

Coherence as a resource for source-independent quantum random-number generation

Jiajun Ma,¹ Aishwarya Hakande,¹ Xiao Yuan,^{2,1,*} and Xiongfeng Ma^{1,†}

¹*Center for Quantum Information, Institute for Interdisciplinary
Information Sciences, Tsinghua University, Beijing 100084, China*

²*Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom*

Measuring quantum states provides a means to generate genuine random numbers. It has been shown that genuine randomness can be obtained even with an uncharacterized source by measuring two incompatible bases [Phys. Rev. X 6, 011020 (2016)]. As coherence is the necessary source for generating randomness, we extend the scheme and propose a framework for quantum random number generation with general uncharacterized coherence resource. The previous scheme can be treated as a special case under the framework by considering a nonlinear uncertainty-relation-based coherence witness. Considering general coherence witnesses, we propose a source-independent random-number generation scheme that achieves a higher randomness generation rate. Our paper highlights the close relation between coherence and random number generation, and may shed light on designing general semi-device-independent quantum information processing protocols.

arXiv:1704.06915v2 [quant-ph] 25 Apr 2019

* xiao.yuan.ph@gmail.com

† xma@tsinghua.edu.cn

I. INTRODUCTION

Random number generation has many important applications in various tasks. In some cases, such as Monte Carlo simulation, it only requires the random numbers to be statistically unbiased. Pseudo random numbers or physical random numbers based on classical mechanics, such as coin flipping and noise measuring, are sufficient. The outcomes of these procedures may appear random, but they are in principle predictable. In cryptography, one of the security foundations lies on the unpredictability of random numbers. For instance, a cryptophytic key requires genuinely random bits. The random numbers via classical mechanic procedures are not suitable for cryptosystems. Therefore, it is important to study the generation of unpredictable (or genuine) random numbers.

According to Born's rule [1], the measurement outcome of a quantum system that is in the superposition of the measurement basis is unpredictable. Based on quantum mechanic principles, there are many quantum random number generation (QRNG) schemes proposed in the past two decades. For a review of the subject, one can refer to Refs. [2, 3] and the references therein. In general, a QRNG setup consists of two parts, a source that contains randomness and a measurement that reads out the randomness. As shown in Fig. 1, the source emits a sequence of quantum signals, and the readout system measures them to produce random outcomes. For instance, if the states from the source form a sequence of qubits $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, and the measurement is a projection onto the $\{|0\rangle, |1\rangle\}$ basis, the outcomes are genuinely random.

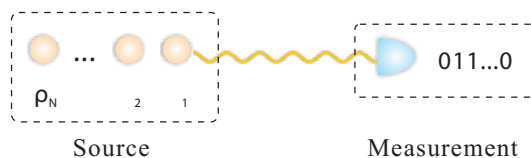


FIG. 1. A quantum random-number generator can be generally decomposed into two parts: source and measurement.

In practice, the source may contain both genuine randomness and classical noise. The latter can normally be influenced by some unexpected environment parameters, such as the temperature. In a cryptographic picture, an adversary Eve may take advantage of the classical noise. Consider a source that emits maximally mixed state $\rho = 1/2(|0\rangle\langle 0| + |1\rangle\langle 1|)$. Although the measurement outcomes on the Z basis appear random, they are not genuinely random. This can be understood in the presence of Eve, who simply prepares $|0\rangle$ or $|1\rangle$ based on a predetermined random string. In this case, the measurement outcomes are entirely predictable. In fact, the predictability in random-number generators becomes a major security issue in current cryptosystems [4].

The key job of randomness analysis is to quantify the genuine randomness so that it can be extracted. One way of randomness quantification relies on modeling the QRNG implementation, which normally requires calibrating the source and measurement devices [5]. In practice, however, the calibration is often hard to perform thoroughly. Once implemented devices deviate from theoretical models, the randomness can be compromised. Several proposals have been proposed to solve this problem. For instance, QRNGs via certifying randomness based on nonlocality tests [6] are called self-testing or device-independent schemes where neither the calibration of the quantum source nor the measurement is required. In practice, realizing a loophole-free Bell test is experimentally challenging, which requires high-fidelity state preparation and detection efficiency, as well as the accurate chronological sequence control [7]. Furthermore, the randomness generation rate is quite limited owing to a small violation of Bell inequalities obtained with the state-of-the-art technology [8, 9].

In real-life applications, a fully device-independent scenario may be too restrictive. By putting certain reasonable assumptions to devices, the performance of QRNGs would become practically acceptable. Along this direction, tremendous efforts have been devoted to find a trade-off between device independence and high randomness generation rate [10–15]. In this paper, we focus on the scenario of generating genuine randomness with well-calibrated measurement devices but uncharacterized sources, namely, source-independent quantum random-number generation (SIQRNG) [11]. With certain reasonable assumptions on measurement devices, the SIQRNG schemes can be very practical. For instance, with a continuous-variable system, the randomness generation rate in such a scenario has achieved the gigabits per second (Gbps) regime [12]. The intuition behind these schemes is the quantum uncertainty relation. Given two complementary measurement bases, X and Z , if the outcome uncertainty of the X -basis measurement is small, its uncertainty of the Z -basis measurement must be large. The information on the complementary basis can be used to reveal the genuine randomness within the source. Since the uncertainty relation is state independent, these QRNG schemes are source independent.

Recently, genuine randomness has been shown to be essentially related to the coherence of the input quantum state in the measurement basis [16, 17]. There, the source is trusted and the extractable randomness is quantified in the asymptotic limit. In this paper, we propose a framework that extends this idea to a more general setting where

the source is uncharacterized and the measurement outcomes are finite. We show that SIQRNG can be realized via measuring the coherence of the input quantum states. Moreover, we propose a method of estimating the coherence via coherence witnesses. By designing a nonlinear coherence witness, we show that the coherence estimation yields the same randomness quantification with the previous uncertainty-relation-based SIQRNG schemes. This coherence witness picture sheds light on the fact that the uncertainty-relation-based SIQRNG does not maximally efficiently extract the genuine randomness from the source. Our randomness analysis is based on the assumption that the source emits signals that are independent and identically distributed (i.i.d.). Furthermore, we propose a new SIQRNG scheme that uses tomography to estimate coherence and thereby generally achieves a higher randomness generation rate than the uncertainty-relation-based ones.

The paper is organized as follows. In Sec. II, we review the preliminaries on coherence measures and SIQRNG. Then, in Sec. III, we introduce a witness to measure the coherence. In Sec. IV, we present the framework of QRNG via measuring coherence, based on which, a SIQRNG protocol is proposed in Sec. V. In Sec. VI, with numerical simulations, we show that our protocol generally achieves a higher randomness generation rate than the uncertainty-relation-based schemes. Finally, we conclude in Sec. VII with discussions on interesting related perspective subjects.

II. PRELIMINARIES: COHERENCE AND SIQRNG

A. Resource theory of coherence

The resource theory of coherence formalizes the intuition that quantum superposition is non-classical [18]. In this framework, with an orthogonal basis $\Pi = \{|i\rangle\}$ as the reference basis, one can define an incoherent state as $\sigma = \sum_i p_i |i\rangle \langle i|$, with $p_i \geq 0$ and $\sum_i p_i = 1$. Incoherent operations are physical realizable operations that map incoherent states to incoherent states. Specifically, they are formed by the Kraus operators, $\{K_i\}$, that satisfy $K_i \mathcal{I} K_i^\dagger \in \mathcal{I}$, where \mathcal{I} is the set of incoherent states.

Under this resource framework, a coherence measure needs to fulfill a few criteria. There are many proposals for coherence measures, such as the l_1 -norm of coherence [18], the coherence of formation [16], and the robustness of coherence [19]. In this paper, we adopt the relative entropy of coherence [18],

$$C(\rho) = H(\Delta_\Pi(\rho)) - H(\rho), \quad (1)$$

where $H(\rho)$ is the von Neumann entropy of ρ and $\Delta_\Pi(\rho) \equiv \sum_i |i\rangle \langle i| \rho |i\rangle \langle i|$ is the dephasing operation in the reference basis Π . Recently, the relative entropy of coherence is linked to the genuine randomness obtained by measuring ρ in the basis Π [17].

B. SIQRNG

In the framework of SIQRNG, Eve prepares a bipartite quantum state of systems A and E , represented by τ_{AE} , where each system consists of N partitions. Considering the most general attack by Eve, the joint state of the $2N$ partitions τ_{AE} can be prepared in an arbitrary bipartite state, where each partition may have arbitrary dimension, including the dimension of 0, and the joint state can be entangled among partitions and between A and E . In this case, τ_{AE} can represent any quantum state with arbitrary dimension. Then Eve sends system A to the legitimate user Alice and retains the rest system E . After receiving A , Alice measures each of the N partitions using a random measurement setting. From Alice's perspective, as illustrated in Fig. 2, the source effectively emits a sequence of quantum states $\{\rho_i\}$, where ρ_i denotes the reduced state of partition $i \in \{1, 2, \dots, N\}$. Note that, in general, different partitions $\{\rho_i\}$ can be entangled with each other as the joint state τ_{AE} is generally entangled. Alice tries to extract genuine randomness from the measurement outcomes. Meanwhile, Eve aims at predicting the random numbers extracted from τ_A with the assistance of τ_E , where $\tau_A = \text{Tr}_E[\tau_{AE}]$ and $\tau_E = \text{Tr}_A[\tau_{AE}]$ are the reduced density matrices of A and E , respectively.

In this paper, we consider a common optical realization of the SIQRNG, as most of the practical QRNGs are implemented with quantum optics [2, 3]. In the measurement setting, we assume Alice uses typical optical components, such as phase modulators and threshold detectors, and she employs random assignment for double-click events. This is a widely used detection model with optical implementation.

In the source-independent scenario, since the source is assumed to be controlled by Eve, τ_A can be prepared in an arbitrary dimensional Hilbert space. As a major component of the security analysis for SIQRNG, we employ the *squashing model* to remove the dimension arbitrariness of τ_A [20, 21]. In the squashing model, Alice first applies a squashing operation which projects the received state to a qubit or a vacuum state and then performs the qubit

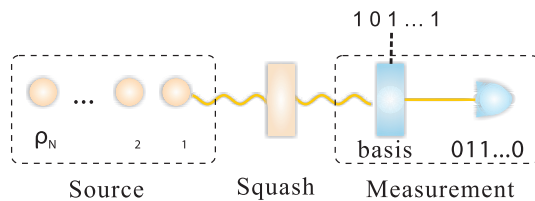


FIG. 2. Source-independent quantum random-number generator. The squashing model is applied to transform the input states into qubits and vacuum states. Then, a basis is chosen to measure the squashed qubits. The dashed line on the measurement side denotes that, in some schemes [11], initial random bits are needed to choose the measurement basis randomly. Note that in some other schemes [12], a beam splitter is used to randomly select the measurement basis, which, of course, might raise the question why one should trust a beam splitter. Such kinds of discussions appeared in Ref. [11].

measurement. In literature [22, 23], the squashing model has been proven to be equivalent to the common optical implementations using threshold detectors and proper postprocessing. Note that the squashing model can be applied here since in the source-independent scenario, the measurement device is assumed to be trusted, thus it can be composed of threshold detectors that fit into the squashing model. Then, in the following randomness analysis, we directly employ this squashing model. In the postprocessing, the squashing model requires Alice to randomly assign the measurement outcome to 0 or 1 for the double-click events, which would affect the net generated randomness. This random assignment issue will be further analyzed in Sec. V B.

After the squashing operation, Alice can project state τ_A to n qubits and $N - n$ vacuum states. Alice performs qubit measurements on the n qubits. The rest $N - n$ vacuum states can be regarded as measurement losses. Since measurement devices are trusted in the SIQRNG scenario, one can assume that the n qubits are fair sampled from the N states [11]. That is, the detection efficiency loophole here is not considered. Here, we remark that our analysis can be applied to other cases, such as the one where Alice uses a photon non-demolition measurement (compatible with the squashing model). In practice, post-selection (discarding losses) is feasible, whereas photon non-demolishing measurement technology is not available with current technologies. If Alice uses post selection, she might need initial randomness in making measurement basis choices. However, in the asymptotic limit, the amount of this consumed initial randomness can be reduced to a negligible ratio. In fact, for the scheme proposed in Table I, Alice can choose $q \rightarrow 0$, and this amount of randomness is sufficient for randomness extraction. The following randomness analysis will focus on the n squashed qubits.

Here, we briefly review the previous SIQRNG scheme [11].

- 1) An untrusted source (controlled by Eve) emits a sequence of quantum states.
- 2) Alice (or Eve) squashes the quantum states into qubits and vacua. Alice discards the vacua and retains the n squashed qubits.
- 3) Alice randomly chooses n_x qubits out of the n squashed qubits and measures them in the X basis. Within n_x outcomes, the ratio of outcome $|-\rangle$ is e_{xb} , which is defined to be the error rate. Ideally, Alice expects the source to emit state $|+\rangle$ and hence the result of $|-\rangle$ is defined as an error.
- 4) Alice measures the rest $n_z = n - n_x$ squashed qubits in the Z basis to obtain n_z raw random bits.
- 5) Alice extracts $n_z(1 - S(e_{xb} + \theta)) - t_e$ random bits from the raw random bits using a universal hashing function, where S is the binary Shannon entropy, θ is the deviation due to statistical fluctuations, and 2^{-t_e} is the failure probability of the randomness extraction.

In above scheme, the security proof techniques of QKD are employed in the randomness analysis. Here is the argument. Ideally, the source should emit states $|+\rangle$. Then, Alice measures them in the Z basis $\{|0\rangle, |1\rangle\}$ to generate random numbers. In the scenario of SIQRNG, the source is allowed to emit arbitrary quantum states in arbitrary dimensions. On the measurement end, one can consider a virtual protocol. First, a squashing model is applied to project the quantum states into a sequence of qubit and vacuum states. Then Alice performs an error correction procedure that transforms all states to $|+\rangle$. Finally she measures them in the Z basis. By designing the error correcting code appropriately, this operation can be commuting with the Z -basis measurement [24]. Also the squashing model is proven to be equivalent to the threshold detection model with appropriate postprocessing [22]. Hence, Alice does not need to perform the virtual protocol, which requires a universal quantum computer. Instead, she can perform the Z -basis measurement first and then apply a randomness extraction on the measurement outcomes. The number of extractable random bits, $n_z(1 - S(e_{xb} + \theta)) - t_e$, is derived from the uncertainty principle together with the X -basis

error rate e_{xb} . The randomness extraction essentially functions the same as the error correction procedure in the virtual protocol.

In the randomness extraction, or the error correction in the virtual protocol, Alice needs to know the error rate in the X basis, $\{|+\rangle, |-\rangle\}$, where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. Thus, in the SIQRNG scheme, Alice needs to randomly test the quantum state in the X basis to estimate its error rate, which is later used for randomness quantification [11].

In this paper, we would show the randomness of SIQRNG from a difference perspective. The idea is based on the recently discovered link between coherence and genuine randomness in a quantum state [16, 17]. Instead of applying error correction, Alice estimates the coherence of the qubit states via measuring randomly sampled qubit states. Then she measures the rest qubits in the Z basis to generate the raw randomness. The estimated coherence is used to bound the genuine randomness of the raw data. In postprocessing, she distills the genuine randomness from the raw data.

III. MEASURING COHERENCE WITH A COHERENCE WITNESS

Normally, to estimate the coherence of an unknown state, one needs to perform a full state tomography to obtain the density matrix ρ . In some cases, full tomography information is not available. For example, as the dimension of the state increases, the number of required measurement in tomography increases quadratically, which becomes challenging for experiments. Thus, it is interesting to estimate the coherence of an unknown state without a full tomography.

A similar problem raises in the field of entanglement measure, where it is expected to estimate the entanglement with a limited number of measurements. The solution there is to employ entanglement witnesses, which are originally designed to justify whether quantum states are entangled or not to estimate entanglement [25–27]. Recently, this idea has been extended to the resource theory of coherence [19], i.e., estimating the coherence with coherence witnesses.

The original coherence witness is a linear function of the density matrix ρ [19]. Here, we extend this notion to an arbitrary function of ρ .

Definition 1. A coherence witness is a function of ρ , $W(\rho)$, that satisfies the following criteria,

1. $\forall \rho \in \mathcal{I}$, $W(\rho) \geq 0$;
2. $\exists \rho \notin \mathcal{I}$, $W(\rho) < 0$.

A linear coherence witness has been shown to be useful to bound the coherence [19]. Here, we design a nonlinear coherence witness to bound the relative entropy of coherence.

Lemma 1. Given a reference basis $\Pi = \{|i\rangle\}$ in a d -dimensional Hilbert space,

$$W_u(\rho) = H(\Delta_{\Xi}(\rho)) - \log_2 d, \quad (2)$$

is an coherence witness, where Ξ and Π are mutually unbiased bases of the same Hilbert space, so that Ξ is maximally incompatible with Π .

Proof. The incoherent state set \mathcal{I} is defined in basis Π . Since Ξ is a mutually unbiased basis of Π , for all $\rho \in \mathcal{I}$, we have $H(\Delta_{\Xi}(\rho)) = \log_2 d$, and hence $W_u(\rho) = 0$. Also, for all $|i'\rangle \in \Xi$, we have $H(\Delta_{\Xi}(|i'\rangle\langle i'|)) = 0$, and hence $W_u(|i'\rangle\langle i'|) = -\log_2 d < 0$. \square

Theorem 1. Given a reference basis $\Pi = \{|i\rangle\}$ and a state ρ in a d -dimensional Hilbert space, the relative entropy of coherence $C(\rho)$ can be bounded by the coherence witness $W_u(\rho)$ defined in Eq. (2),

$$C(\rho) \geq -W_u(\rho) = \log_2 d - H(\Delta_{\Xi}(\rho)), \quad (3)$$

where Ξ is a mutually unbiased basis of Π .

Proof. The dephasing operators, $\Delta_{\Pi}(\rho)$ and $\Delta_{\Xi}(\rho)$, can be viewed as two projection measurements, which have the quantum uncertainty relation [28],

$$H(\Delta_{\Pi}(\rho)) + H(\Delta_{\Xi}(\rho)) \geq -\log_2 c + H(\rho), \quad (4)$$

where $c = \max_{i,i'} |\langle i|i'\rangle|^2$, with $|i\rangle \in \Pi$ and $|i'\rangle \in \Xi$.

The two bases Π and Ξ are mutually unbiased, and hence $c = 1/d$. Rearranging the terms in Eq. (4) and using the definition $C(\rho) = H(\Delta_{\Pi}(\rho)) - H(\rho)$, Eq. (3) is obtained. \square

From Theorem 1, one can estimate the relative entropy of coherence via measuring the state in the complementary basis of the reference basis. This idea is similar to the one employed in the uncertainty-relation-based SIQRNG [11, 12]. There, the intrinsic randomness generated via the Z -basis measurement is estimated by measuring the state in the complementary X basis. In the next section, we propose a framework that formalizes the relation between these two scenarios.

IV. FRAMEWORK OF SIQRNG VIA MEASURING COHERENCE

The task of estimating coherence of an unknown quantum state shares similarities with randomness evaluation in SIQRNG. In both scenarios, the source state is uncharacterized whereas the measurement is trusted. Meanwhile, the amount of genuine randomness within the source can be quantified by the coherence of the quantum state [16, 17]. Therefore, extracting genuine randomness in SIQRNG can be reduced to the problem of estimating coherence within the source. In this section, we would present a framework that links the two tasks.

A. Quantification of randomness

Following the discussion of SIQRNG in Sec. II, we focus on the n squashed qubits, which contribute one raw data bit each. Alice needs to quantify the genuine randomness in the n -bit raw data from the n -qubit state, $\tau_A = \text{Tr}_E(\tau_{AE}) \in \mathcal{H}_2^{\otimes n}$, where \mathcal{H}_2 denotes a two-dimensional Hilbert space. In the partial trace, we put the $N - n$ vacuum states to system E . Note that the n qubits can be correlated with each other, or even with Eve's system τ_E .

Suppose Alice randomly chooses n_z qubits and measures them in the Z basis to generate raw random bits, whereas she measures the rest $n - n_z$ qubits in some other complementary bases for parameter estimation, which would give Alice information about τ_A . Denote the measurement outcome in the Z basis (an n_z -bit string) by K_z . Here, Alice's measurement can be viewed as a dephasing operation on each qubit of subsystem A in the Z basis, $\Delta_{Z^{\otimes n_z}}^A(\tau_{AE})$. Then the randomness contained in K_z is quantified by [29]

$$R^{\varepsilon_1}(K_z) = \min_{\tau_{AE}} H_{\min}^{\varepsilon_1}(A|E)_{\Delta_{Z^{\otimes n_z}}^A(\tau_{AE})}, \quad (5)$$

where the minimization runs over all possible states of Eve that satisfy $\text{Tr}_E(\tau_{AE}) = \tau_A$, and $H_{\min}^{\varepsilon_1}$ is the smooth min-entropy, defined in Appendix A, with a smooth parameter ε_1 .

The min-entropy $R^{\varepsilon_1}(K_z)$ is the key parameter for randomness extraction. With universal hashing [30], such as Teplitz-matrix hashing, one can extract random bits that are ε -close to a uniformly distributed string from Eve's point of view. Here, the security parameter is $\varepsilon = \varepsilon_1 + \varepsilon_2$, with ε_2 as the failure probability introduced in the randomness extraction procedure.

B. Randomness analysis

In the following, we analyze the randomness with the assumption that the n squashed qubits are i.i.d. This assumption is also made in the scenario of collective attacks in QKD, where Eve attacks each signal in an i.i.d. manner. In a more general setting, the n qubits might be entangled, which corresponds to the scenario of coherent attacks in QKD. It is proven that the security parameter for coherent attacks is only polynomially larger than the security parameter for collective attacks [31]. The extra information available to the adversary for coherent attacks can be compensated by slightly reducing the size of the final random bits in the privacy amplification stage. We expect that a similar argument can be employed here to obtain the security proof against the most general sources. We leave the randomness analysis with an correlated source for future study.

With the i.i.d. assumption, the joint state that outputs the raw random bits can be expressed by $\tau_{AE} = \rho_{AE}^{\otimes n_z}$, where ρ_{AE} is the squashed joint quantum state of each signal. From Eq. (5), one can have

$$R^{\varepsilon_1}(K_z) = \min_{\rho_{AE}} H_{\min}^{\varepsilon_1}(A|E)_{\Delta_{Z^{\otimes n_z}}^A(\rho_{AE}^{\otimes n_z})}. \quad (6)$$

For $n_z \geq \frac{8}{5} \log_2 \frac{2}{\varepsilon_1^2}$, the smooth min-entropy can be lower bounded by the conditional von Neumann entropy, defined as $H(A|E)_{\rho_{AE}} = H(\rho_{AE}) - H(\rho_E)$ [32]. Thus one has

$$R^{\varepsilon_1}(K_z) \geq n_z \min_{\rho_{AE}} H(A|E)_{\Delta_{Z^{\otimes n_z}}^A(\rho_{AE})} - 7.09 \sqrt{n_z \log_2 \frac{2}{\varepsilon_1^2}}, \quad (7)$$

whose derivation is shown in Appendix B. Meanwhile, $\min_{\rho_{AE}} H(A|E)_{\Delta_{Z^{\otimes n_z}}^A(\rho_{AE})}$ is related to the relative entropy of coherence of ρ_A [17],

$$\min_{\rho_{AE}} H(A|E)_{\Delta_{Z^{\otimes n_z}}^A(\rho_{AE})} = C(\rho_A), \quad (8)$$

where the reference basis for $C(\rho_A)$ is the Z basis. Inserting Eq. (8) into Eq. (7), one has

$$R^{\varepsilon_1}(K_z) \geq n_z C(\rho_A) - 7.09 \sqrt{n_z \log_2 \frac{2}{\varepsilon_1^2}}. \quad (9)$$

We remark that this expression only holds for $n_z \geq \frac{8}{5} \log_2 \frac{2}{\varepsilon_1^2}$, which is normally satisfied in practice, typically, $n_z \geq 95$ for $\varepsilon_1 = 10^{-10}$.

C. Randomness analysis via a coherence witness

The randomness analysis result Eq. (9) implies that one can estimate the amount of randomness in a quantum state by measuring the coherence of the state. In Sec. III, we have shown that, without full state tomography, one can lower bound the coherence with coherence witnesses. Therefore, there is a close relation between the coherence witness and the SIQRNG: Any coherence witness that is able to lower bound the coherence can be employed to realize a SIQRNG scheme.

As an example, we apply this analysis method to the SIQRNG scheme described in Sec. IIB. The measurement used by Alice in the SIQRNG scheme forms a coherence witness W_u as shown in Eq. (2). Then, applying Theorem 1 to Eq. (9), the number of genuine random bits, denoted by $R_u^{\varepsilon_1}$, is estimated by

$$R_u^{\varepsilon_1} \geq -n_z W_u(\rho_A) - 7.09 \sqrt{n_z \log_2 \frac{2}{\varepsilon_1^2}}, \quad (10)$$

where $W_u(\rho_A) = H(\Delta_X(\rho_A)) - 1$. In the asymptotic limit, when the number of emitted qubits n approaches infinitely large, the randomness generation rate is given by

$$r_u = \frac{R_u^{\varepsilon_1}}{N} \Big|_{N \rightarrow \infty} \geq q\beta(1 - H(\Delta_X(\rho_A))), \quad (11)$$

where $\beta = n/N$ is the transmittance of the signal, and $q = n_z/n$ is the ratio of Z -basis measurement. In this limit, Alice can set $q \rightarrow 1$ to maximize r_u . Note that Eq. (11) coincides with the randomness generation rate evaluated via the complementary uncertainty relation [11].

To summarize, the security analysis in our framework is divided by two steps. First is the squashing operation, which maps uncharacterized signal states to qubit states. Applying the squashing model requires proper postprocessing of measurement outcomes, such as random assignment of double clicks to be discussed in Section VB. Second is the coherence characterization of the squashed qubit states, which quantifies the secure random bits of the statistics obtained by measuring the qubit states.

V. TOMOGRAPHY-BASED SIQRNG

Note that the SIQRNG protocols based on the complementary uncertainty relation do not necessarily extract the maximal amount of the randomness from the source. For instance, suppose that the source emits state $(|0\rangle + i|1\rangle)/\sqrt{2}$, from Eq. (11), one obtains a lower bound of r_u to be 0, thus no genuine randomness can be extracted. Nevertheless, the measurement outcome on the Z basis is in fact genuinely random. In this case, the genuine randomness cannot be revealed by the coherence witness using the X measurement. Instead, the randomness can be witnessed with another complementary basis $Y = \{|\pm i\rangle = (|0\rangle \pm i|1\rangle)/\sqrt{2}\}$. Without *a priori* knowledge of the source, one might choose a bad witness to underestimate the genuine randomness.

For the SIQRNG scheme described in Sec. IIB, by adding one more measurement basis, Alice can obtain a better estimation of the coherence of ρ_A via a full state tomography. Then she can extract more genuine randomness from the raw data. Based on this observation, we propose a SIQRNG protocol based on tomography as presented in Table I.

Note that in Table I the Z -basis measurement data are used for both measurement tomography and randomness generation. We remark that the data used for tomography is, in principle, kept secret from any other party, which means the testing data are not revealed to the eavesdropper. Therefore, in principle, the X - and Y -basis measurement data can also be used to extract extra randomness with the analysis similar to that of the Z -basis measurement data. But in the limit where $q \rightarrow 0$, the amounts of randomness generated by the X - and Y -basis measurement become negligible.

TABLE I. Source-independent quantum random number generation

<p>1. State preparation</p> <p>(a) An untrusted source (might be controlled by Eve) emits N quantum states in arbitrary dimensions, which are sequentially sent to the readout system.</p> <p>(b) A <i>squashing operation</i> transforms the quantum states into n qubits and $N - n$ vacua.</p> <p>2. Measurement</p> <p>(a) The $N - n$ vacua are discarded and the remaining n squashed qubit states are post-selected for randomness generation.</p> <p>(b) Alice randomly chooses n_x, n_y, and n_z qubits for the X-, Y-, and Z-basis measurements, with probability $q_x = q$, $q_y = q$, $q_z = 1 - 2q$, respectively. Denote p_x, p_y, and p_z to be the rates to obtain the outcomes $+\rangle$, $i+\rangle$, and $0\rangle$, respectively.</p> <p>(c) The Z-basis measurement outcomes are recorded as the raw data.</p> <p>3. State tomography With the measurement results, p_x, p_y, and p_z, Alice can estimate the density matrix of the squashed qubits, ρ_A. Note that statistical fluctuations need to be considered here.</p> <p>4. Randomness evaluation and extraction With the information of ρ_A, Alice can bound the genuine randomness of the raw data and apply a proper randomness extractor to obtain the final random bits.</p>

In the protocol, the coherence of the source $C(\rho_A)$ can be accurately estimated with a full tomography of ρ_A . Then, the number of genuine random bits, denoted by $R_t^{\varepsilon_1}$, can be estimated via Eq. (9),

$$R_t^{\varepsilon_1} \geq n_z C(\rho_A) - 7.09 \sqrt{n_z \log_2 \frac{2}{\varepsilon_1}}. \quad (12)$$

In the asymptotic limit, the randomness generation rate, denoted by r_t , is given by

$$r_t \geq \frac{R_t^{\varepsilon_1}}{N} \Big|_{N \rightarrow \infty} = q_z \beta C(\rho_A), \quad (13)$$

where $\beta = n/N$ is the transmittance of the signal and $q_z = n_z/n$ is the ratio of Z -basis measurement. In large data-size limit, Alice can set $q_z \rightarrow 1$.

In tomography, the density matrix, ρ_A can be estimated from measurement outcomes, p_x , p_y , and p_z , defined in Table I. Write ρ_A as $\rho_A = (I + (2\vec{p} - 1) \cdot \vec{\sigma})/2$, where $\vec{p} = (p_x, p_y, p_z)$ and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. Substituting ρ_A into Eq. (1) and (12), one can get

$$C(\rho_A) = H(p_z) - H\left(\frac{p_o + 1}{2}\right), \quad (14)$$

and

$$R_t^{\varepsilon_1} \geq n_z H(p_z) - n_z H\left(\frac{p_o + 1}{2}\right) - 7.09 \sqrt{n_z \log_2 \frac{2}{\varepsilon_1}}, \quad (15)$$

with $p_o = \sqrt{4(p_x^2 + p_y^2 + p_z^2 - p_x - p_y - p_z) + 3}$.

A. Squashing model

As discussed in Sec. II, the squashing model should be applied to project the quantum state from the uncharacterized source into a sequence of qubits. The squashing model requires that, depending on the measurement setting, appropriate postprocessing should be implemented. Here, Alice can employ the postprocessing of the squashing model used in the analysis of quantum state tomography [33].

According to the Supplemental Material of Ref. [33], the double click events should be considered to derive a set of bounds of the obtained statistics. Specifically, for each measurement bases $j \in \{X, Y, Z\}$, let n_j^0 , n_j^1 and n_j^d denote the numbers of the two single-click events and the double-click event, respectively. Then Alice obtains a set of qubit states, \mathcal{S} , that is compatible with the measurement results. That is, they fulfill the following constraints,

$$\frac{n_j^0}{n_j^0 + n_j^1 + n_j^d} \leq p_j \leq \frac{n_j^0 + n_j^d}{n_j^0 + n_j^1 + n_j^d}. \quad (16)$$

In the following, we denote the lower bound for p_j in the above inequalities by p_j^L , and the upper bound by p_j^U .

In our protocol, Alice needs to consider the worst case of $\rho_A \in \mathcal{S}$. That is, she should minimize $C(\rho_A)$ in Eq. (15) over \mathcal{S} . In Appendix C, we show that $C(\rho_A)$ is a unimodal function with respect to each p_j , with the minimal value achieved for $p_j = 1/2$. Without loss of generality, from now on, we assume $p_j^U \geq 1/2$, otherwise Alice can flip the bit label in the j basis. Denote the worst-case value of p_j for the coherence quantification by p_j^w , thus $p_j^w = \max(p_j^L, 1/2)$.

B. Double clicks

As discussed in Secs. II B and Sec. V A, the squashing model requires the random assignment of double-click events. Note that for the double-click events in the X and Y bases, the random assignment postprocessing need not be actually implemented as these measurement outcomes are only used for tomography testing in Eq. (16) where Alice only needs to count the number of double-click events and evaluate the errors introduced by these events. On the other hand, as the measurement outcome in the Z basis is used to generate the raw random bits, Alice should implement the random assignment on the double-click events to map them to single-value outcomes. In the rest of this subsection, we first introduce the random assignment method as directly required by the squashing model. Then we introduce an alternative discard method, which is more practical in experiments.

1. Random assignment method

Here, we consider a postprocessing method that Alice randomly assigns 0 or 1 to the double-click events in the Z basis. After the squashing model analysis, Alice obtains p_z^w as the worst case estimation of p_z . Then, in the random assignment procedure, the probability of assigning value 0 to the double-click events, denoted by p_a , should be compatible with p_z^w ,

$$p_z^w = \frac{n_z^0 + p_a n_z^d}{n_z^0 + n_z^1 + n_z^d}. \quad (17)$$

Thus, p_a is given by

$$p_a = \frac{p_z^w n_z - n_z^0}{n_z^d}. \quad (18)$$

Note that the random assignment method generally consumes extra randomness, which should be taken into account when evaluating the net randomness generation rate. Here, the randomness cost in the double-click assignment procedure is $n_z^d H(p_a)$. Thus the asymptotic net randomness generation rate is given by

$$\begin{aligned} R_t^n &= R_t^{\varepsilon_1} - n_z^d H(p_a) \\ &\geq n_z H(p_z^w) - n_z H\left(\frac{p_z^w + 1}{2}\right) - 7.09 \sqrt{n_z \log_2 \frac{2}{\varepsilon_1^2}} \\ &\quad - n_z^d H(p_a), \end{aligned} \quad (19)$$

where $p_o^w = \sqrt{4(p_x^{w2} + p_y^{w2} + p_z^{w2} - p_x^w - p_y^w - p_z^w) + 3}$.

2. Discard method

In practice, the random assignment might be technically challenging to implement. Thus, we consider a simpler method to deal with the double-click events in the Z basis — discarding all the double-click events. Denote ρ^s and

ρ^d to be the density matrices of the squashing qubit states, single-click, and double-click, respectively. Then, one has $p_d \rho_A^d + (1 - p_d) \rho_A^s = \rho_A$, where $p_d = n_z^d / n_z$ is the ratio of double-click events in the Z basis. Once discarding all the double-click events, the random bits in the remaining data can be lower bounded by

$$R_t^n \geq n_z^s C(\rho_A^s) - 7.09 \sqrt{n_z \log_2 \frac{2}{\varepsilon_1^2}}, \quad (20)$$

where $n_z^s = n_z - n_z^d$ is the number of single-click events in the Z basis. To estimate $C(\rho_A^s)$, one can employ the concavity of relative entropy of coherence,

$$p_d C(\rho_A^d) + (1 - p_d) C(\rho_A^s) \geq C(\rho_A). \quad (21)$$

Thus

$$\begin{aligned} n_z^s C(\rho_A^s) &\geq n_z C(\rho_A) - n_d C(\rho_A^d) \\ &\geq n_z C(\rho_A) - n_d. \end{aligned} \quad (22)$$

where $C(\rho_A^d) \leq 1$ is used in the second inequality. Combining Eqs. (22), (20), and (15), the net randomness generation rate is given by

$$\begin{aligned} R_t^n &\geq n_z H(p_z^w) - n_z H\left(\frac{p_o^w + 1}{2}\right) \\ &\quad - 7.09 \sqrt{n_z \log_2 \frac{2}{\varepsilon_1^2}} - n_z^d \end{aligned} \quad (23)$$

Note that the randomness generation rate by the discard method is generally smaller than that the rate by the random assignment method in Eq. (19).

C. Analysis of statistic fluctuations

In practice, the number of squashed qubits n is finite. Thus the probabilities used for parameter estimation would suffer from statistical fluctuations. Here, we analyze the finite-data-size effect on the estimation of p_j^w . In order to distinguish the probabilities with the measurement rates, denote the expectation values of p_j^w to be \bar{p}_j , which would be inserted into Eq. (14) to evaluate the genuine randomness. Since the qubits are assumed to be i.i.d., Alice can employ the Hoeffding inequality [34] to estimate the discrepancy between p_j^w and \bar{p}_j caused by the statistical fluctuations,

$$\text{Prob}(\bar{p}_j \leq p_j^w - \theta) \leq e^{-2\theta^2 n_j} = \varepsilon_j, \quad (24)$$

Here, we assume $p_j^w - \theta \geq 1/2$, otherwise we take the worst bound $\bar{p}_j = 1/2$. Replacing $\{p_j^w\}$ by $\{p_j^w - \theta\}$ in Eq. (19) or (23), one can obtain the lower bound of randomness with a total failure probability,

$$\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_x + \varepsilon_y + \varepsilon_z, \quad (25)$$

where ε_1 is introduced by the smooth parameter and ε_2 is introduced by randomness extraction.

VI. SIMULATION

In this section, we first analyze the performance of the tomography-based SIQRNG and compare it to the original proposal. Then, by taking account of statistical fluctuations, we optimize the ratio of qubits used for the tomography testing.

A. Comparing to the original SIQRNG

For simplicity, we consider the asymptotic limit where the number of emitted quantum states N is infinitely large. Then, the randomness generation rate for the original protocol is given by Eq. (11), whereas the rate for the

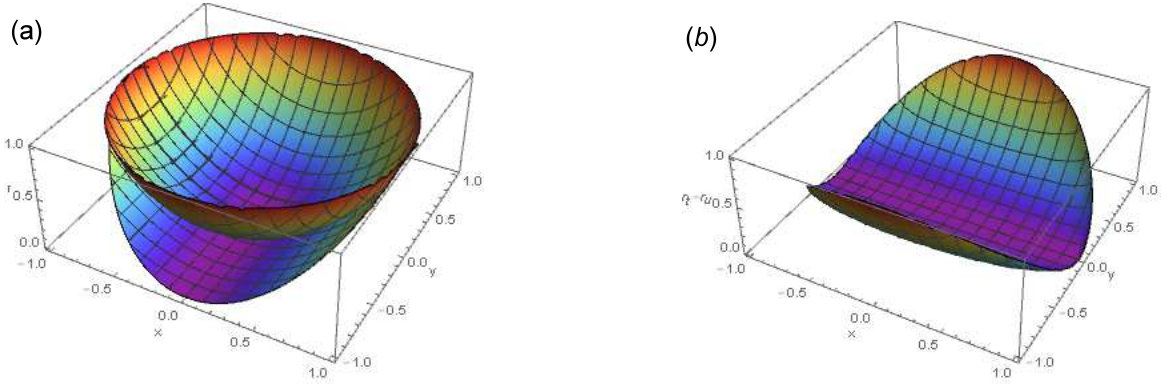


FIG. 3. Comparison of the randomness generation rates with an input qubit state $\rho_A = \frac{I+x\sigma_x+y\sigma_y}{2}$, with $N \rightarrow \infty, q_z = 1, \beta = 1$. (a) The lower surface describes the randomness generation rate for the uncertainty-relation-based scheme r_u as shown in Eq. (26), whereas the upper surface describes the randomness generation rate for the tomography-based scheme r_t as shown in Eq. (27). (b) Illustration of the gap between the two schemes, $r_t - r_u \geq 0$.

tomography-based protocol is given by Eq. (13). In both cases, q_z is set to be 1. Besides, we assume the single photon source is used without considering the photon loss and the detector inefficiency, and hence the transmittance $\beta = 1$. Then, the randomness generation rate for the original protocol is given by

$$r_u|_{q_z=1, \beta=1} \geq 1 - H(\Delta_X(\rho_A)), \quad (26)$$

and the tomography-based protocol by

$$r_t|_{q_z=1, \beta=1} \geq C(\rho_A). \quad (27)$$

In the comparison, we assume the input state has the form of $\rho_A = (I + x\sigma_x + y\sigma_y)/2$, where x and y are two parameters and $x^2 + y^2 \leq 1$. The comparison between Eqs. (26) and (27) is illustrated in Fig. 3. One can clearly see that the tomography-based scheme generally provides a higher randomness generation rate than the original proposal. The larger the parameter y is, the bigger gaps the two schemes have. In general, one can consider a more general state, $\rho_A = (I + x\sigma_x + y\sigma_y + z\sigma_z)/2$, where the gap of randomness generation rate between the two schemes is nonzero as long as $y \neq 0$.

B. Parameter optimization

Now we analyze the performance of the tomography-based protocol by simulating a practical experiment setup. Details of the simulation model is presented in Appendix D. Consider a practical source, consisting of a laser and a polarization modulator, which emits N coherent-state pulses with an intensity of μ_0 . We assume the quantum state is prepared to be $|+\rangle$, which is then transmitted through a depolarization channel. Thus, the received quantum state can be described by

$$\rho_A = p\frac{I}{2} + (1-p)|+\rangle\langle+|, \quad (28)$$

where $p \in [0, 1]$. The readout system consists of a polarization rotator (used for basis selection), a polarization beam splitter, and two threshold detectors with the same detection efficiencies. Denote the total transmittance by η , including the detector efficiency and the coupling efficiency between the source and the detector. It is equivalent to consider a coherent state with intensity of $\mu \equiv \mu_0\eta$. Here, we ignore the detection caused by the dark count since for QRNG dark counts are normally negligible comparing to η . For a more comprehensive model taking account of the dark counts, one can refer to the corresponding QKD model [35].

In this model, we put the misalignment errors into the parameter p . Then, the simulated worst estimations of p_j

are given by,

$$\begin{aligned}\bar{p}_x &= \frac{1 - p - e^{-\mu} + pe^{-\frac{\mu}{2}}}{1 - e^{-\mu}} - \theta, \\ \bar{p}_y &= \frac{e^{-\frac{\mu}{2}} - e^{-\mu}}{1 - e^{-\mu}} - \theta, \\ \bar{p}_z &= \frac{e^{-\frac{\mu}{2}} - e^{-\mu}}{1 - e^{-\mu}} - \theta.\end{aligned}\tag{29}$$

Meanwhile, the number of double click events in the Z basis is

$$n_z^d = N(1 - 2q)(1 + e^{-\eta\mu_0} - 2e^{-\frac{\eta\mu_0}{2}}).\tag{30}$$

The detailed model with the calculations of Eqs. (29) and (30) are shown in Appendix D. Note that all the \bar{p}_j here are less than $1/2$, as required in the randomness analysis shown in Sec. V. Then one can evaluate the extractable randomness from Eq. (23). Here, we adopt the discard method to process the double-click events in the Z basis. In fact, one would obtain the same randomness generation rate by the random assignment method in the case of Eq. (28).

In the simulation, we pick $p = 0$, $p = 0.1$, $p = 0.3$ for the input state of Eq. (28), and set $\varepsilon_x = \varepsilon_y = \varepsilon_z = \varepsilon_1 = \varepsilon_2 = 10^{-10}$ and the number of pulses $N = 10^{10}$. First, by optimizing the tomography testing parameter q , we show the dependence of the randomness generation rate, given by Eq. (23), on the intensity μ in Fig. 4. From the figure, one can see that, initially, the randomness generation rate increases with the intensity of the signal μ due to the increase in the single-click events relative to the no-click events. As μ keeps increasing, the double-click events become dominant. Then, the randomness generation rate starts to decrease. Thus, there is an optimal choice of μ . In experiment, one should characterize total transmittance η first and then set the light intensity μ_0 to make $\mu = \mu_0\eta$ optimal.

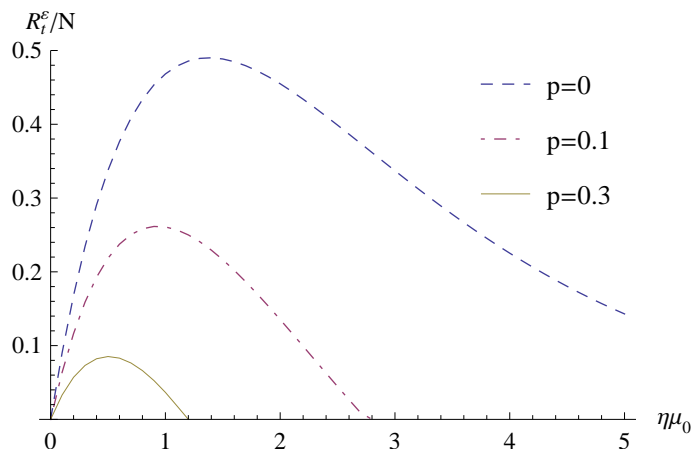


FIG. 4. Randomness generation rate vs. intensity μ with various depolarization parameters p . The optimal μ for $p = 0$, 0.1 , and 0.3 are $\mu = 1.4$, 0.9 , and 0.5 , respectively.

Next, we investigate how the optimal tomography testing parameter, q , varies with the number of pulses, N . Here, we pick $p = 0.1$ and optimize the the intensity μ . When $N \leq 10^{4.8}$, no net random bits can be generated. One can see from Fig. 5(a) that the optimal q starts from 0.14 and drops down close to 0 with the increase in N . This is consistent with the intuition that $q_z = 1 - 2q \rightarrow 1$ as N goes to infinity. Note that the optimization of q assume that $q_x = q_y$ in the tomography-based protocol. In general, one can optimize q_x and q_y separately.

We also investigate how the randomness generation rate varies with the number of pulses N , as shown in Fig. 5(b). Here, we pick $p = 0.1$ and optimize both the intensity μ and the tomography testing parameter q . One can see that no randomness can be obtained as $N \leq 10^{4.8}$, beyond which point the rate increases with N . This increase is mainly contributed by the increase ratio of the Z -basis measurement as N becomes large. This is similar to the biased basis case in QKD [36].

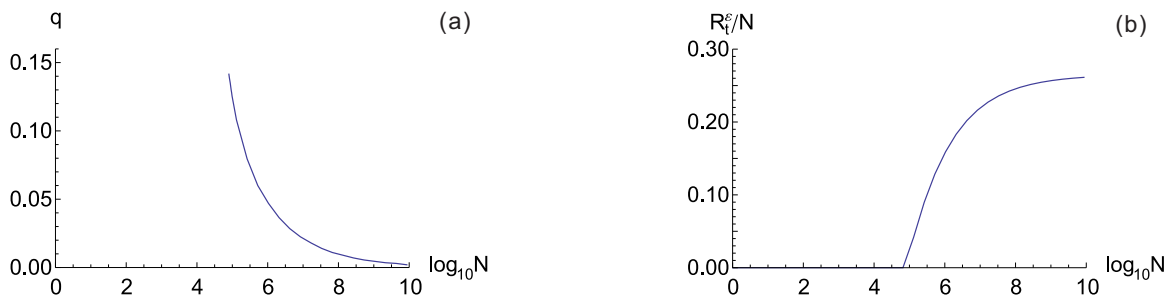


FIG. 5. The performance of the tomography-based SIQRNG by optimizing the basis selection parameter q , where we pick $p = 0.1$ and have the intensity parameter μ optimized. (a) shows the optimal value of q vs the number of pulses N . (b) illustrates the randomness generate rate R_t^E/N vs the number of pulses N .

VII. DISCUSSION

In this paper, we propose a framework for SIQRNG via measuring coherence of an unknown quantum state. We show that the uncertainty-relation-based SIQRNG is essentially related to estimating the relative entropy of coherence with a coherence witness. Furthermore, we propose a SIQRNG scheme based on state tomography. By simulating a typical QRNG setup, we show that our protocol generally enjoys a higher randomness generation rate than the uncertainty-relation-based ones.

The security analysis of QRNG is very similar to that of QKD. The mathematical definition of security in two tasks is essentially the same. For example, privacy amplification in QKD is closely related to the randomness extraction in QRNG. In practice, there are mainly two differences between them. (a) QKD involves two legitimate parties Alice and Bob; thus it requires error correction to ensure the consistency of the random numbers shared between them; whereas QRNG only involves one party, and hence error correction is unnecessary. (b) Local randomness used for encoding and basis selection is free in QKD, whereas randomness is a resource in QRNG. Nevertheless, most of the security analysis techniques in QRNG, including ours, are borrowed from QKD. In terms of the security analysis, the i.i.d. assumption in QRNG is equivalent to the collective attack assumption in QKD, whereas the non-i.i.d. scenario in QRNG corresponds to the coherent attack assumption in QKD. The difference between the two attacks vanishes when the data size goes to infinity [31], and we would expect the same deduction for our framework. In order to link coherence with randomness in the non-i.i.d. case, one needs to consider the one-shot coherence distillation [37, 38]. This is an interesting subject for future study.

Along the direction of this paper, one can realize a SIQRNG scheme by designing the coherence witness that is adapted to specific experimental conditions. Besides, it is promising to extend the framework to high-dimensional QRNG, e.g., schemes based on continuous variables [12] or laser phase fluctuations [39]. A possible challenge of this extension is the development of the high-dimensional squashing model. As the resource framework of coherence is related to the security proof of QRNG, it is also interesting to investigate whether similar relation exists between coherence and QKD.

ACKNOWLEDGMENTS

We thank Zhen Zhang for fruitful discussions. This work was supported by the National Natural Science Foundation of China Grants No. 11674193 and No. 11875173 and the National Key R&D Program of China Grants No. 2017YFA0303900 and No. 2017YFA0304004.

Appendix A: Smooth min-entropy

In this appendix, we provide the definition of smooth min-entropy [29].

Definition 2. Given a bipartite density operator ρ_{AE} , the min-entropy of A conditioned on E is defined as

$$H_{\min}(A|E)_{\rho_{AE}} \equiv -\min_{\sigma_E} D_{\infty}(\rho_{AE} || \mathbb{I} \otimes \sigma_E) \quad (\text{A1})$$

where the minimization ranges over all normalized density operators σ_E on E and

$$D_{\infty}(\tau || \tau') \equiv \min\{\lambda \in \mathbb{R} : \tau \leq 2^{\lambda} \tau'\}. \quad (\text{A2})$$

Then the smooth min-entropy of A conditioned on E is defined as

$$H_{\min}^{\varepsilon}(A|E)_{\rho_{AE}} \equiv \sup_{\rho'_{AE}} H_{\min}(A|E)_{\rho'_{AE}}, \quad (\text{A3})$$

where the supremum ranges over all density operators ρ'_{AE} which are ε -close to ρ_{AE} . Normally, the distance between ρ'_{AE} and ρ_{AE} can be measured by Bures distance $\|\rho - \sigma\|_B = \sqrt{2 - F(\rho, \sigma)}$, where $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$ and $\|\cdot\|_1$ is the l_1 -norm.

Appendix B: Derivation of Eq. (7)

With the i.i.d. assumption, the amount of randomness from transmitted quantum states is given by

$$R^{\varepsilon_1}(K_z) = \min_{\rho_{AE}} H_{\min}^{\varepsilon_1}(A|E)_{\Delta_{Z^{\otimes n_z}}^A(\rho_{AE}^{\otimes n_z})}. \quad (\text{B1})$$

As $n_z \geq \frac{8}{5} \log_2 \frac{2}{\varepsilon_1^2}$, the smooth min-entropy can be lower bounded by the conditional von Neumann entropy, $H(A|E)_{\rho_{AE}} = H(\rho_{AE}) - H(\rho_E)$ [32],

$$H_{\min}^{\varepsilon_1}(A|E)_{\Delta_{Z^{\otimes n_z}}^A(\rho_{AE}^{\otimes n_z})} \geq n_z H(A|E)_{\Delta_Z^A(\rho_{AE})} - \sqrt{n_z} \delta(\varepsilon_1, \eta), \quad (\text{B2})$$

where $\delta(\varepsilon_1, \eta) = 4(\log_2 \eta) \sqrt{\log_2 \frac{2}{\varepsilon_1^2}}$ and $\eta \leq \sqrt{2^{-H_{\min}(A|E)_{\Delta_Z^A(\rho_{AE})}} + \sqrt{2^{H_{\max}(A|E)_{\Delta_Z^A(\rho_{AE})}}} + 1}$, with H_{\min} and H_{\max} being the min-entropy and maximal-entropy, respectively. Here, $H_{\min}(A|E)_{\Delta_Z^A(\rho_{AE})} \geq 0$ and $H_{\max}(A|E)_{\Delta_Z^A(\rho_{AE})} \leq 1$ and hence one has $\eta \leq 2 + \sqrt{2}$. Thus $\delta(\varepsilon_1, \eta) \leq k \sqrt{\log_2 \frac{2}{\varepsilon_1^2}}$, with $k = 4 \log_2(2 + \sqrt{2}) \approx 7.09$. Therefore, inserting Eq. (B2) into Eq. (B1), one has

$$R^{\varepsilon_1}(K_z) \geq n_z \min_{\rho_{AE}} H(A|E)_{\Delta_Z^A(\rho_{AE})} - 7.09 \sqrt{n_z \log_2 \frac{2}{\varepsilon_1^2}}. \quad (\text{B3})$$

Appendix C: Partial derivatives of $C(\rho_A)$

In this appendix, we analyze the partial derivatives of $C(\rho_A)$. From Eq. (15),

$$C(\rho_A) = H(p_z) - H\left(\frac{p_o + 1}{2}\right), \quad (\text{C1})$$

where $p_o = \sqrt{4(p_x^2 + p_y^2 + p_z^2 - p_x - p_y - p_z) + 3}$. Thus,

$$\begin{aligned} \frac{\partial C(\rho_A)}{\partial p_z} &= \frac{\partial}{\partial p_z} \left[H(p_z) - H\left(\frac{p_o + 1}{2}\right) \right] \\ &= \frac{2p_z - 1}{p_o} \log_2 \left(\frac{1 + p_o}{1 - p_o} \right) - \log_2 \left(\frac{p_z}{1 - p_z} \right), \\ \frac{\partial C(\rho_A)}{\partial p_x} &= \frac{2p_x - 1}{p_o} \log_2 \left(\frac{1 + p_o}{1 - p_o} \right), \\ \frac{\partial C(\rho_A)}{\partial p_y} &= \frac{2p_y - 1}{p_o} \log_2 \left(\frac{1 + p_o}{1 - p_o} \right). \end{aligned} \quad (\text{C2})$$

By analyzing above equations, for $j \in \{x, y, z\}$, $\frac{\partial^2 C(\rho_A)}{\partial p_j^2} \geq 0$. Therefore, function $\frac{\partial C(\rho_A)}{\partial p_j}$ is nondecreasing with p_j . Thus,

$$\begin{aligned} \frac{\partial C(\rho_A)}{\partial p_j} &= 0 & p_j &= 1/2, \\ &\leq 0 & p_j &\leq 1/2, \\ &\geq 0 & p_j &\geq 1/2. \end{aligned}$$

Appendix D: Simulation model

In this appendix, we analyze a simulation model with a practical experimental setup. Consider a practical source, consisting of a laser and a polarization modulator, which emits N coherent-state pulses with an intensity of μ_0 . We assume the quantum state is prepared to be $|+\rangle$, which is then transmitted through a depolarization channel. Thus, the received quantum state can be described by $\rho_A = p\frac{1}{2} + (1-p)|+\rangle\langle+|$, where $p \in [0, 1]$. The readout system consists of a polarization rotator (used for basis selection), a polarization beam splitter, and two threshold detectors with the same detection efficiencies. Denote the total transmittance by η , including the detector efficiency and the coupling efficiency between the source and the detector. Here, we ignore the detection caused by dark counts since for QRNG, dark counts are normally negligible comparing to η .

In the following, we aim to evaluate the expected value of all the directly obtained experimental statistics. They include the number of total click events, n_j , the number of single-click events of the two detectors, n_j^0, n_j^1 , and the number of double-click events, n_j^d , when measuring in the $j \in \{X, Y, Z\}$ basis.

Note that the number of photons of the coherent-state pulse follows the Poisson distribution, $P(n) = \frac{e^{-\mu_0} \mu_0^n}{n!}$. With the total transmittance of the system η , the total number of clicks in the X basis is

$$\begin{aligned} n_x &= Nq \sum_n e^{-\mu_0} \frac{\mu_0^n}{n!} [1 - (1-\eta)^n] \\ &= Nq(1 - e^{-\eta\mu_0}), \end{aligned} \quad (D1)$$

where q is the probability of selecting the X basis. Similarly, one has

$$n_y = Nq(1 - e^{-\eta\mu_0}), \quad (D2)$$

$$n_z = N(1 - 2q)(1 - e^{-\eta\mu_0}), \quad (D3)$$

Denote the probability of double clicks when emitting m photons and measuring in the $j \in \{X, Y, Z\}$ basis by $p_{doub}^{i,m}$. Since the polarization of the input state is $p\frac{1}{2} + (1-p)|+\rangle\langle+|$, in which only the component $\frac{1}{2}$ may result in the double-click events, then,

$$\begin{aligned} p_{doub}^{x,m} &= \frac{p}{2^m} \sum_k C_m^k [1 - (1-\eta)^k] [1 - (1-\eta)^{m-k}] \\ &= \frac{p}{2^m} \sum_k [C_m^k - C_m^k (1-\eta)^k - C_m^k (1-\eta)^{m-k} + C_m^k (1-\eta)^m] \\ &= \frac{p}{2^m} [2^m + (1-\eta)^m 2^m - 2(2-\eta)^m] \\ &= p(1 + (1-\eta)^m - 2(1 - \frac{\eta}{2})^m). \end{aligned} \quad (D4)$$

Meanwhile, for measurement basis Y and Z , both component $\frac{1}{2}$ and $|+\rangle\langle+|$ of ρ_A result in the double-click events with equal probability, thus one has

$$\begin{aligned} p_{doub}^{y,m} &= p_{doub}^{z,m} = \frac{1}{2^m} \sum_k C_m^k [1 - (1-\eta)^k] [1 - (1-\eta)^{m-k}] \\ &= 1 + (1-\eta)^m - 2(1 - \frac{\eta}{2})^m. \end{aligned} \quad (D5)$$

Thus, the total number of double clicks in the X basis is

$$\begin{aligned} n_x^d &= Nq \sum_m e^{-\mu_0} \frac{\mu_0^m}{m!} p_{doub}^{x,m} \\ &= Nqp(1 + e^{-\eta\mu_0} - 2e^{-\frac{\eta\mu_0}{2}}). \end{aligned} \quad (D6)$$

And similarly, one has

$$\begin{aligned} n_y^d &= Nq \sum_m e^{-\mu_0} \frac{\mu_0^m}{m!} p_{doub}^{y,m} \\ &= Nq(1 + e^{-\eta\mu_0} - 2e^{-\frac{\eta\mu_0}{2}}), \end{aligned} \quad (D7)$$

$$\begin{aligned} n_z^d &= N(1 - 2q) \sum_m e^{-\mu_0} \frac{\mu_0^m}{m!} p_{doub}^{z,m} \\ &= N(1 - 2q)(1 + e^{-\eta\mu_0} - 2e^{-\frac{\eta\mu_0}{2}}), \end{aligned} \quad (D8)$$

Now we evaluate the number of single-click events corresponding to outcome $|+\rangle$ and $|-\rangle$. Note that, for ρ_A , the component $|+\rangle\langle+|$ never results in outcome $|-\rangle$, whereas the component $\frac{1}{2}$ contributes to the single-click events of $|+\rangle$ and $|-\rangle$ with the equal probability. Thus, one has

$$\begin{aligned} n_x^0 &= (1-p)n_x + \frac{1}{2}(pn_x - n_x^d) \\ &= Nq(1-p - e^{-\eta\mu_0} + pe^{-\frac{\eta\mu_0}{2}}), \\ n_x^1 &= \frac{1}{2}(pn_x - n_x^d) \\ &= Nqp(e^{-\frac{\eta\mu_0}{2}} - e^{-\eta\mu_0}). \end{aligned} \tag{D9}$$

For measurement basis Y and Z , both the component $|+\rangle\langle+|$ and $\frac{1}{2}$ contributes to the single-click events of the two outcomes with the equal probability. Thus, one has

$$\begin{aligned} n_y^0 &= n_y^1 = \frac{1}{2}(n_y - n_y^d) \\ &= Nq(e^{-\frac{\eta\mu_0}{2}} - e^{-\eta\mu_0}), \\ n_z^0 &= n_z^1 = \frac{1}{2}(n_z - n_z^d) \\ &= N(1-2q)(e^{-\frac{\eta\mu_0}{2}} - e^{-\eta\mu_0}). \end{aligned} \tag{D10}$$

Combining the above results, the experimental obtained statistics are given by

$$\begin{aligned} n_x^0 &= Nq(1-p - e^{-\eta\mu_0} + pe^{-\frac{\eta\mu_0}{2}}), \\ n_x^1 &= Nqp(e^{-\frac{\eta\mu_0}{2}} - e^{-\eta\mu_0}), \\ n_x^d &= Nqp(1 + e^{-\eta\mu_0} - 2e^{-\frac{\eta\mu_0}{2}}), \\ n_y^0 &= n_y^1 = Nq(e^{-\frac{\eta\mu_0}{2}} - e^{-\eta\mu_0}), \\ n_y^d &= Nq(1 + e^{-\eta\mu_0} - 2e^{-\frac{\eta\mu_0}{2}}), \\ n_z^0 &= n_z^1 = N(1-2q)(e^{-\frac{\eta\mu_0}{2}} - e^{-\eta\mu_0}), \\ n_z^d &= N(1-2q)(1 + e^{-\eta\mu_0} - 2e^{-\frac{\eta\mu_0}{2}}), \\ n_x &= Nq(1 - e^{-\eta\mu_0}), \\ n_y &= Nq(1 - e^{-\eta\mu_0}), \\ n_z &= N(1-2q)(1 - e^{-\eta\mu_0}). \end{aligned} \tag{D11}$$

Inserting these expressions into Eq. (16), one has

$$\begin{aligned} \frac{1-p - e^{-\eta\mu_0} + pe^{-\frac{\eta\mu_0}{2}}}{1 - e^{-\eta\mu_0}} &\leq p_x \leq \frac{1 - (1-p)e^{-\eta\mu_0} - pe^{-\frac{\eta\mu_0}{2}}}{1 - e^{-\eta\mu_0}}, \\ \frac{e^{-\frac{\eta\mu_0}{2}} - e^{-\eta\mu_0}}{1 - e^{-\eta\mu_0}} &\leq p_y \leq \frac{1 - e^{-\frac{\eta\mu_0}{2}}}{1 - e^{-\eta\mu_0}}, \\ \frac{e^{-\frac{\eta\mu_0}{2}} - e^{-\eta\mu_0}}{1 - e^{-\eta\mu_0}} &\leq p_z \leq \frac{1 - e^{-\frac{\eta\mu_0}{2}}}{1 - e^{-\eta\mu_0}}. \end{aligned} \tag{D12}$$

Following the analysis of the squashing model and statistical fluctuations, the worst case expectation values of p_x , p_y and p_z are thus given by

$$\begin{aligned} \bar{p}_x &= P_x^L - \theta = \frac{1-p - e^{-\eta\mu_0} + pe^{-\frac{\eta\mu_0}{2}}}{1 - e^{-\eta\mu_0}} - \theta, \\ \bar{p}_y &= P_y^L - \theta = \frac{e^{-\frac{\eta\mu_0}{2}} - e^{-\eta\mu_0}}{1 - e^{-\eta\mu_0}} - \theta, \\ \bar{p}_z &= P_z^L - \theta = \frac{e^{-\frac{\eta\mu_0}{2}} - e^{-\eta\mu_0}}{1 - e^{-\eta\mu_0}} - \theta. \end{aligned} \tag{D13}$$

Inserting \bar{p}_j into Eq. (19) or (23), one can estimate the amount of extractable randomness from the raw random bits.

-
- [1] M. Born, *Zeitschrift für Physik* **37**, 863 (1926).
- [2] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, *npj Quantum Information* **2**, 16021 (2016), review Article.
- [3] M. Herrero-Collantes and J. C. Garcia-Escartin, *Reviews of Modern Physics* **89**, 015004 (2017).
- [4] B. Peter, “The nsas work to make crypto worse and better,” <https://arstechnica.com/security/2013/09/the-nsas-work-to-make-crypto-worse-and-better/>.
- [5] X. Ma, F. Xu, H. Xu, X. Tan, B. Qi, and H.-K. Lo, *Phys. Rev. A* **87**, 062327 (2013).
- [6] 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, *et al.*, *Nature* **464**, 1021 (2010).
- [7] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenbergh, R. F. Vermeulen, R. N. Schouten, C. Abellán, *et al.*, *Nature* **526**, 682 (2015).
- [8] C. A. Miller and Y. Shi, arXiv preprint arXiv:1411.6608 (2014).
- [9] R. Arnon-Friedman, R. Renner, and T. Vidick, arXiv preprint arXiv:1607.01797 (2016).
- [10] Z. Cao, H. Zhou, and X. Ma, *New Journal of Physics* **17**, 125011 (2015).
- [11] Z. Cao, H. Zhou, X. Yuan, and X. Ma, *Physical Review X* **6**, 011020 (2016).
- [12] D. G. Marangon, G. Vallone, and P. Villoresi, *Phys. Rev. Lett.* **118**, 060503 (2017).
- [13] F. Bischof, H. Kampermann, and D. Bruß, *Phys. Rev. A* **95**, 062305 (2017).
- [14] J. B. Brask, A. Martin, W. Esposito, R. Houlmann, J. Bowles, H. Zbinden, and N. Brunner, *Physical Review Applied* **7**, 054018 (2017).
- [15] T. Van Himbeek, E. Woodhead, N. J. Cerf, R. García-Patrón, and S. Pironio, *Quantum* **1**, 33 (2017).
- [16] X. Yuan, H. Zhou, Z. Cao, and X. Ma, *Phys. Rev. A* **92**, 022124 (2015).
- [17] X. Yuan, Q. Zhao, D. Girolami, and X. Ma, arXiv preprint arXiv:1605.07818 (2016).
- [18] T. Baumgratz, M. Cramer, and M. B. Plenio, *Phys. Rev. Lett.* **113**, 140401 (2014).
- [19] C. Napoli, T. R. Bromley, M. Cianciaruso, M. Piani, N. Johnston, and G. Adesso, *Physical review letters* **116**, 150502 (2016).
- [20] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, *Quantum Inf. Comput.* **4**, 325 (2004).
- [21] X. Ma, C.-H. F. Fung, and H.-K. Lo, *Phys. Rev. A* **76**, 012307 (2007).
- [22] N. J. Beaudry, T. Moroder, and N. Lütkenhaus, *Physical review letters* **101**, 093601 (2008).
- [23] T. Tsurumarui and K. Tamaki, *Physical Review A* **78**, 032302 (2008).
- [24] P. W. Shor and J. Preskill, *Physical review letters* **85**, 441 (2000).
- [25] K. M. R. Audenaert and M. B. Plenio, *New Journal of Physics* **8**, 266 (2006).
- [26] J. Eisert, F. G. Brandão, and K. M. Audenaert, *New Journal of Physics* **9**, 46 (2007).
- [27] O. Gühne, M. Reimpell, and R. Werner, *Physical review letters* **98**, 110502 (2007).
- [28] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, *Nature Physics* **6**, 659 (2010).
- [29] R. König, R. Renner, and C. Schaffner, *IEEE Transactions on Information theory* **55**, 4337 (2009).
- [30] R. Impagliazzo, L. A. Levin, and M. Luby, in *Proceedings of the twenty-first annual ACM symposium on Theory of computing* (ACM, 1989) pp. 12–24.
- [31] M. Christandl, R. König, and R. Renner, *Physical review letters* **102**, 020504 (2009).
- [32] M. Tomamichel, R. Colbeck, and R. Renner, *IEEE Transactions on Information Theory* **55**, 5840 (2009).
- [33] C.-H. F. Fung, H. Chau, and H.-K. Lo, *Physical Review A* **84**, 020303 (2011).
- [34] W. Hoeffding, *Journal of the American statistical association* **58**, 13 (1963).
- [35] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo, *Phys. Rev. A* **72**, 012326 (2005).
- [36] C.-H. F. Fung, X. Ma, and H. F. Chau, *Phys. Rev. A* **81**, 012318 (2010).
- [37] Q. Zhao, Y. Liu, X. Yuan, E. Chitambar, and X. Ma, *Phys. Rev. Lett.* **120**, 070403 (2018).
- [38] K. Bu, U. Singh, S.-M. Fei, A. K. Pati, and J. Wu, *Phys. Rev. Lett.* **119**, 150405 (2017).
- [39] H. Zhou, X. Yuan, and X. Ma, *Phys. Rev. A* **91**, 062316 (2015).