Semi-device-independent quantum random-number generator resistant to general attacks

Zhenguo Lu,[1,2] Jundong Wu,[1,2] Yu Zhang,[1,2] Shaobo Ren,[1,2] Xuyang Wang,[1,2,3] Hongyi Zhou,[4] and Yongmin Li[1,2,3,*]

[1] *State Key Laboratory of Quantum Optics Technologies and Devices, Institute of Opto-Electronics, Shanxi University, Taiyuan 030006, People's Republic of China*

[2] *Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan 030006, People's Republic of China*

[3] *Hefei National Laboratory, Hefei 230088, People's Republic of China*

[4] *State Key Laboratory of Processors, Institute of Computing Technology, Chinese Academy of Sciences, 100190, Beijing, People's Republic of China*

(Received 2 July 2025; revised 24 March 2026; accepted 17 April 2026; published 8 May 2026)

Quantum random-number generators (QRNGs) produce true random numbers based on the inherent randomness of quantum theory, rendering them a foundational segment of quantum cryptography. Distinguished from trusted-device QRNGs, whose security depends on characterized devices, semi-device-independent (semi-DI) QRNGs permit partial devices to be defective or even maliciously manipulated, achieving a good trade-off between generation rate and security. In this paper, we propose a semi-DI QRNG that resists general attacks while accounting for finite-size effects. The protocol requires no rigorous characterization of the source and measurement devices other than limiting the energy of the emitted states, significantly reducing the demand on practical QRNG systems. Leveraging the tight Kato's inequality for correlated variables, we show that our protocol generates more randomness than it consumes. Furthermore, we demonstrate the scheme using a continuous-variable system with ternary inputs of states. Heterodyne detection is employed to enable phase compensation through data postprocessing, alleviating the stringent requirement on system stability. The system operates at 100 MHz, achieving a net random-number generation rate of 1.165 Mbps at 5.3×10^9 rounds. Our work offers a promising approach to achieving both robust security and a high generation rate with a simple experimental setup.

DOI: 10.1103/gxt7-7zfq

I. INTRODUCTION

True random numbers are essential for cryptography, simulations, and machine learning. The primary metrics of random numbers are statistical uniformity and unpredictability. Random-number generation based on classical physical processes is inherently predictable, whereas quantum random-number generators (QRNGs) leverage the probabilistic nature of quantum state collapse to produce true random numbers that show uniformity and unpredictability [1].

QRNGs with trusted apparatus are most commonly used to generate random numbers and are known for their high generation speed and easy implementation [2–8]. In practical implementations of QRNGs, side-channel information can be leaked to potential adversaries, owing to imperfect devices; this threatens the security and privacy of random generated numbers.

Device-independent (DI) protocols have proven secure with minimal assumptions about the devices used [9–17]. Randomness expansion and randomness amplification are two families of DI protocols. The former requires a perfect randomness input to select measurement settings in a Bell test [9–12], whereas the latter can tolerate an imperfect randomness input [13–17]. However, DI protocols require loophole-free violation of a Bell inequality; these protocols are experimentally challenging to implement and suffer from low random-number generation rates. To balance security and practicality, semi-DI QRNGs that partially trust the device used have been developed [18–32]. Currently, there are two main types of semi-DI QRNG: source-independent (SI) QRNGs [20–27] and measurement-device-independent (MDI) QRNGs [28–32], which avoid the precise characterization of the source and measurement devices, respectively.

Recently, a variety of SI (MDI) QRNG protocols with relaxed security assumptions on the measurement (source) devices have been proposed [33–41]. For example,

*Contact author: yongmin@sxu.edu.cn

 © 2026 American Physical Society

security risks have been examined in SI-QRNGs, arising from afterpulse effects and detection efficiency mismatches in single-photon detectors [33]. Randomness lower bounds in MDI QRNGs with defective single-photon sources have been derived using the decoy state method [35,36]. Semi-DI QRNGs that constrain the overlap or energy of prepared quantum states have also been investigated [37–41]. In these studies, the assumptions of semi-DI frameworks have been relaxed and their security has been improved, while their applicability has been broadened. However, the independent and identically distributed (i.i.d.) assumption is stringently imposed on the measurement devices in current MDI QRNGs. An adversary can employ correlated measurement strategies in different random-number generation rounds at the measurement side; these are known as general attacks [29]. This will result in potentially correlated outputs and thus enhance the adversary's ability to guess the random sequences. A recently proposed numerical framework for security analysis provides lower bounds on the secure randomness of semi-DI QRNGs under general attacks [42].

In this paper, we present a multi-input semi-DI QRNG based on heterodyne detection and demonstrate its security under general attacks. The measurement apparatus is treated as a black box that hides performed quantum measurements to the user. Crucially, our protocol allows an adversary to perform correlated quantum operations at the measurement side without being constrained by the i.i.d. assumption, and takes into account finite-size effects. The upper bound on the energy of the prepared quantum states of the source is constrained without requiring full knowledge of the states emitted in each round. We experimentally validate the semi-DI QRNG with three inputs and a number of outputs in a continuous-variable (CV) system with components that are commercially available off the shelf. Furthermore, as heterodyne detection realizes full phase-space tomography, the phase compensation can be enabled by data postprocessing and a real-time phase stabilization system is unnecessary; this reduces the system complexity and enhances the practicability of the protocol. Finally, since our protocol is, in essence, a form of quantum randomness expansion, we demonstrate that it is capable of generating more randomness than it consumes.

II. PROTOCOL

A. Semi-DI QRNG with ternary inputs

Our protocol consists of three steps: (1) preparation and measurement of quantum states, (2) estimation of the genuinely quantum entropy, and (3) randomness extraction. In step (1), a user randomly designates a prepare-and-measure round as either a test round with probability p_t or a generation round with probability $1-p_t$. In a test round, the user randomly prepares quantum states ρ_x ($x \in \{1,2,3\}$) with probability p_x, which are transmitted to an untrusted measurement station. The measurement device measures the quantum states and outputs the results $y \in \{1,2,3,\ldots,d\}$. In a generation round, the user prepares a fixed state ρ_3, which is also measured by the measurement device to yield the outcomes $y \in \{1,2,3,\ldots,d\}$. This process is repeated N times, and we record the count $N_{y|x}$ representing the number of measurement outcomes y given input state ρ_x during test rounds. In step (2), the user calculates the min-entropy according to Eq. (21) for the finite-size case or Eq. (6) for the asymptotic limit. In step (3), based on the evaluated min-entropy, secure random numbers can be extracted [43] by postprocessing the data from the generation rounds, where a strong extractor is applied.

Similar to previous work [40], our protocol does not require exact characterization of the source; instead, it constrains the lower bound of the overlap of quantum states across different rounds by limiting the energy of the prepared states, as follows:

$$|\langle\psi_x|\psi_{x'}\rangle| \geq 1-2\mu, \qquad x,x' \in \{1,2,3\}, \tag{1}$$

where μ is the upper bound of the mean photon number of the emitted states that can be considered pure states.

In our protocol, (1) we assume an i.i.d. source and suppose that neither the source nor the measurement shows a quantum correlation with the environment, while a classical correlation can exist. (2) Following research on semi-DI QRNGs [37,39,40], we emphasize that the generation of input random numbers (used to select the state and designate the test or generation round in our case) must be independent of both the devices and any adversary.

We now briefly review min-entropy evaluation under collective attacks in the asymptotic limit. In this scenario, the adversary, Eve, executes different but uncorrelated quantum measurements in each round [29]. Under the protocol's assumptions, Eve can manipulate the source and the measurement apparatus to disrupt output randomness. Let us denote Eve's attack strategy by s; then the untrusted measurement apparatus can be characterized by a set of d-dimensional positive operator-valued measure (POVM) operators $\{M_y^s\}$, and the transmitted quantum states can be denoted $|\psi_x^s\rangle$. The lower bound of the inner product of states in Eq. (1) is taken as the worst estimate of randomness destruction. Hence, the states $|\psi_x^s\rangle$ can always be expanded in terms of a set of basis vectors $\{|1\rangle,|2\rangle,|3\rangle\}$ in the Hilbert space, as follows [39]:

$$\begin{aligned}
|\psi_1\rangle &= |1\rangle,\\
|\psi_2\rangle &= \beta|1\rangle + \sqrt{1-\beta^2}|2\rangle,\\
|\psi_3\rangle &= \beta|1\rangle + \beta\sqrt{\frac{1-\beta}{1+\beta}}|2\rangle + \sqrt{\frac{1+\beta-2\beta^2}{1+\beta}}|3\rangle,
\end{aligned} \tag{2}$$

where $\beta = 1 - 2\mu$. Then, Eve's guessing probability P_g can be expressed as

$$P_g = \max_{\{p(s),M_y^s\}} \sum_s p(s) \max_y \mathrm{tr}\left(\rho_3 M_y^s\right), \tag{3}$$

where $p(s)$ represents the probability of Eve adopting measurement strategy s. As shown in Eq. (3), Eve maximizes her guessing power by optimizing the measurement M_y^s and the corresponding distribution $p(s)$. However, as the number of strategies s can be infinite, Eq. (3) is difficult to compute. Following the approach in Bancal *et al.* [44] and Brask *et al.* [37], we divide the measurement operators M_y^s into d groups, each with elements $M_y^{\tilde{s}}$ ($\tilde{s} \in \{1, 2, 3, \ldots, d\}$) and satisfying $\max_y \mathrm{tr}(\rho_3 M_y^{\tilde{s}}) = \mathrm{tr}(\rho_3 M_{\tilde{s}}^{\tilde{s}})$. Therefore, Eq. (3) can be rewritten as

$$P_g = \max_{\{\Pi_y^{\tilde{s}}\}} \sum_{\tilde{s}} \mathrm{tr}\left(\rho_3 \Pi_{\tilde{s}}^{\tilde{s}}\right), \tag{4}$$

where $\Pi_y^{\tilde{s}} = p(\tilde{s}) M_y^{\tilde{s}}$. Meanwhile, the operators $\Pi_y^{\tilde{s}}$ are subject to the following constraints:

$$\begin{aligned} \Pi_y^{\tilde{s}} &\succeq 0, \\ \left(\Pi_y^{\tilde{s}}\right)^\dagger &= \Pi_y^{\tilde{s}}, \\ \sum_y \Pi_y^{\tilde{s}} &= \frac{1}{3}\mathrm{tr}\left(\sum_y \Pi_y^{\tilde{s}}\right) I, \\ \sum_{\tilde{s}=1}^{d} \mathrm{tr}\left(\rho_x \Pi_y^{\tilde{s}}\right) &= p(y|x). \end{aligned} \tag{5}$$

Here, $p(y\,|\,x) = \lim_{N\to+\infty}[N_{y_x}/(p_t p_x N)]$ is the experimentally observed conditional probability. Combining Eqs. (4) and (5), the guessing probability P_g can be determined by semidefinite programming (SDP). The min-entropy describing the lower bound of secure randomness is then calculated as follows:

$$H_{\min} = -\log_2 P_g. \tag{6}$$

Note that, in the asymptotic limit ($N \to +\infty$), the statistical fluctuations in $p(y\,|\,x)$ are negligible and then the optimal probability of test rounds p_t tends to zero. Hence, randomness consumption of the protocol vanishes.

B. Security analysis against general attacks

As previously mentioned, the adversary, Eve, randomly selects a measurement strategy in each round to obtain the side information of the QRNG system. However, by performing correlated quantum operations across a number of rounds, Eve can insert predefined correlations in the outputs, enhancing her guessing ability. Equivalently, Eve executes correlated POVM measurements on the quantum states in different rounds [29]. Moreover, owing to the limited number of rounds, the estimation of conditional probabilities $p(y\,|\,x)$ is inaccurate, making the system vulnerable. Fortunately, analogous to the handling of statistical fluctuations in the source-trusted case [42], we can introduce a stochastic process that witnesses Eve's correlated operations. By employing concentration inequalities to account for correlated variables, the loss of private randomness can then be quantified in both scenarios.

As described by Zhou [42], we first solve the dual problem [45] of Eq. (4):

$$\begin{aligned} \min_{\{H^{\tilde{s}},\omega_{xy}\}} \; & -\sum_{x,y} \omega_{xy} p'(y\,|\,x) \\ \text{s.t.} \quad & (H^{\tilde{s}})^\dagger = H^{\tilde{s}}, \\ & \sum_x \rho_x \left(\delta_{3,x}\delta_{\tilde{s},y} + \omega_{xy}\right) + H^{\tilde{s}} - \frac{1}{3}\mathrm{tr}(H^{\tilde{s}}) I \preceq 0, \end{aligned} \tag{7}$$

to obtain a set of solutions $\{\omega'_{xy}\}$ that satisfy

$$\sum_x \rho_x \left(\delta_{3,x}\delta_{\tilde{s},y} + \omega'_{xy}\right) + H^{\tilde{s}} - \frac{1}{3}\mathrm{tr}(H^{\tilde{s}}) I \preceq 0. \tag{8}$$

As the conditional probabilities $p'(y|x)$ substituted into Eq. (7) are theoretical values computed using a physical model, they are independent of Eve's attack behavior. $p'(y\,|\,x)$ can also be regarded as nominal values [46].

We now define a stochastic process $\{W_n\}$ as follows:

$$W_n = \begin{cases} \dfrac{1}{1-p_t}, & \text{if the } n\text{th round is a generation round and Eve successfully guesses the output;} \\ \dfrac{\omega'_{xy}}{p_t p_x}, & \text{if the } n\text{th round is a test round and the measured output is } y \text{ when emitting } \rho_x; \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

Let $\{\mathcal{F}_{n-1}\}$ be a filtration containing information regarding W_n from the first round to the $(n-1)$th round; we then have

$$E\left(W_n\,|\,\mathcal{F}_{n-1}\right) = \sum_{\tilde{s}} \mathrm{tr}\left(\rho_3 \Pi_{\tilde{s}}^{\tilde{s},\mathcal{F}_{n-1}}\right) + \sum_{x,y,\tilde{s}} \omega'_{xy} \mathrm{tr}\left(\rho_x \Pi_y^{\tilde{s},\mathcal{F}_{n-1}}\right), \tag{10}$$

where $\Pi_y^{\tilde{s},\mathcal{F}_{n-1}}$ is Eve's quantum operation, correlated with the first $n-1$ rounds, that meets the following normalization condition:

$$\sum_y \Pi_y^{\tilde{s},\mathcal{F}_{n-1}} = \frac{1}{3}\mathrm{tr}\left(\sum_y \Pi_y^{\tilde{s},\mathcal{F}_{n-1}}\right) I. \tag{11}$$

Combining Eqs. (8), (10), and (11), we obtain

$$E(W_n \mid \mathcal{F}_{n-1}) \le 0. \tag{12}$$

By the definition of W_n, we have

$$\sum_{n=1}^{N} W_n = \frac{N_g}{1-p_t} + \sum_{x=1}^{3}\sum_{y=1}^{d} \frac{\omega'_{xy}}{p_x p_t} N_{y\,|\,x}, \tag{13}$$

where N_g is the number of successful guesses.

Herein, to achieve a tighter bound of randomness, we employ Kato's inequality, derived by Zhou *et al.* [46], instead of Azuma's inequality used by Zhou [42]. As the first step in applying Kato's inequality, W_n needs to be normalized into the range [0,1]:

$$W_n^{\mathrm{nor}} = \frac{W_n - W_{\min}}{W_{\max} - W_{\min}}, \tag{14}$$

where $W_{\min} = \min W_n$ and $W_{\max} = \max W_n$. By applying Kato's inequality [47] to W_n^{nor}, we have

$$\sum_{n=1}^{N} W_n^{\mathrm{nor}} - \sum_{n=1}^{N} E(W_n^{\mathrm{nor}} \mid \mathcal{F}_{n-1}) \le \left[b + a\left(2\frac{\sum_{n=1}^{N} W_n^{\mathrm{nor}}}{N} - 1\right)\right]\sqrt{N}, \tag{15}$$

which holds with a probability of at least $1-\epsilon$. Here, the parameters $a \in \mathbb{R}$ and $b \ge 0$ satisfy

$$\exp\left(-\frac{2(b^2-a^2)}{\left(1-\frac{4a}{3\sqrt{N}}\right)^2}\right) = \epsilon. \tag{16}$$

From Eq. (12), we have

$$\sum_{n=1}^{N} E(W_n^{\mathrm{nor}} \mid \mathcal{F}_{n-1}) = \sum_{n=1}^{N} \frac{E(W_n - W_{\min} \mid \mathcal{F}_{n-1})}{W_{\max} - W_{\min}} \le -\frac{NW_{\min}}{W_{\max} - W_{\min}}. \tag{17}$$

Thus, given $\sqrt{N} - 2a > 0$, Eq. (15) can be rewritten as

$$W_s^{\mathrm{nor}} \le \frac{\sqrt{N}}{\sqrt{N}-2a}\left[(b-a)\sqrt{N} - \frac{NW_{\min}}{W_{\max} - W_{\min}}\right], \tag{18}$$

where $W_s^{\mathrm{nor}} = \sum_{n=1}^{N} W_n^{\mathrm{nor}}$. Combining Eqs. (13), (14), and (18), we obtain the upper bound N_g^U of the number of successful guesses, namely,

$$N_g^U := \left[\Delta\left(N, \epsilon(a,b)\right) - \sum_{x=1}^{3}\sum_{y=1}^{d} \frac{\omega'_{xy}}{p_x p_t} N_{y\,|\,x}\right](1-p_t), \tag{19}$$

where

$$\Delta(N, \epsilon(a,b)) = \frac{N\left[(b-a)W_{\max} - (b+a)W_{\min}\right]}{\sqrt{N} - 2a}. \tag{20}$$

Therefore, the min-entropy $H'_{\min}$ per generation round against general attacks is determined as

$$H'_{\min} = -\log_2\left(\frac{N_g^U}{N(1-p_t)}\right). \tag{21}$$

Taking into account the randomness consumed in the protocol, the net randomness generation rate R_{net} of our protocol is defined as

$$R_{\mathrm{net}} = R_{\mathrm{gross}} - R_{\mathrm{in}}, \tag{22}$$

where $R_{\mathrm{gross}} = (1-p_t)H'_{\min}$ is the gross randomness produced by the protocol per round, and R_{in} is the input randomness per round. Details of the evaluation of R_{in} are provided in Sec. II C.

C. Implementation

To implement the proposed protocol, the source station randomly prepares coherent states $|\sqrt{\mu}e^{ix(2\pi/3)}\rangle$ with the same probability ($p_x = \frac{1}{3}, x \in \{1,2,3\}$) in the test rounds, and prepares a fixed state $|\sqrt{\mu}\rangle$ in the generation rounds. The prepared states are directed to heterodyne detectors in the measurement station, where the amplitude quadrature X and the phase quadrature P of the optical field are measured. Depending on the region of the phase space in which (X,P) falls, the measured results (X,P) are mapped into four disjoint sets D_y, defined by

$$D_y = \left\{(X,P) \mid \sqrt{X^2+P^2} \in (0,\infty), \tan^{-1}\left(\frac{P}{X}\right) \in \left[\frac{\pi}{2}(y-1), \frac{\pi}{2}y\right)\right\}, \tag{23}$$

where $y \in \{1,2,3,4\}$. The preparation and discretized measurement scheme of our protocol is illustrated in Fig. 1(a). An ideal heterodyne detection can be characterized by a POVM $\Pi_\alpha = (1/\pi)|\alpha\rangle\langle\alpha|$, where $\alpha = X + iP$ is the complex amplitude of the coherent state $|\alpha\rangle$. For an

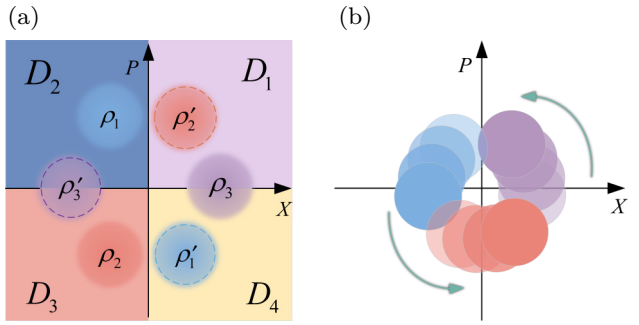

FIG. 1. (a) Complementary modulation and measurement. The modulated quantum states have the same intensity μ but different phases. In each quantum random-number generator execution round, two successive temporal modes of coherent states are modulated with phase difference π . The early mode (ρ_1, ρ_2, or ρ_3) serves as the prepared state in the protocol, while the corresponding late mode (ρ_1', ρ_2', or ρ_3') acts as an auxiliary state to balance the dc level of the homodyne detectors. The measured results of the two quadratures (X, P) are discretized into D_1, D_2, D_3, and D_4. (b) Impact of phase drifts on observed states in phase space, where the prepared states undergo a random rotation over time.

input state $|\sqrt{\mu}e^{ix\frac{2}{3}\pi}\rangle$, the probability density of the outputs (X, P) of the heterodyne detection is given by the well-known Husimi function:

$$Q(\alpha) = \frac{1}{\pi} e^{-\left|X+iP-\sqrt{\eta\mu}e^{ix\frac{2}{3}\pi}\right|^2}, \tag{24}$$

where η is the total detection efficiency. In our case, η can be expressed as $\eta = \eta_q\eta_e$, where η_q is the quantum efficiency of the photodiodes and η_e is the equivalent efficiency induced by the electronic noise [48]. Consequently, the theoretical conditional probability distribution $p'(y\,|\,x)$ can be calculated as

$$\begin{aligned} p'(y\,|\,x) &= \iint_{D_y} Q(\alpha)\, dX\, dP \\ &= \frac{e^{-\eta\mu}}{\pi} \iint_{D_y} e^{-r^2+2r\sqrt{\eta\mu}\cos\left(\varphi-x\frac{2}{3}\pi\right)} r\, dr\, d\varphi. \end{aligned} \tag{25}$$

In cases of collective attacks, we can substitute $p'(y\,|\,x)$ for $p(y\,|\,x)$ in Eq. (5) and obtain the expected performance of the measurement device through SDP. In cases of general attacks, $p'(y\,|\,x)$ can be substituted as the nominal value in Eq. (7) to obtain the solution $\{\omega_{xy}'\}$ of the dual SDP. It is worth noting that ω_{xy}' is closely related to the randomness bound R_{net}. Herein, we introduce a constraint on ω_{xy}', namely $\omega_{xy}' \le 0$, which still satisfies Eq. (8) and, therefore, preserves security. On completion of these constraints, we can apply Kato's inequality to the constructed random variables W_n^{nor}, with the parameters a and b from Eq. (15) optimized to tighten R_{net}, subject to the constraints

$$\begin{aligned} &\exp\left(-\frac{2(b^2-a^2)}{\left(1-\frac{4a}{3\sqrt{N}}\right)^2}\right) = \epsilon, \\ &\sqrt{N}-2a>0, \quad a\in\mathbb{R}, \quad b\ge 0. \end{aligned} \tag{26}$$

Based on the definition of net randomness generation rate R_{net} in Eq. (22), we simulate the effects of key parameters on the performance of our protocol, where we

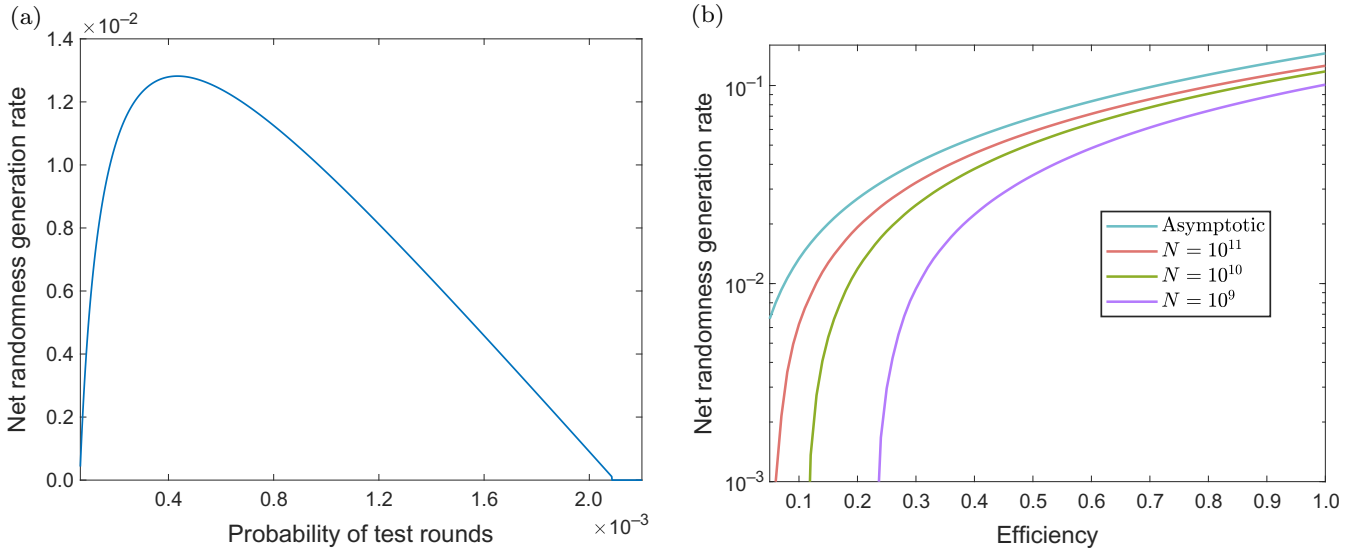

FIG. 2. (a) Simulations of net randomness generation rate vs probability of test rounds, with $\mu = 0.005$, $\eta = 23.2\%$, $N = 5.3 \times 10^9$, and $\epsilon = 10^{-10}$. (b) Simulations of net randomness generation rate vs detection efficiency for different total numbers of rounds, with $\epsilon = 10^{-10}$, optimized mean photon number, and optimized probability of test round.

set $N_{y|x} = p_x p_t N p'(y|x)$ and the violation probability of Kato's inequality $\epsilon = 10^{-10}$. Figure 2(a) shows R_{net} as a function of the probability of test rounds. Figure 6 shows simulations of R_{gross} and R_{net} versus mean photon number. Figure 2(b) plots the relation between R_{net} and η for different total numbers of rounds N, where the mean photon number and probability of test rounds are chosen optimally. As N increases, R_{net} converges to the asymptotic limit $H_{\min}$. However, real-time randomness extraction in real-world applications necessitates finite block lengths to sustain throughput. In this case, both finite-size and non-i.i.d. effects become significant and non-negligible.

As pointed out in Sec. II A, a truly random input is optimal for the execution of our QRNG protocol. Given this, to show the practical utility of our protocol, we need to prove that it generates more randomness than it consumes. In our work, the input randomness serves three tasks: state preparation in test rounds, switch between generation and test rounds, and the seed for the randomness extractor. As we employ a strong extractor, whose seed is considered part of its output, the randomness consumption is thus attributed to the first two components. Therefore, following the analysis of Wang *et al.* [32], one could employ the interval algorithm [49] to convert a uniform bit string into the required biased bit string. In this scenario, the average input randomness per round R_{in} of our protocol (for large N) is given by

$$R_{\text{in}} = h_2(p_t) + p_t H(x), \tag{27}$$

where $h_2(\cdot)$ is the binary entropy function and $H(x)$ is the Shannon entropy of the input distribution $\{p_x\}$.

III. EXPERIMENT

The proposed protocol is demonstrated with a CV fiber system. The experimental setup is illustrated in Fig. 3. The system consists of two main parts: partially trusted quantum state preparation and untrusted quantum state measurement. All optical components are connected using polarization-maintaining fibers and are commercially available.

In the quantum state preparation, a laser module emits continuous-wave light with a wavelength of 1538 nm, which is split into two paths by a 99:1 beam splitter (BS). The major portion serves as the local oscillator (LO) for heterodyne detection, and the minor portion is the signal for the state preparation. In the signal path, the light is injected into a $LiNbO_3$ phase modulator (PM) with a bandwidth of 300 MHz. An arbitrary waveform generator (AWG) operating at 100 MHz randomly prepares coherent states with phases of 0, $\frac{2}{3}\pi$, and $\frac{4}{3}\pi$. The modulated states are then attenuated using a variable optical attenuator (VOA) and split into two paths by a 90:10 BS. 90% of the light is directed into an optical power meter (Thorlabs S132C; resolution, 1 nW), while the remaining 10% is connected to a fixed attenuator (ATT). The prepared states $|\sqrt{\mu}e^{ix(2\pi/3)}\rangle$ are directed to the measurement station. By combining the power meter readout and the precalibrated attenuation factor of the ATT, the intensity of the prepared optical field is characterized. Here, we use a time window of 10 ns (time duration per quantum state) to estimate the mean photon number μ of prepared states. The monitoring of the energy of quantum states is considered trusted.

In quantum state measurement, the LO passes through a VOA that adjusts the light intensity, and is then split into two beams by a 90:10 BS. 10% of the light is directed to the Meter for the purpose of monitoring the LO power. Notice that the protocol requires no precise characterization of the measurement; therefore, the LO is monitored not for security purposes but merely to enhance the performance of the QRNG in this proof-of-principle experiment. The remaining 90% of the LO and the prepared states are injected into the LO port and the signal port of the 90° optical hybrid (90° OH), respectively. Two balanced amplified photodetectors (BDs), each with a bandwidth of 1.6 GHz, measure the two quadratures of the prepared states. By careful calibration, the efficiency of the heterodyne detection in our experiment was found to be 23.2%. The output signals from the BDs are sampled by a Tektronix MSO5204B mixed-signal oscilloscope (MSO) with a sampling frequency of 1 GHz. To realize synchronization sampling, the MSO and AWG share a reference clock and the MSO is triggered by the AWG. Finally, the acquired data are sent to a computer for offline postprocessing. To reconstruct the signal states, 10 samples are averaged to represent a single signal state.

The parameters employed in the experiment are presented in Table I, where the mean photon number of the prepared states μ and the probability of test rounds p_t have been optimized. The proportion of generation rounds is chosen to be large to increase the randomness generation rate. This implies that the modulator spends the majority of its time preparing a fixed state, which induces a dc imbalance in two ac-coupled homodyne detectors and causes distortions of quantum signals [50]. To mitigate this effect, we adopt a complementary modulation scheme, following the approach of Wang *et al.* [32], which is presented in Fig. 1(a). In our scheme, each prepared state is paired with an auxiliary state exhibiting a π phase difference within each round. Thus, the two temporal modes (the prepared and auxiliary state) yield opposite expectation values for the same quadrature measurements, thereby balancing the dc output of two homodyne detectors. In addition, to precisely locate the valid signal segments used for QRNG execution, we insert a predefined reference signal before each valid signal segment. The data synchronization process is implemented by searching the maximum correlation between the reference signal and the detected signal. The data frame configuration of our experimental implementation is shown in Fig. 4.

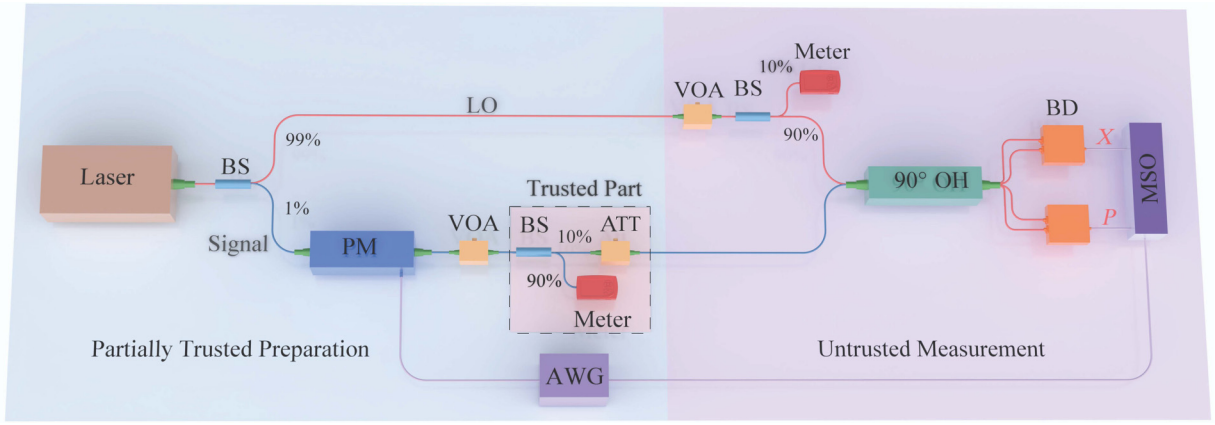

FIG. 3. Experimental setup of semi-device-independent quantum random-number generator. The phase of the input signal is randomly modulated by a PM, while its amplitude is evaluated through the meter reading from the trusted part. After optimizing the intensity of the local oscillator (LO), the prepared state is interfered with the LO in a 90° OH, followed by measurement of two quadratures using a pair of BDs. Through the statistical dependence between the prepared states and measurement outcomes, the lower bound on the generated randomness is established. ATT: fixed optical attenuator; AWG, arbitrary waveform generator; BD, balanced amplified photodetector; BS, beam splitter; Meter: optical power meter; MSO, mixed-signal oscilloscope; PM: phase modulator; VOA, variable optical attenuator; 90° OH, 90° optical hybrid.

In practical implementations, the relative phase θ between the quantum signals and local oscillator (LO) might drift over time, owing to environmental variations or modulator instability. Consequently, the phases of the three prepared coherent states deviate from their ideal values of $\frac{2}{3}\pi$, $\frac{4}{3}\pi$, and 2π, and the observed quantum states rotate in the phase space, as illustrated in Fig. 1(b). This leads to a reduction in the extractable randomness because such a phase drift can be intuitively considered equivalent to classical noise in $N_{y|x}$ in Eqs. (5) and (19). Considering that the time scale of the phase drift is much slower than that of the sampling rate, therefore, the sampled data sequences can be divided into a number of blocks. For each block that contains three quantum signals, the phases of the quantum signals remain nearly constant. To optimize phase-tracking accuracy, we select a block of 10 000 prepared states, based on the experimental observation that the duration of each block (200 μs) is slightly shorter than the characteristic time of the phase fluctuations (of the order of milliseconds), as shown in Fig. 5. By averaging (X, P) of ρ_3 (which is prepared in most of the time) in each block, the shot noise can be suppressed and the phase of ρ_3 can be determined as $\theta = \tan^{-1}(\overline{P}/\overline{X})$. Therefore, the three quantum signals within each block are rotated by $-\theta$ to correct the phases to the configuration shown in Fig. 1(a). This allows us to assign a measured output y to each (X, P) group.

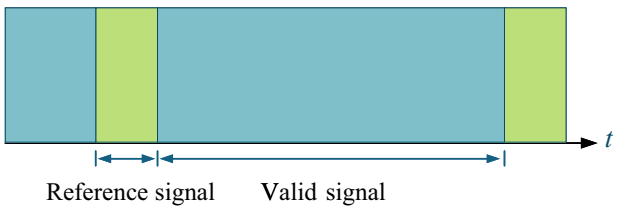

FIG. 4. Data frame configuration of quantum random-number generator (QRNG) experiment. Reference signals for data synchronization are periodically inserted between the valid signals used for QRNG execution. Valid signals consist of the prepared and auxiliary signals required for the complementary modulation.

IV. RESULTS

In our experiment, with mean photon number $\mu = 0.005$, we ran a total of 5.3×10^9 rounds, of which about 1.8×10^6 rounds were selected as test rounds to obtain the statistics $N_{y|x}$. Hence, based on Eqs. (19) and (21), the

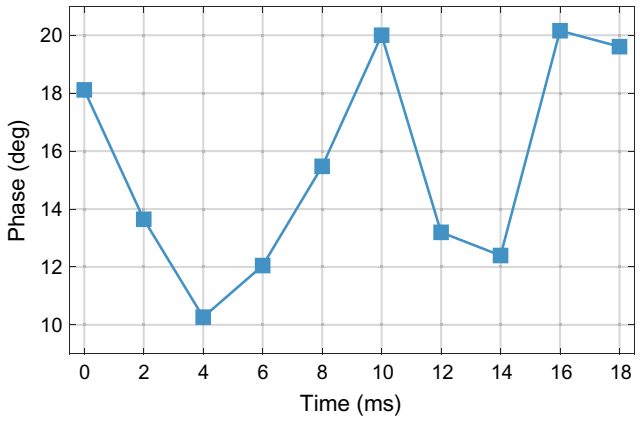

FIG. 5. Phase drift of prepared state ρ_3 vs time. The phase is evaluated at intervals of 2 ms.

TABLE I. Experimental parameters.

Mean photon number μ	0.005
Probability of test rounds p_t	3.464×10^{-4}
Total number of rounds N	5.3×10^{9}
Detection efficiency η	23.2%
Input distribution in test rounds p_x	$\frac{1}{3}, \forall x$
Violation probability of Kato's inequality ϵ	10^{-10}

gross randomness generation rate R_{gross} obtained by our QRNG execution is estimated to be 0.01668 bits/round. Considering the consumed randomness for choosing test rounds and prepared states, the average randomness consumption rate R_{in} in our case, evaluated via Eq. (27), is 0.00503 bits/round. Consequently, our QRNG system achieves a net randomness generation rate of 0.01165 bits per round (green cross in Fig. 6).

Employing a two-bit encoding scheme, we converted the sequence y into a binary sequence (raw bit string). A secure random-number sequence was then generated from the raw bit string, according to H'_{min}, with a strong randomness extractor based on a Toeplitz hash function [43,51,52]. Finally, we performed randomness tests on the generated bit sequence using the NIST Statistical Test Suite and found that the random sequence passed the NIST tests.

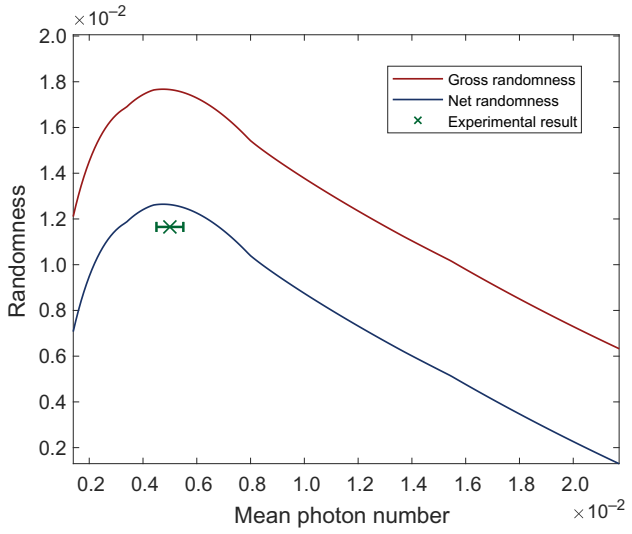

FIG. 6. Randomness vs mean photon number with $\eta = 23.2\%$, $p_t = 3.464 \times 10^{-4}$, $N = 5.3 \times 10^{9}$, and $\epsilon = 10^{-10}$. Red and blue curves: expected gross and net randomness generation rates, respectively; green cross: experimental net randomness generation rate, where the error bar denotes the estimation precision on the mean photon number.

V. CONCLUSION

In this study, we implement a semi-DI QRNG based on relaxed assumptions, using commercially available components. Our protocol requires an easily verifiable energy-bound assumption on the source and allows the adversary to perform non-i.i.d. general attacks at the measurement side. Finite-size effects are also incorporated in the security analysis. Compared with the randomness estimation presented elsewhere [37,39,40], in our work, we explicitly take into account the randomness overhead and demonstrate a positive net randomness generation rate, thereby enhancing its practical feasibility. In addition, heterodyne detection is employed, enabling phase compensation through postprocessing without the need for an active phase stabilization system, reducing the complexity of QRNG implementation.

Operating at 100 MHz, our QRNG system achieves a secure net random-number generation rate of 1.165 Mbps, offering a promising approach for a high-security, robust, and high-generation-rate semi-DI QRNG.

In future work, we intend to improve the system repetition rate to increase the generation rate of the QRNG. It is expected that the protocol can be extended to more input states and discretization outputs in a CV system. In this case, the min-entropy per round can be further increased.

ACKNOWLEDGMENTS

We thank S. Liu and H. Shi for helpful discussions. This work was supported by the National Natural Science Foundation of China (Grants No. 62205188 and No. 62175138) and Quantum Science and Technology–National Science and Technology Major Project (Grant No. 2021ZD0300703).

DATA AVAILABILITY

The data that support the findings of this article are not publicly available on publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors on reasonable request.

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

[2] Y. Shen, L. Tian, and H. Zou, Practical quantum random number generator based on measuring the shot noise of vacuum states, Phys. Rev. A **81**, 063814 (2010).

[3] J. Y. Haw, S. M. Assad, A. M. Lance, N. H. Y. Ng, V. Sharma, P. K. Lam, and T. Symul, Maximization of extractable randomness in a quantum random-number generator, Phys. Rev. Appl. **3**, 054004 (2015).

[4] M. A. Wayne, E. R. Jeffrey, G. M. Akselrod, and P. G. K., Photon arrival time quantum random number generation, J. Mod. Opt. **56**, 516 (2009).

[5] B. Qi, Y.-M. Chi, H.-K. Lo, and L. Qian, High-speed quantum random number generation by measuring phase noise of a single-mode laser, Opt. Lett. **35**, 312 (2010).

[6] T. Symul, S. M. Assad, and P. K. Lam, Real time demonstration of high bitrate quantum random number generation with coherent laser light, Appl. Phys. Lett. **98**, 231103 (2011).

[7] T. Gehring, C. Lupo, A. Kordts, D. Solar Nikolic, N. Jain, T. Rydberg, T. B. Pedersen, S. Pirandola, and U. L. Andersen, Homodyne-based quantum random number generator at 2.9 Gbps secure against quantum side-information, Nat. Commun. **12**, 605 (2021).

[8] B. Bai, J. Huang, G.-R. Qiao, Y.-Q. Nie, W. Tang, T. Chu, J. Zhang, and J.-W. Pan, 18.8 Gbps real-time quantum random number generator with a photonic integrated chip, Appl. Phys. Lett. **118**, 264001 (2021).

[9] S. Pironio, A. Acín, S. Massar, A. B. De La Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Random numbers certified by Bell's theorem, Nature **464**, 1021 (2010).

[10] 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 **562**, 548 (2018).

[11] 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).

[12] P. J. Brown, S. Ragy, and R. Colbeck, A framework for quantum-secure device-independent randomness expansion, IEEE Trans. Inf. Theory **66**, 2964 (2020).

[13] M. Kessler and R. Arnon-Friedman, Device-independent randomness amplification and privatization, IEEE J. Sel. Areas Inf. Theory **1**, 568 (2020).

[14] R. Gallego, L. Masanes, G. De La Torre, C. Dhara, L. Aolita, and A. Acín, Full randomness from arbitrarily deterministic events, Nat. Commun. **4**, 2654 (2013).

[15] F. G. S. L. Brandao, R. Ramanathan, A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, T. Szarek, and H. Wojewodka, Realistic noise-tolerant randomness amplification using finite number of devices, Nat. Commun. **7**, 11345 (2016).

[16] R. Colbeck and R. Renner, Free randomness can be amplified, Nat. Phys. **8**, 450 (2012).

[17] A. Kulikov, S. Storz, J. D. Schär, M. Sandfuchs, R. Wolf, F. Berterottière, C. Hellings, R. Renner, and A. Wallraff, Device-independent randomness amplification, arXiv:2412.17931.

[18] X. Lin, R. Wang, S. Wang, Z.-Q. Yin, W. Chen, G.-C. Guo, and Z.-F. Han, Certified randomness from untrusted sources and uncharacterized measurements, Phys. Rev. Lett. **129**, 050506 (2022).

[19] M. Pivoluska, M. Plesch, M. Farkas, N. Ružičková, C. Flegel, N. H. Valencia, W. McCutcheon, M. Malik, and E. A. Aguilar, Semi-device-independent random number generation with flexible assumptions, npj Quantum Inf. **7**, 50 (2021).

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

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

[22] 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).

[23] D. Drahi, N. Walk, M. J. Hoban, A. K. Fedorov, R. Shakhovoy, A. Feimov, Y. Kurochkin, W. S. Kolthammer, J. Nunn, J. Barrett, and I. A. Walmsley, Certified quantum random numbers from untrusted light, Phys. Rev. X **10**, 041048 (2020).

[24] M. Avesani, H. Tebyanian, P. Villoresi, and G. Vallone, Unbounded randomness from uncharacterized sources, Commun. Phys. **5**, 273 (2022).

[25] L. Li, M. Cai, T. Wang, Z. Tan, P. Huang, K. Wu, and G. Zeng, On-chip source-device-independent quantum random number generator, Photonics Res. **12**, 1379 (2024).

[26] Y. Du, X. Hua, Z. Zhao, X. Sun, Z. Zhang, X. Xiao, and K. Wei, Source-independent quantum random number generators with integrated silicon photonics, Commun. Phys. **8**, 9 (2025).

[27] H. Zhou, Continuous-variable source-independent quantum random number generator with phase-insensitive detection, IEEE J. Sel. Top. Quantum Electron. **31**, 1 (2025).

[28] A. Chaturvedi and M. Banik, Measurement-device–independent randomness from local entangled states, EPL **112**, 30003 (2015).

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

[30] 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).

[31] F. Bischof, H. Kampermann, and D. Bruß, Measurement-device-independent randomness generation with arbitrary quantum states, Phys. Rev. A **95**, 062305 (2017).

[32] C. Wang, I. W. Primaatmaja, H. J. Ng, J. Y. Haw, R. Ho, J. Zhang, G. Zhang, and C. Lim, Provably-secure quantum randomness expansion with uncharacterised homodyne detection, Nat. Commun. **14**, 316 (2023).

[33] X. Lin, S. Wang, Z.-Q. Yin, G.-J. Fan-Yuan, R. Wang, W. Chen, D.-Y. He, Z. Zhou, G.-C. Guo, and Z.-F. Han, Security analysis and improvement of source independent quantum random number generators with imperfect devices, npj Quantum Inf. **6**, 100 (2020).

[34] Y. Li, Y. Fei, W. Wang, X. Meng, H. Wang, Q. Duan, Y. Han, and Z. Ma, Impact of an imperfect sampling device on the security of a continuous-variable source-independent quantum random-number generator with a

phase-randomized local oscillator, Phys. Rev. A **109**, 052401 (2024).

[35] Y.-Q. Nie, H. Zhou, B. Bai, Q. Xu, X. Ma, J. Zhang, and J.-W. Pan, Measurement-device-independent quantum random number generation over 23 Mbps with imperfect single-photon sources, Quantum Sci. Technol. **9**, 025024 (2024).

[36] T. Wu, C.-H. Zhang, X.-Y. Zhou, J. Li, and Q. Wang, Measurement-device-independent quantum random-number generator with source flaws, IEEE Photonics J. **15**, 1 (2023).

[37] 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).

[38] 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).

[39] 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).

[40] M. Avesani, H. Tebyanian, P. Villoresi, and G. Vallone, Semi-device-independent heterodyne-based quantum random-number generator, Phys. Rev. Appl. **15**, 034034 (2021).

[41] T. V. 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).

[42] H. Zhou, Numerical framework for semi-device-independent quantum random-number generators, Phys. Rev. A **107**, 052402 (2023).

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

[44] J.-D. Bancal, L. Sheridan, and V. Scarani, More randomness from the same data, New J. Phys. **16**, 033011 (2014).

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

[46] H. Zhou, T. Sasaki, and M. Koashi, Numerical method for finite-size security analysis of quantum key distribution, Phys. Rev. Res. **4**, 033126 (2022).

[47] G. Kato, Concentration inequality using unconfirmed knowledge, arXiv:2002.04357 [math.PR].

[48] J. Appel, D. Hoffman, E. Figueroa, and A. I. Lvovsky, Electronic noise in optical homodyne tomography, Phys. Rev. A **75**, 035802 (2007).

[49] T. S. Han and M. Hoshi, Interval algorithm for random number generation, IEEE Trans. Inf. Theory **43**, 599 (1997).

[50] T. McDermott, *Wireless Digital Communications: Design and Theory* (Tucson Amateur Packet Radio Corporation, Tucson, AZ, 1998), 2nd ed.

[51] H. Krawczyk, in *Advances in Cryptology – CRYPTO '94*, Lecture Notes in Computer Science, Vol. 839, edited by Y. G. Desmedt (Springer Berlin Heidelberg, Berlin, 1994), p. 129.

[52] H. Krawczyk, in *Advances in Cryptology – EUROCRYPT '95*, Lecture Notes in Computer Science, Vol. 921, edited by L. C. Guillou and J.-J. Quisquater (Springer, Berlin, 1995), p. 301.