden, J. Mod. Opt. **47**, 595 (2000).
[3] A. Uchida, K. Amano, M. Inoue, K. Hirano, S. Naito, H. Someya, I. Oowada, T. Kurashige, M. Shiki, S. Yoshimori, K. Yoshimura, and P. Davis, Nature Photon. **2**, 728 (2008).
[4] C. Gabriel, C. Wittmann, D. Sych, R. Dong, W. Mauerer, U. Andersen, C. Marquardt, and G. Leuchs, Nature Photon. **4**, 711 (2010).
[5] M. Ren, E. Wu, Y. Liang, Y. Jian, W. Guang, and H. Zeng, Phys. Rev. A **83**, 023820 (2011).
[6] http://csrc.nist.gov/groups/ST/toolkit/rng/index.html
[7] R. Colbeck, *Quantum and Relativistic Protocols for Secure Multi-Party Computation*, Ph.D. Thesis, Cambridge University (2007) arXiv:0911.3814.
[8] S. Pironio, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Nature (London) **464**, 1021 (2010).
[9] M. Giustina, A. Mech, S. Ramelow, B. Wittmann, J. Kofler, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S. W. Nam, R. Ursin, and A. Zeilinger, Nature **497**, 227 (2013).
[10] B. G. Christensen, K. T. McCusker, J. B. Altepeter, B. Calkins, T. Gerrits, A. E. Lita, A. Miller, L. K. Shalm, Y. Zhang, S. W. Nam, N. Brunner, C. C. W. Lim, N. Gisin, and P. G. Kwiat, Phys. Rev. Lett. **111**, 130406 (2013).
[11] M. Pawłowski and N. Brunner, Phys. Rev. A **84**, 010302 (2011).
[12] H. W. Li, Z. Q. Yin, Y. C. Wu, X. B. Zou, S. Wang, W. Chen, G. C. Guo, and Z. F. Han, Phys. Rev. A **84**, 034301 (2011).
[13] M. Dall'Arno, E. Passaro, R. Gallego, M. Pawłowski, and A. Acín, Quant. Inf. Comp. **15**, 0037 (2015).
[14] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, J. ACM **49**, 496 (2002).
[15] O. Nieto-Silleras, S. Pironio, and J. Silman, New J. Phys. **16**, 013035 (2014).
[16] J. D. Bancal, L. Sheridan, and V. Scarani, New J. Phys. **16**, 033011 (2014).
[17] P. Mironowicz, H. W. Li, and M. Pawłowski, Phys. Rev. A **90**, 022322 (2014).
[18] http://www.idquantique.com/random-number-generators/products.html
[19] D. G. Grier, Nature **424**, 810 (2003).
[20] G. Lima, A. Vargas, L. Neves, R. Guzmán, and C. Saavedra, Opt. Express **17**, 10688 (2009).
[21] L. Neves, G. Lima, J. G. Aguirre Gómez, C. H. Monken, C. Saavedra, and S. Pádua, Phys. Rev. Lett. **94**, 100501 (2005).
[22] G. Lima, L. Neves, R. Guzmán, E. S. Gómez, W. A. T. Nogueira, A. Delgado, A. Vargas, and C. Saavedra, Opt. Express **19**, 3542 (2011).
[23] C. Qiu, J. Chen, Y. Xia, and Q. Xu, Sci. Rep. **2**, 855 (2012).