Semi-device-independent randomness from *d*-outcome continuous-variable detection

Hamid Tebyanian,[1] Marco Avesani,[1] Giuseppe Vallone,[1,2,3] and Paolo Villoresi[1,3]
[1]*Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo 6B, 35131 Padova, Italy*
[2]*Dipartimento di Fisica e Astronomia, Università di Padova, via Marzolo 8, 35131 Padova, Italy*
[3]*Istituto di Fotonica e Nanotecnologie, CNR, Via Trasea 7, 35131 Padova, Italy*

(Received 12 July 2021; accepted 22 November 2021; published 14 December 2021)

Recently, semi-device-independent protocols have attracted increasing attention, guaranteeing security with few hypotheses and experimental simplicity. In this paper, we demonstrate a many-outcome scheme with binary phase-shift keying (BPSK) for a semi-device-independent protocol based on the energy assumption. We show in theory that the number of certified random bits of the *d*-outcome system outperforms the standard scheme (binary outcomes). Furthermore, we compare the results of two well-known measurement schemes, homodyne detection and heterodyne detection. Taking into account the experimental imperfections, we discuss the experimental feasibility of the *d*-outcome design, and finally, we experimentally validate this approach with an experiment based on BPSK modulation and heterodyne detection.

DOI: 10.1103/PhysRevA.104.062424

I. INTRODUCTION

In the information security age, data privacy and secure communication are of paramount relevance. It is worth stressing the role of genuine random numbers for privacy and security applications. Nearly all of the protocols dealing with privacy and security rely on random numbers, and a protocol's security is directly connected to the quality of the employed random numbers [1]. Thus, owning certified random numbers is a critical component for guarding the information. Pseudo-random-number generators have been popular and widely used in the past few decades. However, the generated numbers are not truly random since the randomness source is based upon a classical phenomenon that is deterministic. In general, random-number generators (RNGs) can be classified into two major groups, classical and quantum. Due to their determinism, classical RNGs cannot offer high levels of security, while quantum random-number generators (QRNGs) are qualified candidates for generating genuine and unpredictable random numbers based on the intrinsic randomness of quantum mechanics [2].

Despite the fact that quantum mechanics ensures the unpredictability of the generated random numbers, experimental imperfections of QRNGs can open a back door for eavesdroppers to attack or manipulate the protocol [3]. For instance, the generator's apparatus can be correlated with an external party or can deviate from the expected behavior. Hence, QRNGs can be categorized into three subgroups, trusted-device, semi-device-independent (semi-DI), and device-independent (DI) QRNGs [4]. Although the trusted-device QRNGs are cheap, fast, and more reliable than the classical generators, they can be compromised due to the security loopholes resulting from trusting the devices. On the other hand, the highest security is achievable by DI QRNGs in which randomness is certified by the violation of a Bell inequality, without any trust on any devices [5].

Besides offering highly secure randomness, DI-QRNG protocols are also robust against experimental imperfections. Unfortunately, the experimental realization of a loophole-free Bell test is extremely hard to accomplish, and only proof-of-principle experiments have been realized, obtaining modest generation rates [6–10]. Taking into account the complexity of this protocol and the low bit rate, DI QRNGs are still very far from being practical. Indeed, security and speed are the two key features of RNGs, and both are needed in practical applications.

Semi-DI protocols are an intermediate approach between DI and trusted-device schemes, which offer an optimal trade-off between generation rate, security, and ease of implementation [4]. Depending on the protocol needs, assumptions can vary; for very secure protocols, there are fewer assumptions on the device, i.e., a single assumption on the overlap or energy of the prepared states [11–17], assumptions related to the dimension of the Hilbert space [18,19], and the requirement of trusted measurement in the case of source-DI protocols [20–23] or a trusted source and in measurement-DI protocols [24,25]. Recently, a new class of protocols was proposed, in which both the source and measurement are untrusted and only a single assumption on the overlap or energy of the prepared states is required [11–16]. These protocols can provide increased security since they reduce the number of assumptions on the devices.

In this work, we investigate the impact of increasing the number of outcomes of the measurement apparatus given a binary-input semi-DI QRNG implemented with optical continuous variables (CVs) [15,16]. The protocol builds upon the prepare-and-measure scheme, with a measurable condition on the maximum energy of the prepared states that implies a lower bound on the state's overlap. The use of CVs allows for high generation rates. Indeed, discrete-variable (DV) implementations have the problem of single-photon

detector saturation (megahertz bandwidth), while CV-detection schemes could employ high-speed detectors with gigahertz bandwidth. Therefore, while the generated bit per measurement could be higher with DV implementations, the absolute generation rate (bit/second) could be remarkably higher with the CV detection schemes. This is the reason why considering CV encoding can improve the absolute generation rate.

The main contribution of our work is the demonstration that by increasing the number of outcomes in the postprocessing stage, i.e., without changing the experimental setup, it is possible to improve the generation rate for two well-known CV-detection schemes, homodyne and heterodyne. At first sight, our result seems to be in contradiction to the results reported in [26]. In [26] it was demonstrated that, for a semi-DI QRNG with n inputs subjected to the overlap bound, no more than $\log_2(n+1)$ random bits can be certified and the measurement apparatus achieving the maximum randomness is obtained by using an $(n+1)$-outcome positive operator-valued measure (POVM). So it could be expected that with two inputs, no more than three outputs should be considered. However, the optimal POVM cannot be easily implemented with standard CV components, and we will show that, if the measurement is realized by using a homodyne or heterodyne detector, increasing the number of outputs to more than three (for two inputs) will improve the generation rate.

In particular, we report the numerical results of the method employed for randomness estimation, from 3 to 14 outcomes, concerning both homodyne and heterodyne detections and then compare them with the binary-outcome result. We will also investigate the generation rate as a function of the efficiency of the used measurement device, showing that the advantage of increasing the number of outcomes decreases with lower efficiency.

II. SEMI-DI-QRNG MODEL

A. Randomness certification framework

The protocol is based on two devices: an untrusted measurement and a partially trusted preparation station. For the latter, we perform a single assumption corresponding to an upper bound on the prepared state's energy. Similar to [11,13,16], we consider here the case in which the preparation and measurement devices cannot share quantum correlation, while they can be correlated classically.

The scheme of this protocol is shown in Fig. 1: a preparation device emits the unknown states ρ_x after receiving the binary input $x \in \{0, 1\}$ from the user. The measurement device has d outputs $b \in \{0, 1, \ldots, d-1\}$. By running the experiment N times it is possible to estimate the conditional probabilities $p(b|x)$.

The measurement device is considered a black box, whose internal working principles are unknown to the user. The preparation device is a "gray box": the internal working principles are unknown, although we assume there are no correlations between the preparation device and any external devices. Moreover, we assume that the prepared states are identically and independently distributed (IID hypothesis). The single experimentally verifiable condition on the preparation section is an upper bound on the energy of the prepared states:

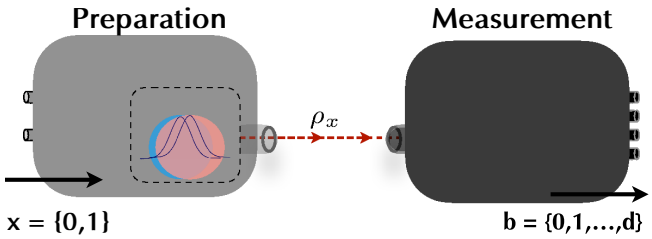

FIG. 1. The general design of the QRNG protocol. Depending on the input x, the unknown state ρ_0 or ρ_1 is transmitted from the preparation part. A single assumption is present on the state's energy. The measurement device, with no assumptions, performs a generic measurement and outputs $b \in \{0, \ldots, d-1\}$.

$$\langle \hat{n} \rangle_{\rho_x} \leqslant \mu. \tag{1}$$

The above condition can easily be checked by using a power meter on the prepared states.

As shown in [27], the conditional min-entropy, namely, the amount of genuine random bits per measurement run, is given by

$$H_{\min} = -\log_2 (P_g), \tag{2}$$

where P_g is the guessing probability, namely, the highest probability that an attacker knowing the internal working principle of the devices can guess the outcomes b, given the input x. We note that, without loss of generality, we can assume that the source generates pure states since mixed states do not provide any advantage to an attacker and, indeed, will lower the guessing probability of an adversary.

It is worth noting that the bound on the energy, whose validity can be checked experimentally, implies a lower bound on the scalar product between the emitted states [13,16], and thus, the approach of [11] can be followed to obtain P_g from the experimental data.

By generalizing the approach of [11] with d outcomes, P_g can be found as the solution of the following semidefinite programming (SDP):

$$\begin{aligned}
&\underset{M_b^{\lambda_0,\lambda_1}}{\text{maximize}}\ \widetilde{P}_g = \frac{1}{2}\sum_{x=0}^{1}\sum_{\lambda_0,\lambda_1=0}^{d-1} \langle \psi_x | M_{\lambda_x}^{\lambda_0,\lambda_1} | \psi_x \rangle \\
&\text{subject to } M_b^{\lambda_0,\lambda_1} = \left(M_b^{\lambda_0,\lambda_1}\right)^\dagger, \\
&\qquad M_b^{\lambda_0,\lambda_1} \geqslant 0, \\
&\qquad \sum_{b=0}^{d-1} M_b^{\lambda_0,\lambda_1} = \frac{1}{2}\mathrm{Tr}[\sum_{b=0}^{d-1} M_b^{\lambda_0,\lambda_1}]\mathbb{1}, \\
&\qquad \sum_{\lambda_0,\lambda_1} \langle \psi_x | M_b^{\lambda_0,\lambda_1} | \psi_x \rangle = p(b|x)\, \forall\, b, x,
\end{aligned} \tag{3}$$

where $M_b^{\lambda_0,\lambda_1}$ are 2×2 operators in the two-dimensional Hilbert space spanned by the orthonormal vectors $|0\rangle$ and $|1\rangle$ and the states $|\psi_x\rangle$ are defined by

$$\begin{aligned}
|\psi_0\rangle &= |0\rangle, \\
|\psi_1\rangle &= (1-2\mu)|0\rangle + 2\sqrt{\mu(1-\mu)}|1\rangle.
\end{aligned} \tag{4}$$

The above states $|\psi_x\rangle$ saturate the bound $|\langle\psi_0|\psi_1\rangle| \geqslant 1-2\mu$ derived from (1) and can be used in the optimization without loss of generality (see [12,16]). In Eq. (3) we assumed that the input states are prepared with equal probability, namely, $p_x = 1/2$.

The variables $\lambda \equiv (\lambda_0, \lambda_1)$ represent the classical information available to anyone knowing the internal workings of the device. The operators $M_b^{\lambda_0,\lambda_1}$ are related to possible physical realizations of the measurement device that are compatible with the observed probabilities $p(b|x)$. More precisely, for each value of the pair $(\lambda_0\,\lambda_1)$, the value $q_\lambda = \frac{1}{2}\mathrm{Tr}[\sum_b M_b^{\lambda_0,\lambda_1}]$ represents the probability that the measurement device is actually implementing the POVM defined by the operators $\{\Pi_b^{\lambda_0,\lambda_1}\}$, where $\Pi_b^{\lambda_0,\lambda_1} = M_b^{\lambda_0,\lambda_1}/q_\lambda$.

It is worth noticing that the above approach is general and does not depend on the actual implementation of the preparation and measurement devices. The min-entropy is directly calculated by using only the value of the energy bound μ and the measured output probabilities $p(b|x)$, independent of their physical realization. We observe that larger $H_{\min}$ can be obtained whenever the probabilities $p(b|x)$ allow us to better distinguish the two input states.

B. Implementation with continuous variables

We now illustrate the amount of randomness that can be obtained by using single-mode optical continuous variables defined by the creation operator $\hat{a}^\dagger$.

1. Preparation

In the preparation part, we employed the binary phase-shift keying (BPSK) system, in which the source, a continuous-wave (CW) laser, emits two coherent states with the same mean photon number and a π phase shift $|\psi_0\rangle = |\alpha\rangle$ and $|\psi_1\rangle = |-\alpha\rangle$. We can use the representation of a coherent state in Fock space to define $|\alpha\rangle$ as $|\pm\alpha\rangle = e^{-\frac{\mu}{2}}\sum_{n=0}^{\infty}\frac{(\pm\sqrt{\mu}e^{i\phi})^n}{\sqrt{n!}}|n\rangle$, where $\alpha = \sqrt{\mu}e^{i\phi}$, μ is the mean photon number, and ϕ is the relative phase between the signal and the local oscillator (LO). We here assume that the LO is chosen such that $\phi = 0$. Note that the input x should be uncorrelated with λ and independent of the devices. Thus, they can be generated from a standard RNG (e.g., pseudo-RNG). We note that the mean photon number for each state $|\psi\rangle$ is upper bounded by the quantity μ given in Eq. (1). We note that states with nonvanishing overlap cannot be deterministically distinguished, unlike orthogonal states.

2. Measurement

Homodyne and heterodyne tomographies are two primary and well-established detection schemes for measuring CV states of light (see Fig. 2). By homodyning, the quantum state is measured from samples obtained from projected Wigner functions, whereas heterodyne detection directly samples phase-space coordinates from the Husimi Q function [28,29]. In regard to semi-DI QRNG protocols, both heterodyne and homodyne detections have been employed on the receiver side, as shown in [16] and [15], respectively. In these works, the (potentially) infinite outcomes of the CV measurement

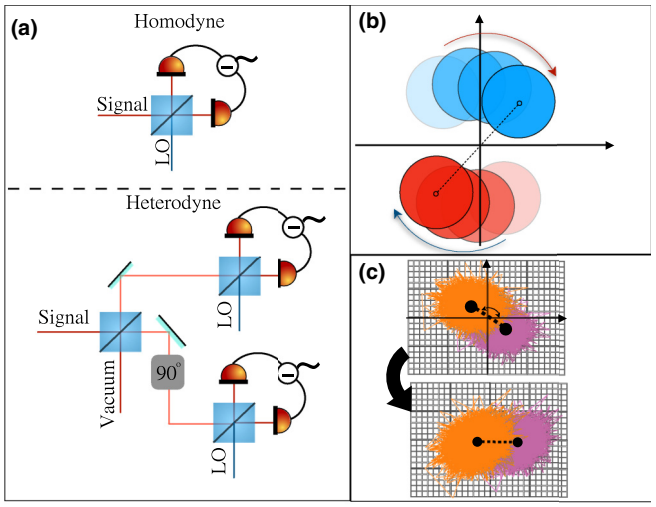

FIG. 2. Homodyne and heterodyne detections. (a) Representation of the two detection schemes. (b) Effects of the phase instability on the received states. (c) Off-line phase compensation for heterodyne detection.

were grouped into two disjoint sets, corresponding to a binary outcome. Here we consider the more general case in which the physical outcomes can be grouped into a larger number of sets.

The POVMs of homodyne and heterodyne receivers can be represented, respectively, by

$$\Pi^{(\mathrm{hom})}(X) = |X\rangle\langle X|,$$
$$\Pi^{(\mathrm{het})}(\beta) = \frac{1}{\pi}|\beta\rangle\langle\beta|, \tag{5}$$

where $|X\rangle$ is the eigenstate of the $\hat{X} = (\hat{a}+\hat{a}^\dagger)/\sqrt{2}$ operator and $|\beta\rangle$ is the coherent state with complex amplitude β.

The corresponding probability densities associated with the measurement of the states $|\pm\sqrt{\mu}\rangle$ are given by

$$\mathcal{P}_\pm^{(\mathrm{hom})}(X) = \sqrt{\frac{2}{\pi}}e^{-2(X\mp\sqrt{\eta\mu})^2},$$
$$\mathcal{P}_\pm^{(\mathrm{het})}(\beta) = \frac{1}{\pi}e^{-(X\mp\sqrt{\eta\mu})^2}e^{-Y^2}, \tag{6}$$

with real X and Y and $\beta = X + iY$. In the above equations we included the overall efficiency η of the channel and of the receiver devices. In order to obtain d possible outcomes $b = 0, 1, \ldots, d-1$ we need to partition the real line X or the phase space β into d disjoint sets.

In the homodyne case, it is necessary to choose $d-1$ increasing real numbers $X_1 < X_2 < \cdots < X_{d-1}$ such that the outcome probabilities for $b = 0, \ldots, d-1$ can be written as

$$p^{(\mathrm{hom})}(b|x) = \frac{1}{\sqrt{\pi}}\int_{X_b}^{X_{b+1}} e^{-[X-(-1)^x\sqrt{2\eta\mu}]^2}dX$$
$$= \frac{1}{2}\{\mathrm{erf}[X_{b+1} - (-1)^x\sqrt{2\eta\mu}]$$
$$- \mathrm{erf}[X_b - (-1)^x\sqrt{2\eta\mu}]\}, \tag{7}$$

TABLE I. Definition of the partitions of the real axis corresponding to different output configurations for the homodyne detection.

d	X_0	X_1	X_2	X_3	X_4	X_5	X_6
2	$-\infty$	0	$+\infty$				
3	$-\infty$	$-L_1$	$+L_1$	$+\infty$			
4	$-\infty$	$-L_1$	0	$+L_1$	$+\infty$		
6	$-\infty$	$-L_2$	$-L_1$	0	$+L_1$	$+L_2$	$+\infty$

with the convention that $X_0 = -\infty$ and $X_d = +\infty$. We note that from Eq. (6) to Eq. (7) we have performed a change of the integration variable.

In the heterodyne case, we may define a partition of the phase space $\{\Lambda_b\}$ with d elements. The output probabilities can be written as

$$p^{(\text{het})}(b|x) = \frac{1}{\pi}\int_{\Lambda_b} e^{-[X-(-1)^x\sqrt{\eta\mu}]^2} e^{-Y^2} dX dY$$
$$= \frac{e^{-\eta\mu}}{\pi}\int_{\Lambda_b} re^{-r^2+2r(-1)^x\sqrt{\eta\mu}\cos\theta} dr d\theta. \quad (8)$$

In the following we will analyze the achievable randomness by considering the above measurements. We will consider cases with an increasing number of outcomes, and we compare them with the results obtained with two outcomes already reported in [15,16].

III. RESULTS

A. Homodyne detection

We start by considering the homodyne detection with perfect efficiency ($\eta = 1$). Due to the symmetry of the prepared states, the partition of the real axis is optimal when it is symmetric around the origin. For instance, the configurations corresponding to two, three, four, and six outcomes are shown in Table I and are illustrated in Fig. 3.

The number of extractable genuine random bits is estimated by numerically solving the dual of the SDP optimization problem given by Eq. (3), constrained by the conditional probabilities $p^{\text{hom}}(b|x)$, obtained from Eq. (7), together with the energy-bound assumption μ. The results are further optimized over the values L_k. The values of the min-entropy as a function of the energy bound μ are shown in Fig. 4 for the two-, four-, and six-outcome cases.

As shown in Fig. 4, by increasing the measurement outcomes, the min-entropy monotonically increases over the entire range of μ, meaning that more randomness can be certified. It is worth noting that, starting from the same physical implementation (homodyne measurement) and changing the postprocessing (namely, by changing the partitions of the outcomes) different values of the min-entropy can be obtained.

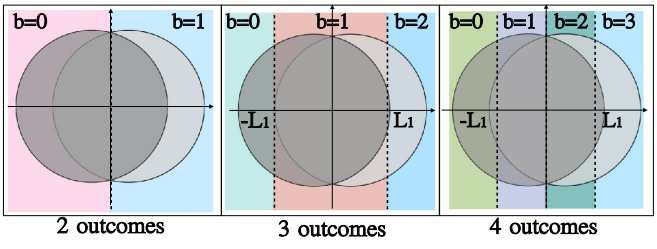

FIG. 3. Homodyne measurement configurations.

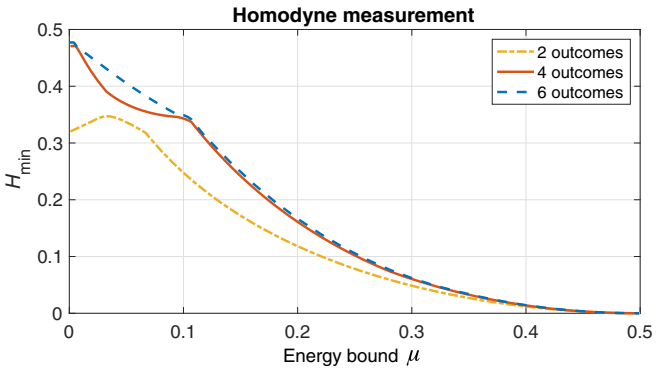

FIG. 4. Min-entropy as a function of the energy bound μ for homodyne detection and different numbers of outcomes.

One could ask what happens with a further increase in the number of outcomes. As shown in Fig. 5, improvements are obtained for small values of μ by increasing the number of outcomes up to 14. In Fig. 6 the best min-entropy (with optimized μ) is shown as a function of the number of outcomes. The data suggest that larger min-entropy will be obtained by further increasing the number of outcomes toward a seemingly asymptotic value of 0.5. However, to rigorously prove the above statement, further analysis, beyond the scope of the present paper, will be needed. To better clarify the d-outcome-approach improvement, we can consider an ideal (noiseless) implementation with a 1.25-GHz repetition rate: in this case the generation rate would be approximately 437 MHz for the standard method (binary outcome). In contrast, with the d-outcome approach, one can improve the generation rate up to 625 MHz without any changes in the experimental setup.

We now present the results obtained with an inefficient system, namely, by considering $\eta < 1$. The parameter η is used to model the effect of different experimental imperfections, such as the losses of the channel, the limited efficiency of the receiver's detectors, and the electronic noise of the detection apparatus. We carried out the same analysis described above by considering different values of η. We show in Fig. 7 the min-entropy as a function of μ for different values of η and for two and four outcomes. The corresponding optimal value

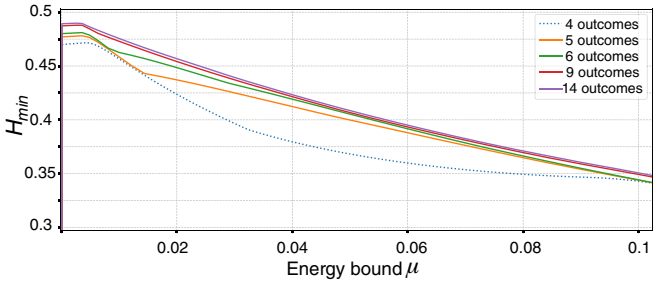

FIG. 5. Min-entropy for a large number of outcomes plotted for small μ values. It should be noted that the curves correspond to the legend from bottom to top.

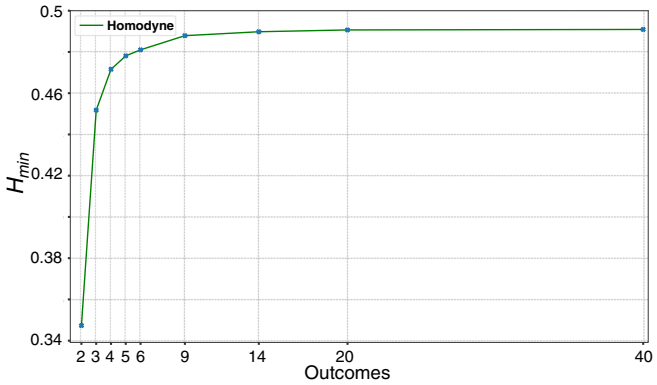

FIG. 6. Maximum min-entropy (with optimized μ and $\eta = 1$) for the different numbers of outcomes for homodyne detection. Note that the lines between the points do not return the min-entropy at noninteger numbers.

of L_1 for the four-outcome case is shown in Fig. 8. From the figures, it can be seen that when the efficiency decreases, the advantage of using more outcomes is less evident, but it is still present.

We note that the "jumps" in the curves for L_1 are due to a "change" in the optimal strategy depending on the value of μ. For each choice of $M_b^{\lambda_0,\lambda_1}$, the probability that a given strategy labeled by (λ_0, λ_1) is used is given by $q_\lambda = \frac{1}{2}\mathrm{Tr}[\sum_{b=0}^{d-1} M_b^{\lambda_0,\lambda_1}]$. A change in strategy means that the optimal values of $q^{(\lambda_0,\lambda_1)}$ have a discontinuity, namely, an abrupt change between two close values of μ. We note that jumps in L_1 correspond to singular points also for $H_{\min}$, in which the derivative of $H_{\min}$ is not defined. This is a clear indication of a change in the optimal strategy. We also underline that such singular points in $H_{\min}$ are also present in the two-outcome results, in which no optimization over L_1 is present.

B. Heterodyne detection

Homodyne detection is sensitive to only one field quadrature, e.g., X_ϕ sampling only a projection of the phase space.

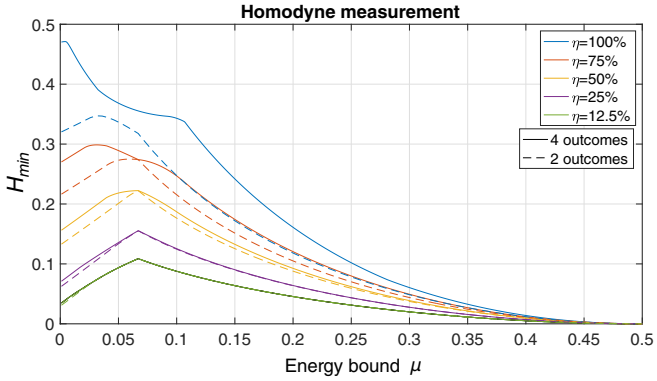

FIG. 7. Min-entropy as a function of the energy bound μ for the homodyne detector. We compared the two-outcome (dashed lines) and four-outcome (solid lines) schemes for different values of the efficiency η. It should be noted that the curves correspond to the legend from top to bottom.

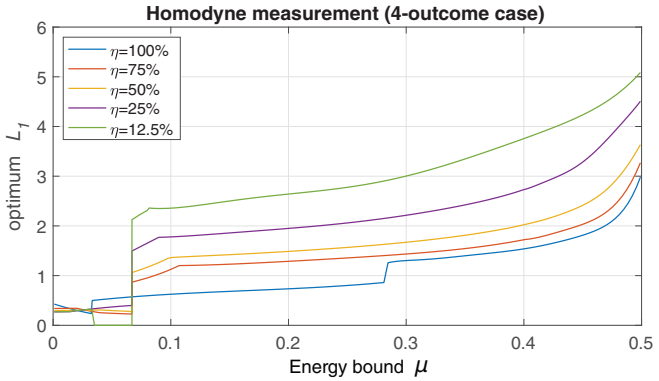

FIG. 8. Optimal value of L_1 for the symmetric four-outcome configuration for different system efficiencies η. It should be noted that the curves correspond to the legend from bottom to top.

Heterodyne detection, on the other hand, performs a joint "noisy" measurement of two conjugated field quadratures, $\tilde{X}_\phi$ and $\tilde{P}_\phi$, thus sampling the entire phase space. The number of possible (and potentially optimal) partitions for heterodyne detection is larger than for homodyne detection due to the increased dimensionality of the measurement.

Like in the homodyne case, it is possible to choose a "strip" partition, namely, the configuration illustrated in Fig. 3: the phase space is subdivided in vertical strips whose boundaries are defined by the increasing real numbers $X_1 < X_2 < \cdots < X_{d-1}$. Looking at Eqs. (7) and (8) it is possible to note that a heterodyne measurement with this configuration and efficiency η is equivalent to the homodyne measurement with efficiency $\eta/2$. Thus, we can directly refer to Fig. 7 for the results.

Other possible configurations are displayed in Fig. 9. By running the SDP for all the configurations represented in Fig. 9, we obtained a min-entropy that is always lower than the one obtained with the configuration shown in Fig. 3.

IV. PRACTICAL CONSIDERATIONS

The main focus of this work is studying the influence of extending the number of outcomes on semi-DI QRNGs based on an energy bound and homodyne or heterodyne detection. We focused on homodyne and heterodyne detections because they are the most common measurement schemes employed in CV protocols. Moreover, recent experiments [15,16], employed these measurement schemes to implement energy-bounded semi-DI QRNG protocols. These works could benefit from this analysis, without any modifications to the experimental

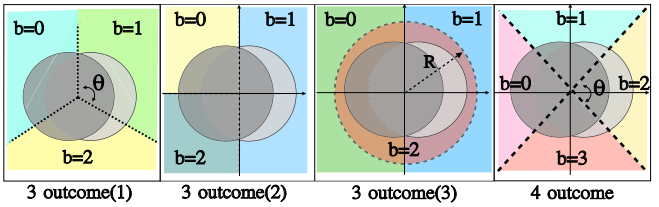

FIG. 9. Alternative partitions of the phase space for heterodyne measurement.

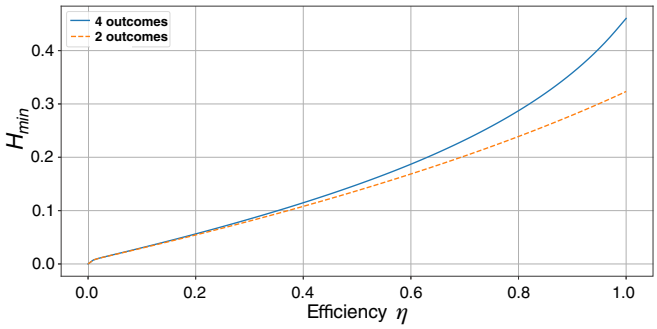

FIG. 10. Min-entropy as a function of efficiency η for the homodyne detection concerning the two- and four-outcome configurations. The mean photon number μ and range L are chosen in a way that the min-entropy is maximized.

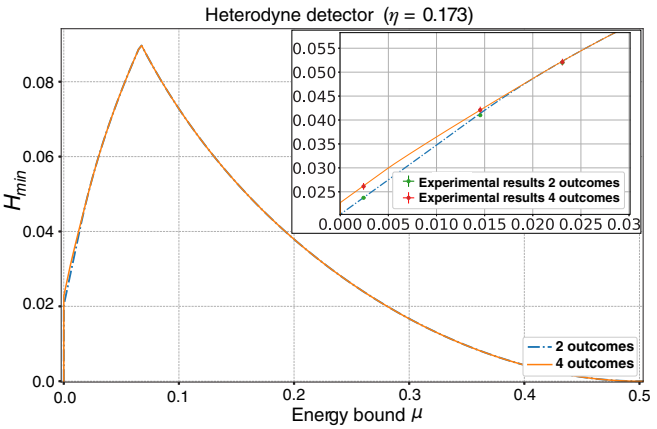

FIG. 11. The min-entropy as a function of the mean photon number. The dash-dotted blue and solid orange curves are the numerical predictions for the two- and four-outcome cases when the detector efficiency is $\eta = 0.173$. The green and red points, however, are the experimental data.

setup. In fact, the presented results show an enhancement of the certifiable min-entropy with respect to the binary case for ideal detection and no losses. However, we note that in practical implementations the expected improvement is reduced. In fact, additional losses, limited detector efficiency and excess noise of the receiver apparatus contribute to a reduction of the correlations $p(b|x)$, limiting the advantage of these schemes, as shown in Fig. 10.

We experimentally investigated the d-outcome approach by analyzing the data obtained in a previous experiment based on the BPSK modulation scheme and heterodyne detection reported in [16]. We also included some data sets that were not presented in the previous paper. We note that in order to implement the d-outcome approach, we need to adjust only the postprocessing stage rather than the experiment setup.

Since the experimental setup was already presented in [16], we refer to it for more experimental details, and we give below the main elements. The setup consists of three parts: source, receiver, and electronics. In the source, a CW laser emits light into a Mach-Zehnder interferometer, where the light is split into two parts: LO and signal. Based on the field-programmable gate array outputs (with 1.25-GHz repetition rate), a phase modulator applies either 0 or π phase shift to the signal: in this way, either $|\alpha\rangle$ or $|-\alpha\rangle$ is generated. Part of the signal is measured by using a beam splitter and a power meter, allowing real-time measurement of the signal energy: this measurement allows a real-time check of the energy assumption and constitutes the only trusted part of the whole experiment. The heterodyne detection then is implemented by means of a 90° optical hybrid followed by two balanced homodyne detectors. An oscilloscope with 4 GHz of analog bandwidth at a sampling rate of 12.5 Gsps and 8-bit resolution collects the data and later sends it to the computer for postprocessing. After phase recovery, obtained as discussed in [16], we applied the strip configuration (see Fig. 3) to the obtained data. Figure 11 presents the min-entropy as a function of μ for binary- and quartet-outcome cases. Since we used a heterodyne detector with 17.3% efficiency, we expect a small improvement using the d-outcome approach for low mean photon numbers μ. Indeed, as shown in the inset, there is a slight increase in the min-entropy for the four outcomes compared to two outcomes when μ is lower than 0.02. Green and red dots represent the experimental data for two- and four-outcome cases, and they are in perfect agreement with the theoretical predictions. The results show that considering the approach presented here, it is possible to achieve a higher generation rate from the same experimental data. Despite the fact that the improvement is small for our experiment with 17.3% detection efficiency, this analysis validates the d-outcome approach.

It is worth noting that the value of L_1 is chosen *a priori*, depending on the value of the assumed bound μ. While in our calculation we assumed that the output of the heterodyne measurement is continuous, any experimental realization in fact implements a discretization of it. The latter effectively implies a discretization of the possible values of L_1. However, as shown in Fig. 11, the discretization has a negligible impact on the results since the experimental points (obtained with the discretized heterodyne measurement) are very well modeled by the continuous heterodyne measurement given in Eqs. (6) and (8). Moreover, it can be checked that small deviations from the optimal L_1 values have a small impact on the value of $H_{\min}$. For instance, in the four-output homodyne case, an error of ± 0.1 in the correct value of L_1 lowers the min-entropy by at most 4%.

We note that, as shown in Fig. 7, there is almost no improvement when the general inefficiency of the experiment η is lower than 12.5%. Any experimental realization that would like to exploit the advantage of a many-outcome configuration should be designed in order to achieve high efficiency.

Although with homodyne detection higher randomness can be certified with respect to heterodyne detection, the former is susceptible to errors in the setting of the phase ϕ between the signal and the LO. Indeed, phase errors induce information loss in homodyne detection, whose magnitude depends on the active phase stabilization response time and precision. It is possible to show that a homodyne detection with phase error $\delta\phi$ is equivalent to a homodyne detection with no phase error and efficiency $\eta = |\cos(\delta\phi)|$. In Fig. 12 we show the optimal min-entropy for a four-outcome homodyne detection as a function of the phase error. As an example, if the phase

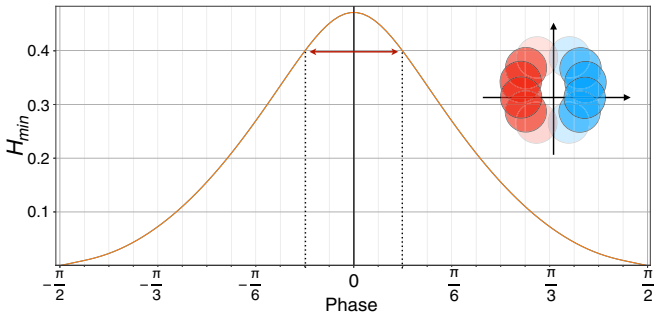

FIG. 12. Optimal min-entropy as a function of phase error for four-outcome homodyne detection.

error is below 15°, the min-entropy may fluctuate between 0.47 and 0.4. On the other hand, heterodyne detection is robust with respect to phase error as long as the sampling rate is much larger than the phase drift: in the latter case, phase-compensation techniques can be used to track and correct phase fluctuations, with minimal impact on the min-entropy. As described in [16], for the heterodyne detection phase drifts can be compensated via software during the postprocessing of the data [see also Fig. 2(c)].

V. CONCLUSION

We have demonstrated a semi-DI QRNG with d outcomes for binary-encoded optical coherent states based on heterodyne or homodyne detection. We compared our results with the binary-outcome case, and we showed the number of certified random bits is improved by increasing the number of outcomes. In this framework, we observed that the homodyne receiver beats the heterodyne receiver in terms of generated randomness. Numerically, we found an asymptotic upper bound of 0.5 as the number of random bits per measurement in the limit of infinite outcomes.

Moreover, in the heterodyne case we found partitioning the phase space into a vertical strip allows a higher generation rate with respect to other configurations (see Fig. 7). Physically, this could be interpreted as better discrimination between the two input states with the strip configuration compared to other phase-space partitions.

From previous analysis [26], it is known that the maximum entropy for a binary-input setup is $\log_2(3) \simeq 1.5849$, while our analysis seems to indicate that with homodyne and heterodyne measurements one can never exceed 0.5 bit of randomness per measurement. We leave for future works the formal proof of the above observation.

It is worth noting that the improvement is significant for perfect detection efficiency, while it decreases in the case of losses. Hence, owning efficient and low-noise detectors is essential for exploiting d-outcome configurations and obtaining higher randomness with respect to the binary-outcome setting. Finally, we illustrated how to apply the d-outcome configuration to the experimental data.

ACKNOWLEDGMENTS

This work was supported by Fondazione Cassa di Risparmio di Padova e Rovigo with the project QUASAR funded within the call "Ricerca Scientifica di Eccellenza 2018"; MIUR (Italian Ministry for Education) under the initiative "Departments of Excellence" (Law No. 232/2016); and the EU Horizon 2020 program under the Marie Sklodowska Curie action, project QCALL (Grant No. GA 675662).

APPENDIX A: DUAL SDP

In the present section, we report how to dualize the primal form of the SDP equation (3). The SDP duality gives an approach to upper bound the optimal value of maximization problems or a lower bound for minimization problems [30]. The dual SDP has several advantages over the primal version. First, the dual-optimization problem returns an upper bound on the guessing probability, while the primal problem returns a lower bound. Thus, even if the solver does not converge to the exact optimal point, the dual solution will never overestimate the true content of randomness, providing reliable bounds. Second, for real-time operation, the dual problem enables us to recompute (suboptimal) bounds without the need to run a full optimization, reducing the resources needed for the entropy estimation. Finally, in the dual problem the finite-size effects can be taken into consideration efficiently, thanks to the linear dependance of $p(b|x)$ in the objective function. Note that in a real experiment, the conditional probabilities $p(b|x)$ are calculated over finite raw data; thus, finite-size effects must be accounted for when estimating the bound.

By using the Lagrangian duality [30], with an approach similar to the one used in [11], the dualized SDP can be written as

$$P_g^* = \min_{H^{\lambda_0,\lambda_1},\nu_{bx}} \left[-\sum_{x=0,1}\sum_{b=0}^{d-1} \nu_{bx} p(b|x) \right] \tag{A1}$$

subjected to

$$H^{\lambda_0,\lambda_1} = (H^{\lambda_0,\lambda_1})^\dagger, \times \sum_x \rho_x \left(\frac{1}{2} \sum_{b=0}^{d-1} \delta_{\lambda_x,b} + \nu_{bx} \right) + H^{\lambda_0,\lambda_1} - \frac{1}{2}\mathrm{Tr}[H^{\lambda_0,\lambda_1}]\mathbb{1} \leqslant 0, \tag{A2}$$

where $H_b^{\lambda_0,\lambda_1}$ are 2×2 Hermitian matrices.

As we can see, the objective function of dual SDP is a linear function of the conditional probability distribution $p(b|x)$, and it does not appear in the constraints. Hence, after solving the dual SDP one time and obtaining a valid set of parameters ν_{bx}^*, it is possible to obtain a (suboptimal) bound for a new set of experimental probabilities $p(b|x)$ by evaluating the objective linear function with the set of parameters ν_{bx}^*. This estimation does not require the full optimization of the SDP, which can be slow and could limit the rate in a real-time operation. A similar approach is not possible with the primal version that needs to run a full optimization of the SDP for every new set of $p(b|x)$.

[1] M. Stipcevic, Quantum random number generators and their applications in cryptography, in *Advanced Photon Counting Techniques VI*, edited by M. A. Itzler, Proc. SPIE 8375, 20 (2012).

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

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

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

[5] Y. Liu, Q. Zhao, M.-H. Li, J.-Y. Guan, Y. Zhang, B. Bai, W. Zhang, W.-Z. Liu, C. Wu, X. Yuan, H. Li, W. J. Munro, Z. Wang, L. You, J. Zhang, X. Ma, J. Fan, Q. Zhang, and J.-W. Pan, Device-independent quantum random-number generation, Nature (London) **562**, 548 (2018).

[6] 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 (London) **556**, 223 (2018).

[7] F. J. Curchod, M. Johansson, R. Augusiak, M. J. Hoban, P. Wittek, and A. Acín, Unbounded randomness certification using sequences of measurements, Phys. Rev. A **95**, 020102(R) (2017).

[8] 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).

[9] M.-H. Li, X. Zhang, W.-Z. Liu, S.-R. Zhao, B. Bai, Y. Liu, Q. Zhao, Y. Peng, J. Zhang, Y. Zhang, W. J. Munro, X. Ma, Q. Zhang, J. Fan, and J.-W. Pan, Experimental Realization of Device-Independent Quantum Randomness Expansion, Phys. Rev. Lett. **126**, 050503 (2021).

[10] Y. Zhang, L. K. Shalm, J. C. Bienfang, M. J. Stevens, M. D. Mazurek, S. W. Nam, C. Abellán, W. Amaya, M. W. Mitchell, H. Fu, C. A. Miller, A. Mink, and E. Knill, Experimental Low-Latency Device-Independent Quantum Randomness, Phys. Rev. Lett. **124**, 010505 (2020).

[11] 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).

[12] 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).

[13] T. Van Himbeeck and S. Pironio, Correlations and randomness generation based on energy constraints, arXiv:1905.09117.

[14] 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).

[15] 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).

[16] M. Avesani, H. Tebyanian, P. Villoresi, and G. Vallone, Semi-Device-Independent Heterodyne-Based Quantum Random-Number Generator, Phys. Rev. Appl. **15**, 034034 (2021).

[17] H. Tebyanian, M. Zahidy, M. Avesani, A. Stanco, P. Villoresi, and G. Vallone, Semi-device independent randomness generation based on quantum state's indistinguishability, Quantum Sci. Technol. **6**, 045026 (2021).

[18] 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).

[19] P. Mironowicz, G. Cañas, J. Cariñe, E. S. Gómez, J. F. Barra, A. Cabello, G. B. Xavier, G. Lima, and M. Pawłowski, Quantum randomness protected against detection loophole attacks, Quantum Inf. Process. **20**, 39 (2021).

[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] Z. Cao, H. Zhou, X. Yuan, and X. Ma, Source-Independent Quantum Random Number Generation, Phys. Rev. X **6**, 011020 (2016).

[22] D. G. Marangon, G. Vallone, and P. Villoresi, Source-Device-Independent Ultrafast Quantum Random Number Generation, Phys. Rev. Lett. **118**, 060503 (2017).

[23] M. Avesani, H. Tebyanian, P. Villoresi, and G. Vallone, Unbounded randomness from uncharacterized sources, arXiv:2010.05798.

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

[26] M. Ioannou, J. B. Brask, and N. Brunner, Upper bound on certifiable randomness from a quantum black-box device, Phys. Rev. A **99**, 052338 (2019).

[27] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, Leftover hashing against quantum side information, IEEE Trans. Inf. Theory **57**, 5524 (2011).

[28] C. R. Müller, C. Peuntinger, T. Dirmeier, I. Khan, U. Vogl, C. Marquardt, G. Leuchs, L. L. Sánchez-Soto, Y. S. Teo, Z. Hradil, and J. Řeháček, Evading Vacuum Noise: Wigner Projections or Husimi Samples?, Phys. Rev. Lett. **117**, 070801 (2016).

[29] N. Walker, Quantum theory of multiport optical homodyning, J. Mod. Opt. **34**, 15 (1987).

[30] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2004).