Certified Quantum Random-Number Generator Based on Single-Photon Entanglement

Nicolò Leone,[1,*] Stefano Azzini,[1] Sonia Mazzucchi,[2] Valter Moretti,[2] and Lorenzo Pavesi[1]

[1] *Nanoscience Laboratory, Department of Physics, University of Trento, Italy*
[2] *Department of Mathematics and TIFPA-INFN, University of Trento, Italy*

(Received 20 September 2021; revised 21 November 2021; accepted 20 January 2022; published 3 March 2022)

Quantum entanglement represents an ideal resource to guarantee the security of random numbers employed in many scientific and cryptographic applications. However, entanglement-based certified random number generators are particularly challenging to implement. Here, we demonstrate a certified quantum random number generator based on momentum-polarization entangled single-photon states. The use of single-photon entanglement allows an attenuated laser source to be employed and a simple setup where only linear optical components are utilized. For the latter, a semi-device-independent modeling of the photonic quantum random number generator is developed, which certifies a minimum entropy of $(2.5 \pm 0.5)\%$, corresponding to a generation rate of 4.4 kHz. At the expenses of a higher level of trust in the system, the certified minimum entropy can be increased to $(30.1 \pm 0.5)\%$, implying a generation rate of 52.7 kHz. Our results show that a simple optical implementation combined with an accurate modeling provide an entanglement-based high-security quantum random number generator using imperfect devices.

DOI: 10.1103/PhysRevApplied.17.034011

I. INTRODUCTION

Random numbers represent a fundamental resource in several applications, in particular numerical simulations, internet of things, and cryptography [1]. In the latter, the level of unpredictability of the sequence of random bits is a fundamental aspect, because it guarantees the security of the cryptographic protocols. For these reasons, randomness certification has become a crucial feature, i.e., being able to ensure that the random numbers are uniformly distributed, uncorrelated, and unpredictable. The validation of the first two features can be done by running suitable statistical tests on the numbers sequence, whereas ensuring the unpredictability represents a more challenging task. In cryptography, the main figure of merit for the randomness as well as for the security of the output string is the *min-entropy* [2]. Its certification for a random number generator (RNG) allows the use of a suitable randomness extractor to obtain from the raw bits a sequence of uniform random bits [3]. RNGs can be classified into three main categories from the security point of view. Pseudo-random number generators (PRNGs) are essentially based on algorithms, hence their security is based on assumptions on the computational power of an adversary. True random number generators (TRNGs) are based on physical phenomena which are hard to predict. Even if TRNGs are, in principle, more secure than PRNGs, it is rather difficult to provide a robust model of their entropy source because they are based on noisy and/or chaotic phenomena. In particular, fluctuations of the working conditions could lower significantly the entropy of the bit string. Finally, RNGs based on quantum physics, where the entropy source is a quantum process, are called quantum random number generators (QRNGs) [1,4]. The probabilistic nature of the measurement outcomes of quantum mechanical phenomena gives an important security advantage to QRNGs with respect to TRNGs. Moreover, this natural source of entropy can be modeled, allowing for an easy and robust estimation of the min-entropy, which is independent from the presence of additional classical noise [5]. Among the different types of mechanisms for a QRNG, the *device-independent* QRNG is considered the most secure: the randomness certification of the generated sequence is obtained independently of any modeling of the employed devices, which are considered as black boxes. In principle, device-independent protocols should be able to handle scenarios in which device imperfections are present and adversaries can access the employed devices and program them [6,7]. The main example of device-independent random number generation protocol relies on Bell's inequality violation. Indeed, it is well known that when a Bell's inequality is violated, the quantum correlation between measurements of local observables cannot be reproduced by a local hidden variable theory. In other words, the results of a quantum

[*] Corresponding author. nicolo.leone@unitn.it

measurement cannot be predetermined [8], hence they are intrinsically random [7] and unpredictable. Remarkably, it is possible to quantify the min-entropy associated with the amount of violation of a Clauser-Horne-Shimony-Holt (CHSH) inequality [9], thus providing an estimate of the amount of *true quantum randomness* the device is able to provide [10–12]. However, it is important to point out that the requirement of device independence is rather demanding from an experimental point of view, requiring also loophole-free Bell tests. A few proof-of-principle experiments with entangled photon pairs have been carried out in this direction[13–17]. In particular, Refs. [16,17] have recently reported remarkable improvements in the random bit throughput, but this kind of experiment remains a technological challenge. Indeed, they typically require complicated setups having spacelike separated detection stages equipped with expensive high-efficiency detectors. This makes the deployment of device-independent QRNG still prohibitive at the moment.

Through the years, another class of QRNG device has been largely developed, allowing the above-mentioned technological hurdles to be overcome at the expense of a few assumptions on the experimental setup: these are the *semi-device-independent* QRNGs. In these solutions, the min-entropy is guaranteed by the fundamental principles of quantum mechanics, but they gain in ease of implementation by introducing additional assumptions on the theoretical modeling of the QRNG. This relaxes the requirements on the employed devices leading to less-complicated implementations. As an example, many of them assume that one or more components of the setup are trusted, e.g., the source [18–21] or the measurement apparatus [22–24]. Others exploit energy-constrained quantum states [25–29], overlap between wavefunctions [30,31], bounded dimensionality [32,33], quantum steering [34], or quantum contextuality [35] as means to ensure that a certain level of min-entropy is achievable. Note that even a device-independent protocol can be considered as a semi-device-independent protocol if just a few assumptions are introduced. The ease of implementation, the high throughput, together with the security of semi-device-independent protocols make this class of QRNGs a valuable resource for applications. In this paper, we report a semi-device-independent QRNG based on single-photon entanglement (SPE), which is a particular kind of entanglement between distinct degrees of freedom of the same photon [36]. The certification scheme we demonstrate relies on the violation of a CHSH inequality using single-photon entangled states of momentum and polarization and on a model of the experimental setup. The model is based on the memory effects introduced by detectors and on the polarization dependence of the optical components (i.e., beam splitters and mirrors). The introduction of such a modeling allows the experimental implementation to be kept simple. Indeed, the use of SPE presents several advantages with respect to inter-photon entanglement. First, SPE states can be generated using classical light sources and off-the-shelf linear optical components. Second, it has been reported that SPE is more robust under decoherence and dephasing [37]. Thirdly, due to the local nature of this quantum phenomenon, coincidence detections performed at spacelike separated measurement stages are not necessary to test the CHSH inequality. Indeed, only single-photon detection events have to be collected using single-photon avalanche diodes (SPADs). Therefore, our experiment represents one of the attempts to make QRNGs based on photonic entanglement more accessible. The aim of the proposed semi-device-independent protocol is essentially robustness against unwanted and undetectable flaws of the system rather than security against a malicious eavesdropper, under the implicit assumption that the provider of the device is trusted. The estimate of min-entropy provided by our protocol relies on a minimal set of assumptions including the stationarity of the working conditions and the partial characterization of a few components of the setup.

The paper is organized as follows. In Sec. II, we introduce SPE for photons and we describe the general methodology used to generate the random numbers. The experimental setup is also described. In Sec. III, the entropy certification protocol is analyzed focusing on the CHSH inequality. The presence of the polarization nonidealities and of the memory effects is also discussed here. In Sec. IV, the experimental data supporting our claim for a certified QRNG based on SPE are shown and discussed, whereas conclusions are finally drawn in Sec. V, where a few future perspectives of this work are also given.

II. SETUP FOR QUANTUM RANDOM NUMBERS GENERATION

SPE is a type of entanglement in which a single particle, e.g., a photon, has two internal degrees of freedom entangled. We consider single-photon entangled states of the form:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|0V\rangle + |1H\rangle\right), \tag{1}$$

where the photon momentum ($|0\rangle$ and $|1\rangle$) is correlated with the photon polarization ($|V\rangle$ and $|H\rangle$). The wavefunction $|\psi\rangle$ belongs to the space $\mathbb{C}^2_M \otimes \mathbb{C}^2_P$, with obvious meaning of the subscripts.

The generation mechanism of the random numbers is schematized in Fig. 1. First, a SPE state of single photon of the form (1) is generated [Fig. 1(a)]. Second, the state undergoes separate rotations in momentum and polarization by angles ϕ and θ, respectively [Fig. 1(b)]. Third, the resulting state is measured and a random symbol is generated, according to which detector clicks [Fig. 1(c)]. The randomness is intrinsic to the quantum measurement process: the rotations modify the expectation values

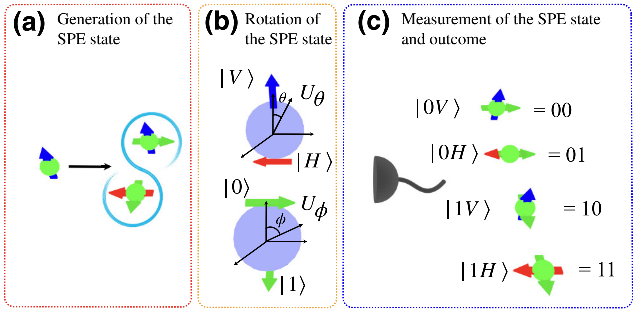

FIG. 1. Method to generate quantum random numbers: (a) a single-photon entangled (SPE) state is generated; (b) the two degrees of freedom of the SPE state are rotated by ϕ and θ, through unitary rotation operators (U_ϕ, U_θ), represented on the Bloch sphere; (c) single-photon avalanche diodes (SPADs) measure the rotated SPE state, i.e., the result of the SPE state projection on one of the four basis states ($|0V\rangle, |1V\rangle, |0H\rangle, |1H\rangle$); two bits are generated according to a coding that identifies the SPAD which has measured the photon. Therefore, a sequence of random numbers is the outcome of a sequence of quantum state projections.

of the projector operators over the four different states $|0V\rangle, |1V\rangle, |0H\rangle, |1H\rangle$ composing the Hilbert space $\mathbb{C}^2_M \otimes \mathbb{C}^2_P$ (see Appendix A for details). Then, by repeating the procedure n times, each time by varying the pair of angles (ϕ, θ), a sequence of n random symbols is obtained. Finally, the symbols are translated into binary numbers according to a coding [Fig. 1(c)]. This random number generation mechanism is the outcome of the projection operation of the state (1) over a rotated basis corresponding to a pair (ϕ, θ).

To implement this, we use the setup of Fig. 2(a). The state $|\psi\rangle$ is generated by using linear optical components and by operating on the single photons emitted by an attenuated green He-Ne laser [38]. First, a Glan-Thompson polarizer fixes the initial polarization state to $|0V\rangle$, namely single photons propagating through the $|0\rangle$ direction with $|V\rangle$ polarization. Then, by using a beam splitter, two half-wave plates, and two mirrors, the desired SPE state (1) is formed [red box in Fig. 2(a)]. Two waveplates are used to obtain equal phase retardation, one is rotated at $\pi/2$ angle to rotate the polarization, the second is placed at 0 angle. Additional phase mismatches are compensated for by adjusting the phase ξ by moving one of the mirrors. Note that, by exchanging the role of the wave plates in the paths or by setting the proper phase ξ, it is possible to obtain any Bell state.

The rotation of the SPE state is obtained by a Mach-Zehnder interferometer (MZI) and two half-wave plates [orange box in Fig. 2(a)]. The MZI rotates the momentum by an angle ϕ, whereas the two half-wave plates rotate the polarization by an angle θ. Note that the actual phase difference between the two arms of the MZI is 2ϕ.

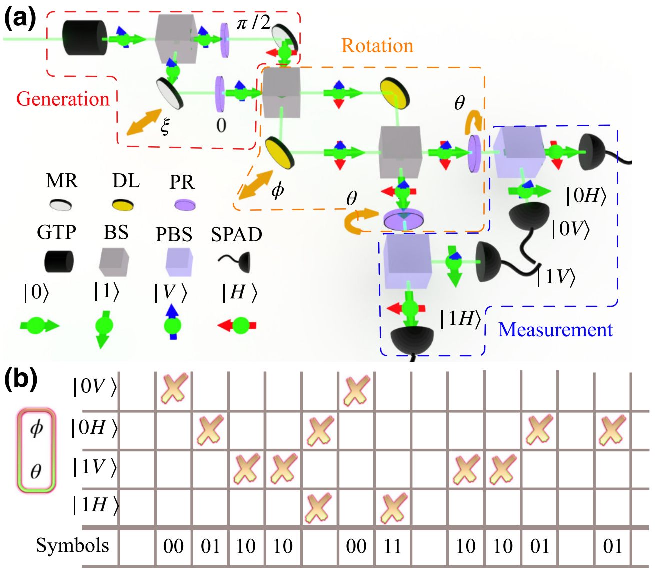

FIG. 2. (a) Experimental setup used to generate random numbers. It is composed by the generation (red box), the rotation (orange box) and the measurement (blue box) stages. The green line shows the optical path of the photons, represented by a tiny green sphere. The photon polarization degree of freedom is indicated by vertical blue and horizontal red arrows. Green arrows represent the photon momentum. In the rotation stage and in the measurement stage the beam splitters, the mirrors, and the single-photon avalanche diodes are characterized. The used linear optical components are labeled according to: GTP, Glan-Thompson polarizer; BS, beam splitter; MR, mirror; DL, delay line (three mirrors, see the Appendix); PBS, polarized beam splitter; PR, polarization rotator (half-wave plate); SPAD, single-photon avalanche diode; ξ, angle used to correct for phase differences in the generation; ϕ, momentum rotation angle; θ, polarization rotation angle. (b) Example of a generated raw random number sequence. This is produced by translating in bits the temporal sequence of symbols generated by coding the event "which SPAD detected the photon," given a certain pair (ϕ, θ). Multiple detection events and time bins with no detection are discarded.

The measurement of the rotated SPE state is done by using the two output ports of the MZI, two polarized beam splitters and four SPADs [blue box in Fig. 2(a)], whose properties are well known (trusted detectors in terms of efficiency, dark counts, dead time, and after-pulsing). The measure projects the rotated SPE state on the four possible single-photon states $|0V\rangle, |0H\rangle, |1V\rangle, |1H\rangle$. The sequence of single-photon detection events is registered by a time-tagging electronics and corresponds to the sequence of random numbers produced by our setup [Fig. 2(b)].

By using the same setup, we measured a Bell's inequality violation which witnessed entanglement between the momentum and the polarization degrees of freedom of the SPE states (1) [38]. It is here relevant to observe that, in the case of single-particle entanglement, the intrinsic randomness of outcomes is related to *contextuality* instead of nonlocality, as is for entanglement of a pair of particles [39]. A theory is contextual if the outcome of the measurement of any observable depends on the choice of other (compatible) observables simultaneously measured on the system. Absurdly, assuming contextuality in our experiment, the value of the momentum observable could depend on which basis is chosen for the polarization measurement. However, Bell's inequality violation excludes all noncontextual realistic hidden variable theories to explain our experiment [36,38]. Note that a realistic theory assumes that the outcomes of an experiment are predetermined. Therefore, the evidence of Bell's inequality violation implies that the measurement outcomes are not predetermined. In this sense, they are intrinsically random as assumed by standard quantum mechanics, which is nonrealistic (outcomes are not predetermined) and noncontextual (outcomes of the measurement of an observable do not depend on which other compatible observables are simultaneously measured). In contrast, in the case of entanglement of two particles, Bell's inequality violation rules out all *local* realistic hidden variable theories. In fact, quantum mechanics is local (no superluminal transmission of information is permitted by entanglement correlations) and nonrealistic.

III. PROTOCOL FOR ENTROPY CERTIFICATION

Having generated a sequence of numbers by a quantum process, the next step is to certify their randomness. This is usually done by evaluating the amount of quantum entropy intrinsic in the generated sequence [1]. A protocol to estimate the amount of quantum randomness in a device-independent QRNG exploiting the quantum correlations of a nonlocal entangled state between two spacelike separated particles was proposed in [10]. It is based on two different local measurement operations (U_x^1, U_y^2), applied to each particle, that give as output two binary results called a and b. These two operations depend on the value of two input bits (x, y), provided by another QRNG. Accumulated a sequence $R = (a_1, b_1; \ldots; a_n, b_n)$ of n outcomes of the form (a, b) for a certain input random sequence $S = (x_1, y_1; \ldots; x_n, y_n)$ of bits (x, y), the quantum correlations are evaluated by the CHSH correlation function:

$$I = \sum_{x,y} (-1)^{xy} (\mathbb{P}(a = b|x, y) - \mathbb{P}(a \neq b|x, y)), \quad (2)$$

where $\mathbb{P}(a = b|x, y)$ is the probability to measure $a = b$ given (x, y) and $\mathbb{P}(a \neq b|x, y)$ is the probability to measure $a \neq b$ given (x, y). Systems with a classical deterministic description satisfy $|I| \leq 2$, whereas certain measurements performed on entangled states violate this inequality, and the randomness of the sequence R is ensured. Moreover, the min-entropy $H_{\min}(R|S)$ of the outcome sequence R, given the initial sequence S, is given by [10]

$$H_{\min}(R|S) = -n \log_2 \left[\mathbb{P}_{\text{guess}}(a, b|x, y)\right], \quad (3)$$

where the guessing probability $\mathbb{P}_{\text{guess}}(a, b|x, y)$ is the largest probability to measure any outcome (a, b) given any (x, y). This is upper-bounded by [10]

$$\mathbb{P}_{\text{guess}}(a, b|x, y) \leq \frac{1}{2} + \frac{1}{2}\sqrt{2 - \frac{I^2}{4}}. \quad (4)$$

The inequality (4) gives an estimation of the effectiveness of the optimal strategy to guess the sequence R knowing the sequence S, whereas (3) provides an upper bound for the number of uniform random bits that can be extracted from the raw sequence of data R [2]. Randomness extraction is done by using known techniques [3], such as the Toeplitz extractor [40].

In our experiments, the angles (ϕ, θ) play the role of the inputs (x, y), therefore we labeled them (ϕ_x, θ_y) when we set them to evaluate (2). Then, for each measurement choice (ϕ_x, θ_y), a pair of outcomes (a, b) is produced, corresponding to a detection event in the path a with polarization b, where we identify $a = 0$ to $|0\rangle$, $a = 1$ to $|1\rangle$, $b = 0$ to $|V\rangle$, and $b = 1$ to $|H\rangle$. Measuring a sequence of outcomes, i.e., a sequence of single-photon detection events for a given input sequence (ϕ_x, θ_y), the Bell's inequality violation is evaluated through an estimator $\hat{I}$ of (2):

$$\hat{I}(\phi_0, \phi_1, \theta_0, \theta_1) = \sum_{x,y} (-1)^{xy} (\hat{\mathbb{P}}(a = b|\phi_x, \theta_y) - \hat{\mathbb{P}}(a \neq b|\phi_x, \theta_y)), \quad (5)$$

the probabilities $\hat{\mathbb{P}}(a, b|\phi_x, \theta_y)$ are computed as the maximum likelihood estimators of the probabilities $\mathbb{P}(a, b|\phi_x, \theta_y)$. As the inputs (ϕ_x, θ_y) are set during a scan, memory effects could correlate successive measurements outcomes. Memory might be caused by the nonidealities of the SPADs, e.g., dead time and after-pulsing, and by the use of an attenuated light source yielding stochasticity of the photons arrival times. Because of this, $\hat{\mathbb{P}}(a, b|\phi_x, \theta_y)$ is estimated by a Markovian model parameterized on the average photon flux detected by the SPADs, the after-pulsing probability and the SPAD dead time (see the appendix and [41]). The resulting guessing probability is defined $\mathbb{P}^*_{\text{guess}}(a, b|\phi_x, \theta_y)$. In our setting, SPAD dark counts are negligible compared with the photon flux from the source.

The protocol described in [10] relies on a loophole-free Bell test, hence it requires to randomly switch the observables to be measured. This requirement is important in a device-independent scenario where an eavesdropper is allowed to control the detectors or the source. Indeed, in this case the knowledge of the measurement basis would allow Eve to program the devices or to prepare the state in a way that would mimic Bell's inequality violation. However, if we assume that the provider of the device is trusted, the semi-device-independent protocol is essentially aimed to provide an estimate of min-entropy robust against unwanted flaws of the system. In this regard, the random switch of measurement basis is not necessary, because (4) is robust under classical side information, the latter including also the choice of the measurement basis, of the state and, if the latter is mixed, its decomposition into a mixture of pure states. In addition, in the device-independent protocol presented in Ref. [10], the random choice of ϕ and θ plays an important role in the construction of an estimator for the Bell parameter I in (2), allowing memory effects in the experimental devices to be tackled. In the present work, these issues are addressed by constructing estimators for the 16 quantum probabilities present in I by means of the Markov model (see Appendix D). It is important to point out that our protocol relies also on the assumption that the generation as well as the rotation parameters of the system are stable during the acquisition time. This requirement actually rules out the possibility that the state can change according to the choice of the measurement basis and it is reasonable in a trusted-provider setting. That said, random fluctuations of the generation parameters are still allowed and do not affect the estimate (4), provided that the noise is stationary. The last issue to compute (5) by using the SPE state as in (1) is the possible existence of a communication channel between the two degrees of freedom. In the setup, nonidealities of the beam splitters and of the mirrors of the MZI make momentum and polarization degrees of freedom no longer independent. Therefore, the actual characteristics of the optical elements (polarization-dependent reflectance, transmittance, and absorption) have to be measured and taken into account. Specifically, projection-valued measures (PVMs) describing the measurement operation cannot be written in the product form $\{P_{\phi_x}\otimes P_{\theta_y}\}_{x,y=0,1}$, where P_{ϕ_x} and P_{θ_y} are the projection operators for a given (ϕ_x,θ_y). To deal with this, we numerically evaluate an upper bound e_P for the difference between the ideal probabilities obtained by $\{P_{\phi_x}\otimes P_{\theta_y}\}$ and the estimated probabilities (named real probabilities in the following) obtained by modeling the experimental setup (see the Appendix C). e_P is calculated by considering any possible input state $\rho\in\mathbb{C}^2_M\otimes\mathbb{C}^2_P$, with $\phi,\theta\in[0,2\pi]$. As a consequence, an upper bound e_I to the difference $|\hat{I}_{\mathrm{ideal}}-\hat{I}_{\mathrm{real}}|$ between the ideal and the real CHSH correlation functions is computed as well. Thus, as long as these bounds are satisfied, (4) becomes

$$\mathbb{P}_{\mathrm{guess}}(a,b|\phi_x,\theta_y)\leq\frac{1}{2}+\frac{1}{2}\sqrt{2-(|\hat{I}_{\mathrm{real}}|-e_I)^2/4}+e_P. \tag{6}$$

Eventually, taking into account the memory effects, $\mathbb{P}^*_{\mathrm{guess}}(a,b|\phi_x,\theta_y)$ is estimated as

$$\begin{aligned}&\mathbb{P}^*_{\mathrm{guess}}(a,b|\phi_x,\theta_y)\\&\quad\leq\mathbb{M}\left(\frac{1}{2}+\frac{1}{2}\sqrt{2-(|\hat{I}_{\mathrm{real}}|-e_I)^2/4}+e_P\right),\end{aligned} \tag{7}$$

where $\mathbb{M}$ is a function which results from the Markovian model (see Appendix D and [41]). From $\mathbb{P}^*_{\mathrm{guess}}$, we calculate $H^*_{\min}$ associated with each measurement outcome as

$$H^*_{\min}=-\log_2\left[\mathbb{P}^*_{\mathrm{guess}}(a,b|\phi_x,\theta_y)\right], \tag{8}$$

and the min-entropy of the whole sequence R as

$$\begin{aligned}H_{\min}(R|S)&=nH^*_{\min}\\&=-n\log_2\left[\mathbb{P}^*_{\mathrm{guess}}(a,b|\phi_x,\theta_y)\right].\end{aligned} \tag{9}$$

As far as $H_{\min}(R|S)>0$, we can extract an unbiased sequence of random numbers from the raw data.

Up to now, the only assumption on the input state ρ regards its stationarity: the state, totally unknown, is fixed during the entire acquisition time. However, it is possible to increase $H^*_{\min}$ by introducing a few assumptions about it: these can lower the two upper bounds e_I and e_P. Let us assume we have nonideal optical components (see Fig. 3) and the presence of an eavesdropper. In this case, at the input of the first beam splitter of the MZI [orange box in Fig. 2(a)], ρ can be modeled as

$$\rho(v,\delta,\pi_1,\pi_2)=R(\pi_1,\pi_2)\rho_s(v,\delta)R(\pi_1,\pi_2)^\dagger. \tag{10}$$

Here $R(\pi_1,\pi_2)$ represents unwanted rotation of the wave plates by unknown angles $\pi_1\in[0,2\pi]$ and $\pi_2\in[0,2\pi]$ in the generation stage [red box of Fig. 2(a)], and $\rho_s(v,\delta)$ is the actual entangled state:

$$\rho_s(v,\delta)=v\left(|\psi(\delta)\rangle\langle\psi(\delta)|\right)+\frac{1-v}{4}\mathbb{I}_4, \tag{11}$$

where $v\in[0,1]$ is a visibility parameter which accounts for the nonideality of the setup and the different sources of noise [38], $\delta\in[0,2\pi]$ an additional phase,

$$|\psi(\delta)\rangle=\left(t_{0n}|0V\rangle+t_{1n}e^{i\delta}|1H\rangle\right), \tag{12}$$

$$t_{0n}=\frac{t_0}{\sqrt{t_0^2+t_1^2}},\qquad t_{1n}=\frac{t_1}{\sqrt{t_0^2+t_1^2}}, \tag{13}$$

and $\mathbb{I}_4$ is the identity matrix. The π_1,π_2,δ,v values might be due to the characteristics of the optical components or

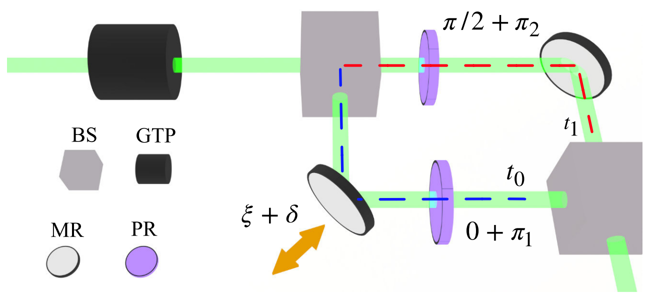

FIG. 3. Generation stage of the setup with indicated the non-idealities of the optical components. Here δ is an additional phase-shift between the two paths, π_1 and π_2 are additional angles of rotation of the waveplates, and t_0 and t_1 are transmission coefficients of the blue dashed and red dashed paths, respectively. The same optical component symbols as in Fig. 2 have been used.

Assumptions:

- Characterized beam splitters and mirrors in the rotation stage; characterized detectors in the measurement stage
- Stationarity of generation and rotation parameters during acquisition time
- (Partial knowledge of the input state ρ)

Protocol:

- Acquisition: for each (ϕ_x, θ_y) set by (x, y), check which of the four detectors reveals one photon. This determines (a, b)
- From the data and the Markovian model, estimate $\hat{\mathbb{P}}(a, b\,|\,\phi_x, \theta_y)$
- Estimate the CHSH correlation parameter $\hat{I}(\phi_0, \phi_1, \theta_0, \theta_1)$
- Calculate e_P and e_I based on the level of security (ρ)
- If $\mathbb{P}^*_{\text{guess}} < 1$, calculate $H^*_{\min}$
- Randomness extraction**

FIG. 4. Schematic summary of the assumptions and the main steps of the protocol for calculating the min entropy. **Note that despite this last step being necessary for extracting the sequence of random numbers, it is not implemented in our experiment.

to the action of the eavesdropper. t_0 and t_1 are the transmission coefficients of the optical paths $|0\rangle$ and $|1\rangle$ as shown in Fig. 3. t_{0n} and t_{1n} are the normalized transmission coefficients which account for the probability of photon transmission on a path with respect to the probabilities that the photon is transmitted on either of the paths. Every time one of the $t_{0n}, t_{1n}, v, \delta, \pi_1, \pi_2$ parameters is fixed, an assumption on the QRNG is introduced reducing its generality. Indeed, the eavesdropper could design an attack to change ρ with respect to (10), potentially jeopardizing the overall security. However, as long as an attack of this type is not feasible or is highly unlikely, the knowledge of the values of these parameters can increase $H^*_{\min}$ with uncompromising security. In the ideal case, neither experimental errors nor eavesdroppers are present, and the state (10) reduces to

$$\rho(v = 1, \delta = 0, \pi_1 = 0, \pi_2 = 0) = |\psi(0)\rangle\langle\psi(0)|. \quad (14)$$

Finally, our entropy certification protocol, being the result of a modeling based on the characteristics of the measurement setup [41], can be classified as semi-device-independent, providing a fair lower bound to the amount of measurable $H_{\min}$. The protocol and the relative assumptions are summarized in Fig. 4.

IV. EXPERIMENTAL DEMONSTRATION OF A CERTIFIED QRNG

The maximum violation of the Bell's inequality is obtained for $\{(\phi_x, \theta_y)\}_{x,y=0,1}$, where $\phi_0 = (3/8)\pi$ and $\phi_1 = (5/8)\pi$ for momentum, $\theta_0 = 0$ and $\theta_1 = (\pi/4)$ for polarization are chosen [38]. To set these values, we fixed $\theta_{0,1}$ and vary ϕ by a piezoelectric transducer actuated mirror (Fig. 2), so that sinusoidal sequences of counts (empty squares in Fig. 5) are measured and the ϕ calibration is obtained (solid lines in Fig. 5). Then, four time sequences of single-photon count rates for each pair (ϕ_x, θ_y) are acquired for 50 s with time bins of 1 μs (solid dots inside green boxes in Fig. 5). From these sequences, the clicking SPAD is identified in each time bin and the corresponding symbol is stored in the random symbol time sequence [Fig. 2(b)]. Multi-photon detection events within the same time bin are removed in postprocessing, which can be seen as a form of coincidence analysis. They constitute the $(12.0 \pm 0.4)\%$ of the raw data and are mainly due to the statistics of emission of the source. The whole numbers of counts per each SPAD acquired over 50 s integration windows are reported in Table I (see Appendix E for experimental details). The relative raw probabilities, estimated over a 10 ms acquisition time, are plotted in Fig. 6 for each (ϕ_x, θ_y). The observed slight probability increase or decrease can be attributed to the temporal stability of the setup, e.g., the piezoelectric transducer actuated mirrors. This affects $H_{\min}$ but it is mitigated by the certification protocol. More stable setup could result in a sequence of random numbers with higher $H_{\min}$. Table II lists the average raw probabilities. These have been estimated by dividing the 50 s acquisition interval in 10 s subintervals, computing the probabilities in the subintervals and getting the mean of these. This procedure yields also a $\simeq 0.2\%$ standard error on the raw probabilities. Table II also reports within parentheses the average probabilities $\hat{\mathbb{P}}(a, b|\phi_x, \theta_y)$ corrected by the Markovian model, which are equal to the raw probabilities within the errors. $\hat{\mathbb{P}}(a, b|\phi_x, \theta_y)$ are used to compute $\hat{I}$, for which we obtain

$$|\hat{I}(\phi_0, \phi_1, \theta_0, \theta_1)| = 2.656 \pm 0.003. \quad (15)$$

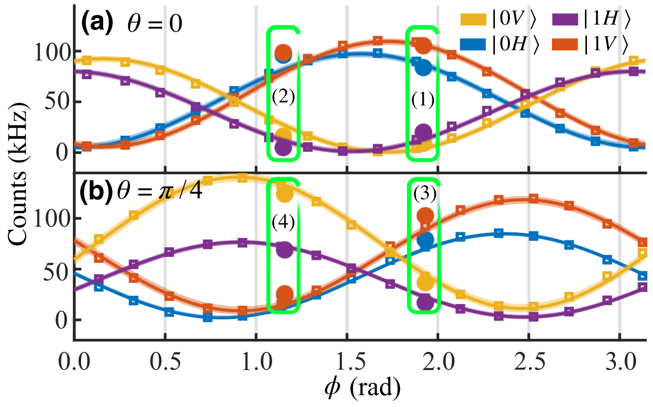

FIG. 5. Measured count rates (empty squares) as a function of ϕ and for (a) $\theta = 0$ and (b) $\theta = \frac{\pi}{4}$. The colors refer to the different states: yellow $|0V\rangle$, purple $|1H\rangle$, blue $|0V\rangle$, and red $|1V\rangle$. The solid lines are a sinusoidal fits (the 99% confidence intervals is shown by the shaded area around the solid line). The green boxes highlight the working points at which the 50 s long data acquisitions are performed. (1) corresponds to (ϕ_1, θ_0), (2) to (ϕ_0, θ_0), (3) to (ϕ_1, θ_1), and (4) to (ϕ_0, θ_1). The measured total count rates at these points are represented by the solid dots.

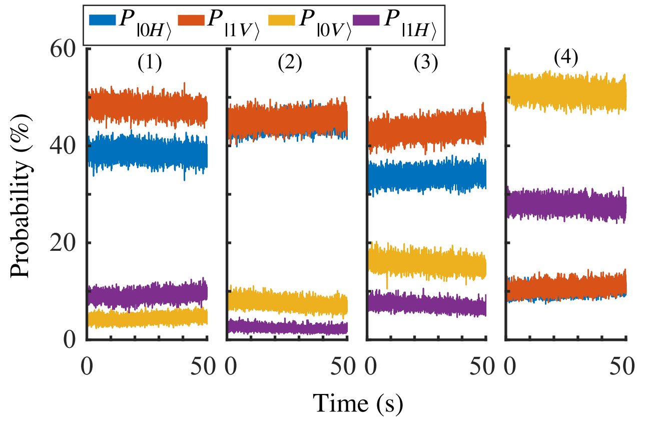

FIG. 6. Probabilities as a function of time for each measurement outcome (yellow $|0V\rangle$, purple $|1H\rangle$, blue $|0V\rangle$, and red $|1V\rangle$) at the four working points (1), (2), (3), and (4) of Fig. 5 and Table I, corresponding to (ϕ_1, θ_0), (ϕ_0, θ_0), (ϕ_1, θ_1), and (ϕ_0, θ_1), respectively. The estimates have been done considering time intervals of 10 ms.

From $\hat{I}(\phi_0, \phi_1, \theta_0, \theta_1)$ and considering an ideal setup, by (4) and (8) we obtain an overall certified min-entropy $H^*_{\min} = (42.8 \pm 0.4)\%$.

However, if we consider the setup nonidealities, we find $e_I = 0.332 \pm 0.008$ and $e_P = 0.080 \pm 0.002$ which yield a $H^*_{\min} = (2.5 \pm 0.5)\%$, according to (7) and (8). This number results from the assumption of a completely unknown input state and, therefore, represents the most conservative, i.e., most secure, estimate. Note that the upper bound to $\mathbb{P}_{\text{guess}}(a, b|x, y)$ given by (4) represents the best estimate for the marginal guessing probability, e.g., $\mathbb{P}_{\text{guess}}(b|y) = \max \sum_a \mathbb{P}_{\text{guess}}(a, b|x, y)$, where the marginal refers to a single degree of freedom. Therefore, from an operative point of view, we write as bit $b = 0$ each photon detected with vertical polarization $|V\rangle$ independently of its momentum state, and as bit $b = 1$ each photon detected with horizontal polarization $|H\rangle$ independently of its momentum state. This means that, having acquired a total amount of $\simeq 35 \times 10^6$ raw data, we can obtain $\simeq 0.88 \times 10^6$ certified random bits after the extraction procedure. Considering that the entire acquisition process requires 200 s, we can get a certified random bits generation rate $\simeq 4.4$ kHz, neglecting the randomness extraction phase.

Table III lists the increase in $H^*_{\min}$ as we increase the level of trust on the input state. Four different levels of trust are considered and the resulting upper bounds e_I and e_P together with the corresponding minimum entropies $H^*_{\min}$ and random number generation rates are reported. Fixing more parameters of the input state (10) means better knowledge of the working conditions of the setup, i.e., a better confidence level or trust, which means making more assumptions on the overall behavior of the system. Please note that the fixed parameters in the following discussions are actually measured from our experimental implementation and reported in Appendix B. In the fully trusted setup, i.e., when we know all the parameters but the visibility v,

TABLE I. Experimental counts. The number of counts acquired during the 50 s long acquisitions for the four experimental realizations $\{(\phi_x, \theta_y)\}_{x,y=0,1}$. The numbers $(i) = 1, 2, 3, 4$ refer to the different green boxes in Fig. 5.

Channel	$(\phi_0, \theta_0)(2)$	$(\phi_1, \theta_0)(1)$	$(\phi_0, \theta_1)(4)$	$(\phi_1, \theta_1)(3)$
$\lvert 0V\rangle$	643 132	371 255	4 754 594	1 426 837
$\lvert 1H\rangle$	202 823	779 771	2 589 956	652 294
$\lvert 0H\rangle$	3 804 170	3 311 003	964 121	3 078 159
$\lvert 1V\rangle$	3 855 004	4 108 774	996 276	3 945 250
Total	8 505 129	8 570 803	9 304 947	9 102 540

TABLE II. Measured raw and corrected probabilities. Mean values of the experimental ($\tilde{\mathbb{P}}$) probabilities for the four pairs of angles $\{(\phi_x, \theta_y)\}_{x,y=0,1}$. Within parenthesis the maximum-likelihood ($\hat{\mathbb{P}}$) probabilities are reported which result from the Markovian model. Errors on the probabilities of $\simeq 0.2\%$ are estimated based on repeated measurements. The numbers within parenthesis in the columns header refer to the measurement points shown in Fig. 5.

Channel	$(\phi_0, \theta_0)(2)$	$(\phi_1, \theta_0)(1)$	$(\phi_0, \theta_1)(4)$	$(\phi_1, \theta_1)(3)$
$\lvert 0V\rangle$	7.6(7.6)	4.3(4.3)	51.1(51.1)	15.7(15.7)
$\lvert 1H\rangle$	2.4(2.4)	9.1(9.1)	27.8(27.8)	7.2(7.2)
$\lvert 0H\rangle$	44.7(44.7)	38.6(38.6)	10.4(10.4)	33.8(33.8)
$\lvert 1V\rangle$	45.3(45.3)	47.9(48.0)	10.7(10.7)	43.3(43.4)

TABLE III. Min-entropy $H^*_{\min}$ and random bits throughput for different levels of trust of our SPE-based QRNG. We report the value of e_P, e_I, the min-entropy $H^*_{\min}$, and the random bit throughput (assumed instantaneous extraction procedure) for different levels of model for the input state ρ. Each row corresponds to a level of modeling, starting from the most general. In the first column there are the parameters over which the maximization of e_P and e_I is performed. In the second column, we reported the parameters fixed in the model. In the second row t_{0n}, t_{1n} are fixed following (13) and using the experimental value measured. In the third row, we fix $\pi_1 = \pi_2 = 0$. Lastly, in the fourth row δ is fixed to zero.

Variable	Fixed	$e_P \times 10^{-2}$	$e_I \times 10^{-2}$	$H^*_{\min}$	Random bits generation rate [kHz]
ρ general	–	8.0 ± 0.2	33.2 ± 0.8	$(2.5 \pm 0.5)\%$	4.4
δ, π_1, π_2, v	t_{0n}, t_{1n}	8.0 ± 0.2	26.4 ± 0.8	$(6.3 \pm 0.6)\%$	11.0
δ, v	$t_{0n}, t_{1n}, \pi_1, \pi_2$	7.8 ± 0.2	1.2 ± 0.2	$(26.9 \pm 0.5)\%$	47.1
v	$t_{0n}, t_{1n}, \pi_1, \pi_2, \delta$	6.6 ± 0.2	0.26 ± 0.07	$(30.1 \pm 0.5)\%$	52.7

(10) becomes

$$\rho(v, \delta = 0, \theta_1 = 0, \theta_2 = 0) \\ = v\,(|\psi(0)\rangle\langle\psi(0)|) + \frac{1-v}{4}\mathbb{I}_4, \qquad (16)$$

which differs from the ideal case, (14), for the presence of a parameter $v \leq 1$. In this fully trusted case, the maximum amount of min-entropy results to be $H^*_{\min} = (30.1 \pm 0.5)\%$.

If the control of the phase δ is removed, δ is free to vary within $[0, 2\pi]$. This could be considered as an additional phase coming from an unwanted erroneous or fraudulent control of the phase ξ (see Fig. 2). In this case, we obtain $H^*_{\min} = (26.9 \pm 0.5)\%$. Then, we can relax also the assumptions on the two polarization rotators (wave plates) present in the generation stage: an external eavesdropper could selectively rotate each one of them by an angle π_1 and/or π_2. This leads to $H^*_{\min} = (6.3 \pm 0.6)\%$.

V. CONCLUSION

In this work, we present a certified QRNG exploiting single-photon entangled states from an attenuated laser beam. Our results show that a simple setup combined with an accurate modeling of its optical elements and detectors can provide an entanglement-based high-security QRNG using imperfect devices. Our protocol can be classified as semi-device-independent because the entropy certification scheme relies on a certain number of assumptions. In particular, the model relies on the detailed knowledge of the technical specifications of the optical components of the setup and of the detectors. The first specifications are necessary to estimate e_P and e_I, the second specifications are related to some parameters present in the Markovian model. In addition, we assume that the generation and rotation parameters of the system are stable during the acquisition time. No additional assumptions on the state of the photons entering in the setup and on the detailed behavior of the measuring apparatus are needed, yielding an estimate of min-entropy robust under unwanted flaws of the system. We remark that, due to the less-than-unity detector efficiency, the model implicitly relies on the fair sampling assumption.

We demonstrate a certified quantum min-entropy $H^*_{\min} = 2.5\%$, whose value can be increased to 30.1% by raising the level of trust in the experimental setup. The certification relies on the violation of a Bell's inequality in the CHSH form. In our certification scheme, we take into account the nonidealites of our measurement setup: the presence of after-pulsing and dead time of SPADs, together with the polarization nonidealities of beam splitters and mirrors composing the MZI. The detectors introduce, in the sequence of outcomes, memory effects which are considered by a Markovian model, finally resulting in negligible corrections to the estimated probabilities. The polarization-based nonidealities are considered by calculating two upper bounds, e_I and e_P, of the difference between the ideal correlation function/probabilities and the measured ones. This accurate modeling allows one to consider either a most secure scenario (less trusted) where minimum assumptions are done, or a trusted setup scenario where the full knowledge of the components and of their setting is assumed. Consequently, we can move from a certified random number generation rate of 4.4 kHz to 52.7 kHz.

The generation rates, here reported, do not represent an improvement compared with other much faster semi-device-independent QRNGs [26,29,30]. Nevertheless, they can be further improved by increasing the laser photon flux up to the linearity limit of the SPADs, i.e., 1 MHz, or by reducing the time bin duration in the acquisition, which decreases the number of multi-photon detection events rejected by the protocol. With these improvements, we estimate a random bits generation of up to 500 kHz for the fully trusted case with the actual experimental setup. Having access to detectors with a lower dead time will further increase the rate of the QRNG. Moreover, the use of degrees of freedom in a setup without possible communication channels or the use of higher-quality optical components will lower the upper bounds e_I and e_P, resulting in a higher value of min-entropy up to the ideal value of almost 43%. In particular, the improvements highlighted above can find a natural implementation in a suitably designed integrated photonic platform, whose compactness and ease of integration with electronics will help making our type of photonic QRNG a potentially high-speed device.

To conclude, our work presents an experimental proof of a QRNG based on SPE. By exploiting the contextual phenomenology of entangled single-photon states, we propose an entropy certification based on the violation of a CHSH inequality using an experimental setup far simpler than typical device-independent ones.

ACKNOWLEDGMENTS

We acknowledge helpful discussions with C. Agostinelli on the code implementation of the Markovian model. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820405. N.L. was supported by a fellowship of Q@TN within the PAT(AI) grant.

The original experiment was conceived by all the authors. N.L. and S.A. performed the experiment in the laboratory, under the supervision of L.P. S.M. developed the Markovian model used in the experiment. S.M. and N.L. carried out the analysis of the imperfection with the help of V.M and S.A. All the authors discussed the results and contributed to the writing of the manuscript.

APPENDIX A: SINGLE-PHOTON ENTANGLEMENT, IDEAL IMPLEMENTATION

In our experimental implementation [setup shown in Fig. 2(a)] we use the momentum and polarization degrees of freedom of a single photon. This system can be represented as a 2-qubit system: the two directions of propagation $\{0, 1\}$ are two orthonormal basis vectors in the momentum Hilbert space $\mathcal{H}_M$ whereas the two polarizations $\{V, H\}$, vertical and horizontal, correspond to two orthonormal basis vectors in the Hilbert space $\mathcal{H}_P$. Consequently, the four vectors $\{|0H\rangle, |0V\rangle, |1H\rangle, |1V\rangle\}$ are an orthonormal basis in the tensor product Hilbert space $\mathcal{H}_M \otimes \mathcal{H}_P$. The two polarization beam splitters and the four detectors in the measurement stage [Fig. 2(a)] implement the PVM corresponding to the four orthogonal projectors:

$$|0V\rangle\langle 0V|, \quad |0H\rangle\langle 0H|, \quad |1V\rangle\langle 1V|, \quad |1H\rangle\langle 1H|. \tag{A1}$$

Defined $\{P^M_{+1}, P^M_{-1}\}$ ($\{P^P_{+1}, P^P_{-1}\}$) the PVM associated with the momentum observable $O^M = \sigma_z$ (polarization observable $O^P = \sigma_z$), the projectors in (A1) can be written in product form as $P^M_a \otimes P^P_b$, where the superscripts M (P) denote momentum (polarization), and $a, b \in \{+1, -1\}$. For any unit vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$, it is possible to define $U_{\boldsymbol{u}} : \mathcal{H}_M \to \mathcal{H}_M$ and $U_{\boldsymbol{v}} : \mathcal{H}_P \to \mathcal{H}_P$, unitary operators that transform the PVM of the operator $O^M = \sigma_z$ and $O^P = \sigma_z$ into the PVM $\{P^{\boldsymbol{u}}_{+1}, P^{\boldsymbol{u}}_{-1}\}$ and $\{P^{\boldsymbol{v}}_{+1}, P^{\boldsymbol{v}}_{-1}\}$ of the operators $\boldsymbol{u} \cdot \sigma$ and $\boldsymbol{v} \cdot \sigma$:

$$P^{\boldsymbol{u}}_{+1} = U^\dagger_{\boldsymbol{u}} P^M_{+1} U_{\boldsymbol{u}}, \quad \mathbb{P}^{\boldsymbol{u}}_{-1} = U^\dagger_{\boldsymbol{u}} P^M_{-1} U_{\boldsymbol{u}}. \tag{A2}$$

$$P^{\boldsymbol{v}}_{+1} = U^\dagger_{\boldsymbol{v}} P^P_{+1} U_{\boldsymbol{v}}, \quad \mathbb{P}^{\boldsymbol{v}}_{-1} = U^\dagger_{\boldsymbol{v}} P^P_{-1} U_{\boldsymbol{v}}. \tag{A3}$$

Here $\sigma = (\sigma_x, \sigma_y, \sigma_z)$, where $\{\sigma_i\}_{i=x,y,z}$ are the Pauli matrices, and $\boldsymbol{u} \cdot \sigma := u_x\sigma_x + u_y\sigma_y + u_z\sigma_z$. For a generic state ρ in $\mathcal{H}_M \otimes \mathcal{H}_P$, the probability of the outcomes (a, b) for the joint measurements of the observable $O^{\boldsymbol{u}} = \boldsymbol{u} \cdot \sigma$ (for the momentum degree of freedom) and of the observable $O^{\boldsymbol{v}} = \boldsymbol{v} \cdot \sigma$ (for the polarization degree of freedom) is given by

$$\begin{aligned}
\mathbb{P}(a, b|\rho, \boldsymbol{u}, \boldsymbol{v}) &= \mathrm{Tr}[\rho P^{\boldsymbol{u}}_a \otimes P^{\boldsymbol{v}}_b] && \text{(A4)}\\
&= \mathrm{Tr}[\rho U^\dagger_{\boldsymbol{u}} P^M_a U_{\boldsymbol{u}} \otimes U^\dagger_{\boldsymbol{v}} P^P_b U_{\boldsymbol{v}}] && \text{(A5)}\\
&= \mathrm{Tr}[\rho (U^\dagger_{\boldsymbol{u}} \otimes U^\dagger_{\boldsymbol{v}}) (P^M_a \otimes P^P_b) (U_{\boldsymbol{u}} \otimes U_{\boldsymbol{v}})] && \text{(A6)}\\
&= \mathrm{Tr}[(U_{\boldsymbol{u}} \otimes U_{\boldsymbol{v}}) \rho (U^\dagger_{\boldsymbol{u}} \otimes U^\dagger_{\boldsymbol{v}}) P^M_a \otimes P^P_b]. && \text{(A7)}
\end{aligned}$$

The last line shows that the statistical distribution of the measurement of $\{P^{\boldsymbol{u}}_a \otimes P^{\boldsymbol{v}}_b\}_{a,b=\pm 1}$ on a state ρ coincides with the distribution of outcomes of the measurement described by the PVM (A1) over the rotated state $\rho_{\boldsymbol{u},\boldsymbol{v}}$, which is given by

$$\rho_{\boldsymbol{u},\boldsymbol{v}} = U_{\boldsymbol{u}} \otimes U_{\boldsymbol{v}}\, \rho\, U^\dagger_{\boldsymbol{u}} \otimes U^\dagger_{\boldsymbol{v}}. \tag{A8}$$

The first part of the preparation stage shown in Fig. 2(a) (orange box), consisting of a MZI, realizes the rotation operator $U_{\boldsymbol{u}} \otimes I$, whereas the second part, consisting in two half-wave plates, implements the operator $I \otimes U_{\boldsymbol{v}}$. Actually, the first part sets the measurement basis of the momentum observable $O^{\boldsymbol{u}} = \boldsymbol{u} \cdot \sigma$, whereas the second part does the same for the polarization observable $O^{\boldsymbol{v}} = \boldsymbol{v} \cdot \sigma$.

The rotation operation for the momentum $U_{\boldsymbol{u}}$, in the case in which lossless balanced beam splitters are used, is unitary and can be obtained as $U_{\boldsymbol{u}} = V_{\mathrm{BS}} V(\phi) V_{\mathrm{MR}} V_{\mathrm{BS}}$, where:

$$V_{\mathrm{BS}} = \begin{pmatrix} \sqrt{0,5} & i\sqrt{0,5} \\ i\sqrt{0,5} & \sqrt{0,5} \end{pmatrix},$$

$$V(\phi) = \begin{pmatrix} e^{2i\phi} & 0 \\ 0 & 1 \end{pmatrix},$$

$$V_{\mathrm{MR}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The last operator V_{MR} represents the swapping action of the mirrors between the two beam splitters, in which an

TABLE IV. Measured transmission coefficients relative to the optical paths $|0\rangle$ and $|1\rangle$ in the generation stage. For each optical path in the generation stage of Fig. 2(a), the measured values at 543.5 nm are reported.

$\|t_0\|^2$	0.421 ± 0.002
$\|t_1\|^2$	0.456 ± 0.001

identical optical response for the vertical and horizontal polarization is assumed. Computing $U_{\boldsymbol{u}}$, we obtain:

$$U_{\boldsymbol{u}} = \begin{pmatrix} \frac{i+ie^{2i\phi}}{2} & \frac{e^{2i\phi}-1}{2} \\ \frac{1-e^{2i\phi}}{2} & \frac{i+ie^{2i\phi}}{2} \end{pmatrix} = ie^{i\phi}\begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix}.$$

The last formula shows that the rotation angle is induced by the phase mismatch between the two arms of the MZI.

In the polarization Hilbert space, the rotation operator $U_{\boldsymbol{v}}$, has a form:

$$U_{\boldsymbol{v}} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix},$$

which is analogous to $U_{\boldsymbol{u}}$.

APPENDIX B: SINGLE-PHOTON ENTANGLEMENT, ACTUAL IMPLEMENTATION

Here the actual parameters for the various optical components of the setup [Fig. 2(a)] can be found. We list the values of the power reflection and transmission coefficients for vertical and horizontal polarization. These have been measured on the used components at the working wavelength of 543.5 nm. Finally, we report the dead time value, the after-pulsing probability, and the dark counts rate of our SPADs.

1. Generation stage

The coefficients t_0 and t_1 of Eq. (13) are calculated as

$$t_0 = \sqrt{\frac{I_0}{I_{\text{in}}}}, \quad t_1 = \sqrt{\frac{I_1}{I_{\text{in}}}} \tag{B1}$$

and their values are reported in Table IV. Here I_0, I_1 and I_{in} refer to the optical power measured in the optical path $|0\rangle$ and $|1\rangle$ of Fig. 3. These coefficients take into account the transmission of the first beam splitter, the reflection of the mirrors and the transmission of the half-wave plates of the generation stage.

TABLE V. Power transmission and reflection coefficients for the two beam splitters composing the MZI and for the two polarizations. The measurements were done at 543.5 nm and for two incident light polarizations (V vertical, H horizontal).

	BS_1	BS_2
$\|t_V\|^2$	0.502 ± 0.005	0.476 ± 0.003
$\|r_V\|^2$	0.423 ± 0.003	0.416 ± 0.001
$\|t_H\|^2$	0.511 ± 0.002	0.4865 ± 0.001
$\|r_H\|^2$	0.349 ± 0.001	0.3583 ± 0.0007

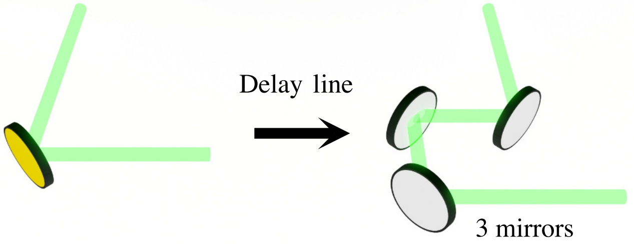

FIG. 7. Delay Line. Scheme of the delay line (DL) mirrors of the MZI of Fig. 2(a). The delay line mirrors are actually composed by three mirrors as shown here.

2. Rotation stage

In Table V we report the measured values of transmission and reflection coefficients of the two beam splitters that compose the MZI in Fig. 2(a), for vertical and horizontal polarization.

The mirrors in the MZI of Fig. 2(a) are composed by a series of mirrors to ease the optical alignment, as reported in Fig. 7. In the figure they are named delay line because they allow an increase/decrease of one path length with respect to the other. The measured power transmission coefficients for the two polarizations of these components are reported in Table VI.

3. Detection stage

In Table VII we report the typical values of dead time, after-pulsing probability, and dark count rate for our SPAD, as given by the manufacturer.

APPENDIX C: TREATMENT OF THE NONIDEALITIES OF THE OPTICAL COMPONENTS IN ESTIMATING THE GUESSING PROBABILITIES

The device-independent bound developed in [10] for the guessing probability relies on the product form of the PVM

TABLE VI. Power transmission coefficients for the two delay lines used in the MZI. For each delay line in the MZI (DL in Fig. 2) we report its measured transmission coefficient for the selected polarization.

	DL_1	DL_2
$\|\gamma_V\|^2$	0.898 ± 0.005	0.872 ± 0.006
$\|\gamma_H\|^2$	0.798 ± 0.004	0.771 ± 0.002

TABLE VII. SPAD characteristics. Typical values of dead time (T_d), after-pulsing probability(p_a), and dark count rate (DCR) for the SPADs used in the experiment.

	T_d [ns]	p_a [%]	DCR [Hz]
SPAD	22	0.5	< 100

$\{P_a^{\boldsymbol{u}} \otimes P_b^{\boldsymbol{v}}\}_{a,b=\pm 1}$ describing the joint measurement of the two observables $O^{\boldsymbol{u}}$ and $O^{\boldsymbol{v}}$. When considering spacelike separated or shielded systems, this condition is fulfilled unless the Bell test is affected by the locality loophole. However, more realistic and technologically accessible implementations require an additional set of rather mild assumptions. As discussed in the main text, we assume that the manufacturer of the devices is trusted and that the technical features of the components fulfill a few additional requirements. Specifically, all components present some nonidealities and these must be taken into account in an estimation of the effective guessing probability. In [41] a theoretical justification of the methodology here described is provided.

For the implemented QRNG based on SPE, the product form of the rotation operator $U_{\boldsymbol{u}} \otimes U_{\boldsymbol{v}}$ is essential to guarantee the product form of the PVM $\{P_a^{\boldsymbol{u}} \otimes P_b^{\boldsymbol{v}}\}_{a,b=\pm 1}$. As discussed in Appendix A, the rotation operation $U_{\boldsymbol{u}} \otimes U_{\boldsymbol{v}}$ is obtained using a MZI and two half-wave plates, which independently rotate the momentum and polarization qubits. The MZI implements the operation $U_{\boldsymbol{u}} \otimes I$, and the two half-wave plates $I \otimes U_{\boldsymbol{v}}$. Although the real description of the two half-wave plates does not present issues regarding the implementation of the product form of $I \otimes U_{\boldsymbol{v}}$, the real description of the MZI presents some problems regarding the realization of $U_{\boldsymbol{u}} \otimes I$. In particular, the beam splitters (BSs) and the mirrors (MRs) in the MZI [see Fig. 2(a)] present some polarization nonidealities. Specifically, the transmission and the reflection coefficients for the two polarizations, vertical and horizontal, are different. The matrix representation of the beam splitter in the basis $\{|0V\rangle, |0H\rangle, |1V\rangle, |1H\rangle\}$ can be modeled as

$$U_{\rm BS}^{\rm real} = \begin{pmatrix} t_V & 0 & ir_V & 0 \\ 0 & t_H & 0 & ir_H \\ ir_V & 0 & t_V & 0 \\ 0 & ir_H & 0 & t_H \end{pmatrix}, \qquad \begin{array}{l} |t_V|^2 + |r_V|^2 \le 1 \\ |t_H|^2 + |r_H|^2 \le 1 \end{array} \tag{C1}$$

where $t_{V,H}, r_{V,H}$ are the transmission and reflection coefficients of the beam splitters for the corresponding polarization. Analogously, the matrix representing the action of the mirrors can be written as

$$U_{\rm MR}^{\rm real} = \begin{pmatrix} 0 & 0 & \gamma_{V_1} & 0 \\ 0 & 0 & 0 & \gamma_{H_1} \\ \gamma_{V_2} & 0 & 0 & 0 \\ 0 & \gamma_{H_2} & 0 & 0 \end{pmatrix}, \qquad \begin{array}{l} |\gamma_{x_y}|^2 \le 1, \\ \forall x \in \{V, H\}, \forall y \in \{1,2\}, \end{array} \tag{C2}$$

where γ_{x_y} is the transmission coefficient for the polarization x of the mirror y of the MZI. Due to these nonidealities, the rotation operator $U_{\phi,\theta}^{\rm real}$ [42], is no longer the tensor product $U_\phi \otimes U_\theta$ as in the ideal case. Moreover, the rotation operator $U_{\phi,\theta}^{\rm real}$ is not unitary: due to the presence of losses (absorption and scattering for the different optical components) some of the photons are actually lost during the propagation into the experimental setup. Consequently, the statistics of detection outcomes is determined only by the revealed photons and the probabilities of the possible outcomes (a, b) have to be estimated as

$$\mathbb{P}^{\rm real}(a, b|\rho, \phi, \theta) = \frac{{\rm Tr}[U_{\phi,\theta}^{\rm real}\rho(U_{\phi,\theta}^{\rm real})^\dagger P_a^M \otimes P_b^P]}{{\rm Tr}[U_{\phi,\theta}^{\rm real}\rho(U_{\phi,\theta}^{\rm real})^\dagger]}, \tag{C3}$$

where

$$U_{\phi,\theta}^{\rm real} = (I_2 \otimes U_\theta)U_{{\rm BS}_2}^{\rm real}(V(\phi) \otimes I_2)U_{\rm MR}^{\rm real}U_{{\rm BS}_1}^{\rm real}. \tag{C4}$$

In the last equation we have considered that we have two different beam splitters with different values of t_V, t_H, r_V, r_H:

$$U_{BS_1}^{\rm real} = \begin{pmatrix} t_{V_1} & 0 & ir_{V_1} & 0 \\ 0 & t_{H_1} & 0 & ir_{H_1} \\ ir_{V_1} & 0 & t_{V_1} & 0 \\ 0 & ir_{H_1} & 0 & t_{H_1} \end{pmatrix}, \quad U_{BS_2}^{\rm real} = \begin{pmatrix} t_{V_2} & 0 & ir_{V_2} & 0 \\ 0 & t_{H_2} & 0 & ir_{H_2} \\ ir_{V_2} & 0 & t_{V_2} & 0 \\ 0 & ir_{H_2} & 0 & t_{H_2} \end{pmatrix}. \tag{C5}$$

Due to all these nonidealities, the probabilities $\mathbb{P}^{\rm real}(a, b|\rho, \phi, \theta)$ differ from (A7) or (A4).

We recall that the uniform bound developed in [10] refers to the maximum attainable (realization-independent) guessing probability for the outcomes of measurements of observables of the product form. To cope with this fact, we numerically compute a bound, e_P, for the difference between the real probabilities (C3) and the probabilities:

$$\mathbb{P}^{\rm ideal}(a, b|\rho, \phi, \theta) = {\rm Tr}[\tilde{U}_{\phi,\theta}^{\rm ideal}\rho(\tilde{U}_{\phi,\theta}^{\rm ideal})^\dagger P_a^M \otimes P_b^P], \tag{C6}$$

with

$$\tilde{U}_{\phi,\theta}^{\rm ideal} = (I \otimes U_\theta)\tilde{U}_{{\rm BS}_2}^{\rm ideal}(V(\phi) \otimes I_2)(V_{\rm MR} \otimes I_2)\tilde{U}_{{\rm BS}_1}^{\rm ideal}. \tag{C7}$$

The unitary operators $\tilde{U}_{{\rm BS}_1}^{\rm ideal}$ and $\tilde{U}_{{\rm BS}_2}^{\rm ideal}$ are defined as

$$\tilde{U}^{\text{ideal}}_{\text{BS}_1} = \begin{pmatrix} \cos\alpha & 0 & i\sin\alpha & 0 \\ 0 & \cos\alpha & 0 & i\sin\alpha \\ i\sin\alpha & 0 & \cos\alpha & 0 \\ 0 & i\sin\alpha & 0 & \cos\alpha \end{pmatrix},$$
$$\tilde{U}^{\text{ideal}}_{\text{BS}_2} = \begin{pmatrix} \cos\beta & 0 & i\sin\beta & 0 \\ 0 & \cos\beta & 0 & i\sin\beta \\ i\sin\beta & 0 & \cos\beta & 0 \\ 0 & i\sin\beta & 0 & \cos\beta \end{pmatrix}. \tag{C8}$$

Please note that, for each choice of $\alpha, \beta \in [0, \pi/2]$, the two operators $U^{\text{ideal}}_{\text{BS}_{1,2}}$ are in the product form, which means that the operator $\tilde{U}^{\text{ideal}}_{\phi,\theta}$ is in product form. The probabilities $\mathbb{P}^{\text{ideal}}(a, b|\rho, \phi, \theta)$ correspond, indeed, to observables in product form. This fact enters directly in the calculation of e_P, defined as

$$e_P = \min_{\{\alpha,\beta\}} \left[\max_{\{\phi,\theta,\rho,a,b\}} \left| \frac{\text{Tr}[U^{\text{real}}_{\phi,\theta}\rho (U^{\text{real}}_{\phi,\theta})^\dagger P^M_a \otimes P^P_b]}{\text{Tr}[U^{\text{real}}_{\phi,\theta}\rho (U^{\text{real}}_{\phi,\theta})^\dagger]} - \text{Tr}[\tilde{U}^{\text{ideal}}_{\phi,\theta}\rho (\tilde{U}^{\text{ideal}}_{\phi,\theta})^\dagger P^M_a \otimes P^P_b] \right| \right]. \tag{C9}$$

Indeed, because the bound of [10] is realization independent, we can choose the best pair of (α, β) that minimizes the maximum distance over the parameters $\{\phi, \theta, \rho, a, b\}$.

The same reasoning can be applied to the correlation function I. Indeed we can define

$$I^{\text{ideal}}(\phi_0, \phi_1, \theta_0, \theta_1) = \sum_{x,y} (-1)^{xy} \left(\mathbb{P}^{\text{ideal}}(a = b|\rho, \phi_x, \theta_y) - \mathbb{P}^{\text{ideal}}(a \neq b|\rho, \phi_x, \theta_y)\right). \tag{C10}$$

and

$$I^{\text{real}}(\phi_0, \phi_1, \theta_0, \theta_1) = \sum_{x,y} (-1)^{xy} (\mathbb{P}^{\text{real}}(a = b|\rho, \phi_x, \theta_y) - \mathbb{P}^{\text{real}}(a \neq b|\rho, \phi_x, \theta_y)). \tag{C11}$$

As before, we define

$$e_I = \min_{\{\alpha,\beta\}} \left[\max_{\{\phi_0,\phi_1,\theta_0,\theta_1,\rho,a,b\}} \left| I^{\text{ideal}}(\phi_0, \phi_1, \theta_0, \theta_1) - I^{\text{real}}(\phi_0, \phi_1, \theta_0, \theta_1) \right| \right], \tag{C12}$$

where we take the minimum value over (α, β) for the maximum over the parameters $\{\phi_0, \phi_1, \theta_0, \theta_1, \rho, a, b\}$.

Lastly we need to model the state ρ [Eq. (10)]:

$$\rho(v, \delta, \pi_1, \pi_2) = R(\pi_1, \pi_2)\rho_s(v, \delta)R(\pi_1, \pi_2)^\dagger. \tag{C13}$$

Here, we give the explicit description of $R(\pi_1, \pi_2)$:

$$R(\pi_1, \pi_2) = \begin{pmatrix} \cos(\pi_1) & \sin(\pi_1) & 0 & 0 \\ -\sin(\pi_1) & \cos(\pi_1) & 0 & 0 \\ 0 & 0 & \cos(\pi_2) & \sin(\pi_2) \\ 0 & 0 & -\sin(\pi_2) & \cos(\pi_2) \end{pmatrix}. \tag{C14}$$

In addition, we also write $\rho_s(v, \delta)$ in the matrix form:

$$\rho_s(v, \delta) = \begin{pmatrix} v|t_{0n}|^2 + (1 - v) & 0 & 0 & v t_{0n}(t_{1n}e^{i\delta})^* \\ 0 & 1 - v & 0 & 0 \\ 0 & 0 & 1 - v & 0 \\ v t^*_{0n} t_{1n} e^{i\delta} & 0 & 0 & v|t_{1n}|^2 + (1 - v) \end{pmatrix}, \tag{C15}$$

where we recall that t_{0n} and t_{1n} are the normalized transmission coefficients of the two paths $|0\rangle$ and $|1\rangle$, as reported in Eq. (12). In the most general scenario, ρ can be written as

$$\rho = (A + iB)^\dagger (A + iB) \tag{C16}$$

where A and B are

$$A = \begin{pmatrix} x_1 & x_5 & x_8 & x_{10} \\ 0 & x_2 & x_6 & x_9 \\ 0 & 0 & x_3 & x_7 \\ 0 & 0 & 0 & x_4 \end{pmatrix},$$

$$B = \begin{pmatrix} 0 & x_{11} & x_{14} & x_{16} \\ 0 & 0 & x_{12} & x_{15} \\ 0 & 0 & 0 & x_{13} \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{C17}$$

To ensure that ρ is semi-definite positive and normalized, we have used the Cholesky decomposition [43] and the spherical 16-dimensional coordinates $\{x_i\}_{i=1\ldots16}$. The latter are unequivocally fixed by the choice of the radius r, which is fixed to the value 1 and by 15 angles $\{\eta_i\}_{i=1\ldots15}$:

$$\begin{aligned} x_1 &= \cos(\eta_1), \\ x_2 &= \cos(\eta_2)\sin(\eta_1), \\ x_3 &= \cos(\eta_3)\sin(\eta_2)\sin(\eta_1), \\ x_4 &= \cos(\eta_4)\sin(\eta_3)\sin(\eta_2)\sin(\eta_1), \\ x_5 &= \cos(\eta_5)\sin(\eta_4)\cdots\sin(\eta_1), \\ \ldots \\ x_{15} &= \cos(\eta_{15})\sin(\eta_{14})\cdots\sin(\eta_1), \\ x_{16} &= \sin(\eta_{15})\cdots\sin(\eta_1). \end{aligned} \tag{C18}$$

To ensure that the diagonal terms x_1, x_2, x_3, and x_4 are positive, we restrict the angles $\{\eta_i\}_{i=1\ldots4}$ to $[0, \pi/2]$. The other angle domains are $\{\eta_i\}_{i=5\ldots14} \in [0, \pi]$ and $\eta_{15} \in [0, 2\pi]$.

1. Numerical evaluation

The maximization procedure over the parameters $\{\phi_0, \phi_1, \theta_0, \theta_1,\ \rho, a, b\}$ for e_I and $\{\phi, \theta, \rho, x, y\}$ for e_P has been performed using sequential quadratic programming (SQP). This is implemented using the Global Optimization Toolbox of the MultiStart solver by MATLAB, using the standard SQP algorithm. The operators $U^{\rm real}_{\rm BS_1}$, $U^{\rm real}_{\rm BS_2}$, and $U^{\rm real}_{\rm MR}$ are built using the measured values of transmission and reflection coefficients for our experimental setup. As the problem necessitates to minimize the upper bounds for (α, β) we have followed the procedure:

(a) fix a pair of (α, β);

(b) perform a rough maximization over the variable parameters, calculating e_I and e_P;

(c) calculate the guessing probability;

(d) repeat.

The pair of angles (α, β) which gives the lowest guessing probability is the best. To ensure that the result we have found is accurate, we have repeated the optimization starting from several initial points: the maximum value obtained is considered to be the global maximum of the function. We have decided to set the number of starting points to 3×10^3 for the optimization of e_P and 10^4 for e_I. The Matlab solver converges to the values reported in the main text. The errors on e_P and e_I are estimated in the following way: for each experimental quantity measured (transmission and reflection coefficients of the optical components) we create a normal distribution having as mean the measured value, and as standard deviation the error of the measurements. From each of these distributions we randomly extract one value and we calculate e_P and e_I keeping fixed the values of the variable parameters that maximize e_I and e_P obtained before. In this way, we can find an error which depends only on the experimental errors over the measurement performed on the optical components. This operation is repeated 4×10^3 times and we take as error the standard deviation of these obtained values. We must remark that the important part of this numerical evaluation is the maximization part, which is necessary to guarantee a bound which is independent of the parameters, whereas the minimization affects only the efficiency of the QRNG and not its security.

APPENDIX D: MARKOVIAN MODEL FOR DETECTORS NONIDEALITIES

Here, we introduce the Markovian model used to estimate the detection probabilities $\mathbb{P}^{\rm real}(a, b|\rho, \phi, \theta)$ [see (C3)] from the experimental data.

Set a particular choice of the pair of parameters ϕ, θ, we denote the four probabilities $\{\mathbb{P}^{\rm real}(a, b|\rho, \phi, \theta)\}_{a,b=\pm1}$ with the symbols $p_j, j = 1, 2, 3, 4$, to ease the following discussion. In addition to the notational simplicity, this choice is motivated by the consideration that each of the four values provides the probability that the incoming photon is detected by the jth detector present in the final stage of the experimental setting.

In a realistic implementation, as discussed in the main text, the presence of after-pulsing and detector dead time produces correlations among the subsequent readouts of the detectors, which in fact can no longer be consider independent. These memory effects can be described in terms of a Markovian model, which is based on the following set of assumptions:

(a) the source has a Poissonian emission statistics with effective intensity parameter $\lambda_e = \eta\lambda$, where $\eta < 1$ is the detector efficiency, assumed equal for the four SPADs;

(b) p_a, the probability of after-pulsing, is 0.5% or less;

(c) the dead time of detectors, T_d, is negligible with respect to the average inter-arrival time $1/\lambda_e$; more precisely, we assume that the product $\lambda_e T_d$ is of the order 10^{-2} or less;

(d) the time of after-pulsing T_a and the dead time T_d are of the same order ($T_a \sim T_d$).

In the following the sequence of subsequent readouts of the detectors will be described in terms of a sequence $\{\eta_n\}_n$ of random variables with four possible outcomes $i = 1, 2, 3, 4$ (i.e., $\eta_n = i$ means that the nth incoming photon is detected by the ith SPAD). According to the approximations above, it is possible to prove (see [41] for further details) that the sequence $\{\eta_n\}$ is a stationary Markov chain, i.e., the probability of predicting the nth readout given the previous $(n-1)$ is equal to

$$\begin{aligned}&\mathbb{P}(\eta_n = i_n|\eta_{n-1} = i_{n-1}, \dots \eta_0 = i_0)\\&= \mathbb{P}(\eta_n = i_n|\eta_{n-1} = i_{n-1})\\&= \mathbb{P}(\eta_1 = i_n|\eta_0 = i_{n-1}), \qquad \forall i_0, \dots i_{n-1},\ i_n = 1,2,3,4.\end{aligned}$$

Further, the transition probabilities $\mathbb{P}(\eta_{n+1} = j\,|\eta_n = i) = P_{ij}$ are given by

$$P_{ij} = p_a\delta_{ij} + (1-p_a)\left((1-\lambda_e T_d)p_j + \lambda_e T_d q_{ij}\right), \quad \text{(D1)}$$

where $q_{ij} = p_i^2\delta_{ij} + (1+p_i)p_j(1-\delta_{ij})$, δ being the Kronecker delta. An unbiased estimator for the four theoretical probabilities p_i, $i = 1,2,3,4$, can be constructed via the maximum likelihood principle. Given a realization of the Markov chain described above, i.e., a sequence of outcomes $\{x_i\}_{i=1,\dots,n}$, with $x_i = 1,2,3,4$, its probability is given by $p_{x_1}\prod_{i=1}^{n-1} P_{x_i x_{i+1}}$ and the corresponding log-likelihood is equal to

$$\begin{aligned}l(P) := \log(p_{x_1}\prod_{i=1}^{n-1} P_{x_i x_{i+1}}) = \log(p_i)\\ + \sum_{i,j=1,2,3,4} N_{ij}\log(P_{ij}), \qquad \text{(D2)}\end{aligned}$$

where N_{ij} is the number of transitions from i to j. The function $l(P)$ has to be maximized under the four constraints $\sum_j P_{ij} = 1,\ i = 1,2,3,4$, providing unbiased estimators $\hat{p}_i$ for the probabilities p_i, $i = 1,2,3,4$. The corresponding confidence intervals are evaluated according to the techniques presented in [44].

In fact, the estimated values $\hat{p}_i$ as well as the corresponding confidence intervals are computed using a code written using the statistical programming language R.

The Markovian model plays an important role also in the estimation of the guessing probability present in the sequence of raw data. Indeed, the semi-device-independent entropy certification protocol described in the main text yields a uniform bound for the theoretical detection probabilities (C3). In particular, setting $p_{\max} := (1/2) + (1/2)\sqrt{2 - (|\hat{I}_{\rm real}| - e_I)^2/4 + e_P}$ we have $p_j \le p_{\max}$ for any choice of the pair ϕ, θ and for any detection channel $j = 1, \dots, 4$.

On the other hand, taking into account the nonidealities of the detectors and the resultant correlations between subsequent photon detections, the effective guessing probability present in the sequence of the raw data has to be estimated as the maximum value of the probability of the nth readout given the previous ones:

$$p^*_{\rm guess} := \sup_{i_0,\dots,i_n} \mathbb{P}(\eta_n = i_n|\eta_{n-1} = i_{n-1}, \dots \eta_0 = i_0). \quad \text{(D3)}$$

By the Markov property, this reduces to

$$\begin{aligned}p^*_{\rm guess} &= \sup_{i,j} P_{ij}\\&= \sup_{i,j}\left(p_a\delta_{ij} + (1-p_a)\left((1-\lambda_e T_d)p_j + \lambda_e T_d q_{ij}\right)\right).\end{aligned}$$

By using the inequality $\sup_j p_j \le p_{\max}$ we finally get $p^*_{\rm guess} \le M(p_{\max})$, where the function M is defined as

$$\begin{aligned}M(p_{\max}) := \sup_{\substack{i,j=1,\dots,4\\ p_j\le p_{\max}}} &\left(p_a\delta_{ij} + (1-p_a)\left((1-\lambda_e T_d)p_j\right.\right.\\&\left.\left. + \lambda_e T_d q_{ij}\right)\right) = \max\{p_a + (1-p_a)\\&\times\left((1-\lambda_e T_d)p_{\max} + \lambda_e T_d p^2_{\max}\right), (1-p_a)\\&\times(p_{\max} + \lambda_e T_d p_{\max}(1-p_{\max}))\}.\end{aligned}$$

APPENDIX E: METHODS

The measurements are performed using as a light source an attenuated single-mode continuous wave green He:Ne laser, emitting at 543.5 nm, with nominal output power of 4 mW. The laser is fiber-coupled and attenuated by a variable optical attenuator before entering the experimental setup. The phase ξ and ϕ are controlled by using two piezoelectric transducer actuated mirrors with feedback loop to ensure time stability. Four lenses collect the four PVMs and fiber-couple them to four different SPADs, whose efficiencies have been equalized previously using four variable optical attenuators, which compensate also for different losses due to fibers and lenses. For each experimental realization $\{(\phi_x, \theta_y)\}_{x,y=0,1}$, the observation windows of 50 s are measured using time bins of 1 μs. The time bin of 1 μs is chosen mainly for two reasons. First, 1 μs is smaller than the typical delay time between two subsequent detection events ($(1/\lambda_e) \simeq 5\ \mu$s), such that within one time bin multi-photon occurrences are of the order of only 10% of the raw data. Second, 1 μs time bin allows us to have a length of the time sequence which can be stored in a standard PC memory. Time bins with multiple detection events or no detection at all are discarded from the collected string of raw events. Except for the laser, the whole setup is optically shielded from the rest of the lab by enclosure in a black box. In addition, each SPAD is optically shielded and physically separated from the three others by means of opaque boxes, in order to avoid any cross-talk in the measurements.

[1] M. Herrero-Collantes and J. C. Garcia-Escartin, Quantum random number generators, Rev. Mod. Phys. **89**, 015004 (2017).

[2] R. Konig, R. Renner, and C. Schaffner, The operational meaning of min- and max-entropy, IEEE Trans. Inf. Theory **55**, 4337 (2009).

[3] N. Nisan and A. Ta-Shma, Extracting randomness: A survey and new constructions, J. Comput. Syst. Sci. **58**, 148 (1999).

[4] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, Quantum random number generation, Npj Quantum Inf. **2**, 16021 (2016).

[5] D. Frauchiger, R. Renner, and M. Troyer, True randomness from realistic quantum devices, arXiv:1311.4547 (2013).

[6] A. Acín and L. Masanes, Certified randomness in quantum physics, Nature **540**, 213 (2016).

[7] S. Pironio, The certainty of quantum randomness (2018).

[8] J. Bell, *The Theory of Local Beables. Speakable and Unspeakable in Quantum Mechanics* (Springer, Cambridge, United Kingdom, 1974).

[9] 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 (1969).

[10] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, and T. A. Manning *et al.*, Random numbers certified by Bell's theorem, Nature **464**, 1021 (2010).

[11] A. Acín, S. Massar, and S. Pironio, Randomness versus nonlocality and entanglement, Phys. Rev. Lett. **108**, 100402 (2012).

[12] S. Pironio and S. Massar, Security of practical private randomness generation, Phys. Rev. A **87**, 012336 (2013).

[13] L. Shen, J. Lee, L. P. Thinh, J.-D. Bancal, A. Cerè, A. Lamas-Linares, A. Lita, T. Gerrits, S. W. Nam, V. Scarani, and C. Kurtsiefer, Randomness extraction from Bell violation with continuous parametric down-conversion, Phys. Rev. Lett. **121**, 150402 (2018).

[14] Y. Liu, X. Yuan, M.-H. Li, W. Zhang, Q. Zhao, J. Zhong, Y. Cao, Y.-H. Li, L.-K. Chen, H. Li, T. Peng, Y.-A. Chen, C.-Z. Peng, S.-C. Shi, Z. Wang, L. You, X. Ma, J. Fan, Q. Zhang, and J.-W. Pan, High-speed device-independent quantum random number generation without a detection loophole, Phys. Rev. Lett. **120**, 010503 (2018).

[15] P. Bierhorst, E. Knill, S. Glancy, Y. Zhang, A. Mink, S. Jordan, A. Rommal, Y.-K. Liu, B. Christensen, S. W. Nam, M. J. Stevens, and L. K. Shalm, Experimentally generated randomness certified by the impossibility of superluminal signals, Nature **556**, 223 (2018).

[16] W.-Z. Liu, M.-H. Li, S. Ragy, S.-R. Zhao, B. Bai, Y. Liu, P. J. Brown, J. Zhang, R. Colbeck, J. Fan, Q. Zhang, and J.-W. Pan, Device-independent randomness expansion against quantum side information, Nat. Phys. **17**, 448 (2021).

[17] L. K. Shalm, Y. Zhang, J. C. Bienfang, C. Schlager, M. J. Stevens, M. D. Mazurek, C. Abellán, W. Amaya, M. W. Mitchell, M. A. Alhejji, H. Fu, J. Ornstein, R. P. Mirin, S. W. Nam, and E. Knill, Device-independent randomness expansion with entangled photons, Nat. Phys. **17**, 452 (2021).

[18] Z. Cao, H. Zhou, X. Yuan, and X. Ma, Source-independent quantum random number generation, Phys. Rev. X **6**, 011020 (2016).

[19] D. G. Marangon, G. Vallone, and P. Villoresi, Source-device-independent ultrafast quantum random number generation, Phys. Rev. Lett. **118**, 060503 (2017).

[20] M. Avesani, D. G. Marangon, G. Vallone, and P. Villoresi, Source-device-independent heterodyne-based quantum random number generator at 17 Gbps, Nat. Commun. **9**, 5365 (2018).

[21] T. Michel, J. Y. Haw, D. G. Marangon, O. Thearle, G. Vallone, P. Villoresi, P. K. Lam, and S. M. Assad, Real-time source-independent quantum random-number generator with squeezed states, Phys. Rev. Appl. **12**, 034017 (2019).

[22] G. Vallone, D. G. Marangon, M. Tomasin, and P. Villoresi, Quantum randomness certified by the uncertainty principle, Phys. Rev. A **90**, 052327 (2014).

[23] Z. Cao, H. Zhou, and X. Ma, Loss-tolerant measurement-device-independent quantum random number generation, New J. Phys. **17**, 125011 (2015).

[24] Y.-Q. Nie, J.-Y. Guan, H. Zhou, Q. Zhang, X. Ma, J. Zhang, and J.-W. Pan, Experimental measurement-device-independent quantum random-number generation, Phys. Rev. A **94**, 060301(R) (2016).

[25] T. Van Himbeeck, E. Woodhead, N. J. Cerf, R. García-Patrón, and S. Pironio, Semi-device-independent framework based on natural physical assumptions, Quantum **1**, 33 (2017).

[26] D. Rusca, T. van Himbeeck, A. Martin, J. B. Brask, W. Shi, S. Pironio, N. Brunner, and H. Zbinden, Self-testing quantum random-number generator based on an energy bound, Phys. Rev. A **100**, 062338 (2019).

[27] T. Van Himbeeck and S. Pironio, Correlations and randomness generation based on energy constraints, arXiv preprint arXiv:1905.09117 (2019).

[28] M. Avesani, H. Tebyanian, P. Villoresi, and G. Vallone, Semi-device-independent heterodyne-based quantum random number generator, arXiv:2004.08344 (2020).

[29] D. Rusca, H. Tebyanian, A. Martin, and H. Zbinden, Fast self-testing quantum random number generator based on homodyne detection, Appl. Phys. Lett. **116**, 264004 (2020).

[30] J. B. Brask, A. Martin, W. Esposito, R. Houlmann, J. Bowles, H. Zbinden, and N. Brunner, Megahertz-rate semi-device-independent quantum random number generators based on unambiguous state discrimination, Phys. Rev. Appl. **7**, 054018 (2017).

[31] N. Leone, D. Rusca, S. Azzini, G. Fontana, F. Acerbi, A. Gola, A. Tontini, N. Massari, H. Zbinden, and L. Pavesi, An optical chip for self-testing quantum random number generation, APL Photonics **5**, 101301 (2020).

[32] H.-W. Li, Z.-Q. Yin, Y.-C. Wu, X.-B. Zou, S. Wang, W. Chen, G.-C. Guo, and Z.-F. Han, Semi-device-independent random-number expansion without entanglement, Phys. Rev. A **84**, 034301 (2011).

[33] T. Lunghi, J. B. Brask, C. C. W. Lim, Q. Lavigne, J. Bowles, A. Martin, H. Zbinden, and N. Brunner, Self-testing quantum random number generator, Phys. Rev. Lett. **114**, 150501 (2015).

[34] D. H. Smith, G. Gillett, M. P. de Almeida, C. Branciard, A. Fedrizzi, T. J. Weinhold, A. Lita, B. Calkins, T. Gerrits, H. M. Wiseman, S. W. Nam, and A. G. White, Conclusive quantum steering with superconducting transition-edge sensors, Nat. Commun. **3**, 625 (2012).

[35] A. A. Abbott, C. S. Calude, and K. Svozil, A quantum random number generator certified by value indefiniteness, Math. Struct. Comput. Sci. **24**, e240303 (2014).

[36] S. Azzini, S. Mazzucchi, V. Moretti, D. Pastorello, and L. Pavesi, Single-particle entanglement, Adv. Quantum Technol. **3**, 2000014 (2020).

[37] P. Saha and D. Sarkar, Robustness measure of hybrid intra-particle entanglement, discord, and classical correlation with initial Werner state, Quantum Inf. Proc. **15**, 791 (2016).

[38] M. Pasini, N. Leone, S. Mazzucchi, V. Moretti, D. Pastorello, and L. Pavesi, Bell-inequality violation by entangled single-photon states generated from a laser, an LED, or a halogen lamp, Phys. Rev. A **102**, 063708 (2020).

[39] V. Moretti, *Fundamental Mathematical Structures of Quantum Theory* (Springer, Printforce, the Netherlands, 2019), Chap. 5.

[40] X. Ma, F. Xu, H. Xu, X. Tan, B. Qi, and H.-K. Lo, Postprocessing for quantum random-number generators: Entropy evaluation and randomness extraction, Phys. Rev. A **87**, 062327 (2013).

[41] S. Mazzucchi, N. Leone, S. Azzini, L. Pavesi, and V. Moretti, Entropy certification of a realistic quantum random-number generator based on single-particle entanglement, Phys. Rev. A **104**, 022416 (2021).

[42] To simplify the notation, we replace $\mathbf{u}$ with the momentum rotation angle ϕ, and $\mathbf{v}$ with the polarization rotation angle θ.

[43] G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins University Press, Baltimore, Maryland, 1996), 3rd ed.

[44] P. Billingsley, Statistical methods in Markov chains, Ann. Math. Stat. **32**, 12 (1961).