

This is a repository copy of *Homodyne-based quantum random number generator at 2.9 Gbps secure against quantum side-information*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/169326/>

Version: Accepted Version

Article:

Gehring, Tobias, Lupo, Cosmo orcid.org/0000-0002-5227-4009, Kordts, Arne et al. (6 more authors) (2021) Homodyne-based quantum random number generator at 2.9 Gbps secure against quantum side-information. Nature Communications. 605. ISSN 2041-1723

<https://doi.org/10.1038/s41467-020-20813-w>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Ultra-fast real-time quantum random number generator with correlated measurement outcomes and rigorous security certification

Tobias Gehring,^{1,*} Cosmo Lupo,^{2,3} Arne Kordts,¹ Dino Solar Nikolic,¹ Nitin Jain,¹
Tobias Rydberg,¹ Thomas B. Pedersen,⁴ Stefano Pirandola,² and Ulrik L. Andersen^{1,†}

¹*Center for Macroscopic Quantum States (bigQ), Department of Physics,
Technical University of Denmark, Fysikvej, 2800 Kgs. Lyngby, Denmark*

²*Department of Computer Science, University of York, York YO10 5GH, United Kingdom*

³*Department of Physics and Astronomy, University of Sheffield, United Kingdom*

⁴*Cryptomathic A/S, Jaegersgade 118, 8000 Aarhus C, Denmark*

(Dated: April 1, 2020)

Quantum random number generators (QRNGs) promise perfectly unpredictable random numbers. However, the security certification of the random numbers in form of a stochastic model often introduces assumptions that are either hardly justified or indeed unnecessary. Two important examples are the restriction of an adversary to the classical regime as well as negligible correlations between consecutive measurement outcomes. Additionally, non-rigorous system characterization opens a security loophole. In this work we experimentally realize a QRNG that does not rely on the aforementioned assumptions and whose stochastic model is established by a rigorous – metrological – approach. Based on quadrature measurements of vacuum fluctuations, we demonstrate a real-time random number generation rate of 8 GBit/s. Our security certification approach offers a number of practical benefits and will therefore find widespread applications in quantum random number generators. In particular, our generated random numbers are well suited for today’s conventional and quantum cryptographic solutions.

Random numbers are ubiquitous in modern society. They are used in numerous applications ranging from cryptography, simulations and gambling, to fundamental tests of physics. For most of these applications, the quality of the random numbers is of utmost importance. If, for instance, cryptographic keys originating from random numbers are predictable, it will have severe consequences for the security of the internet. To ensure the security of cryptographic encryption, the random numbers must be truly random, i.e. completely unpredictable to everyone and thus private, and their randomness must be certified by establishing a stochastic model [1, 2].

True unpredictability and privacy of the generated numbers can be attained through a quantum measurement process: By performing a projective measurement on a pure quantum state, and ensuring that the state is not an eigenstate of the measurement projector, the outcome is unpredictable and thus true random numbers can be generated. Moreover, since a pure state cannot be correlated to any other state in the universe, the generated numbers will be private.

Numerous different types of quantum random number generators (QRNGs) have been devised exploiting the quantum uncertainty in photon counting measurements, phase measurements or quadrature measurements [3–5]. One particular approach of increasing interest due to its high practicality is the optical quadrature measurements of the vacuum state by means of a simple homodyne detection [6–8]. This approach combines simplicity, cost-

effectiveness, chip-integrability and extraordinary high generation speed.

Previous QRNG implementations – independently of the method they are based on – fall short of one or more of the following critical issues:

Adversaries are often assumed to have restricted power: It is often assumed that they only have access to classical side-information and thus have no quantum capabilities. Recently, this issue has been addressed for a source-independent QRNG [8, 9], which however requires a more complex measurement apparatus than simple homodyne detection.

Furthermore, it has often been assumed that measurements are uncorrelated in time [6–8, 10–18], despite the fact that the finite bandwidth of a real detection system introduces correlations. Aliasing in the sampling procedure may minimize correlations as well as suitable post-processing algorithms, however such measures usually throttle the overall rate considerably or remove the correlations only partially.

Most previous implementations did not use a conservative and rigorous approach – a metrology-grade approach [19] – to characterize the parameters of the stochastic model that determines the amount of randomness. A rigorous characterization of the system is however of utmost importance as any parameter uncertainty introduces a non-zero probability for system failure, i.e. the probability that the actual device does not follow the stochastic model describing the underlying physical random number generation process. Knowing the failure probability for the system is critical to its certification.

Finally, high-speed (GBit) randomness extraction using an information theoretically secure randomness extractor has only been demonstrated recently [16–18, 20]

* tobias.gehring@fysik.dtu.dk

† ulrik.andersen@fysik.dtu.dk

and thus many reported QRNGs either achieve only moderate speeds or do not even extract random numbers in real-time.

Here, we report on a QRNG which solves all these aforementioned issues simultaneously for the first time. Using a QRNG based on the quadrature measurement of vacuum fluctuations, we 1) compute a lower bound on the extractable randomness against a quantum-enabled adversary, 2) account for correlated samples resulting from the finite bandwidth of the measurement apparatus, and 3) perform a metrology-grade characterization of the measuring homodyne detector system to establish the stochastic model. Finally, as a result, we produce random numbers in real-time with a rate of 8 Gbit/s using a Toeplitz randomness extractor on a fast field-programmable-gate-array (FPGA).

I. SETTING THE STAGE

A schematic of our QRNG is shown in Fig. 1. An arbitrary quadrature of the vacuum state is measured using a balanced homodyne detector comprising a bright reference beam, a symmetric beam splitter and two photodiodes [21]. The measurement outcomes ideally are random with a Gaussian distribution associated with the Gaussian Wigner function of the vacuum quantum state [22]. The measured distribution, however, contains two additional independent Gaussian noise sources; excess optical noise and electronic noise, thereby contributing two side channels. These must be accounted for in estimating the min-entropy of the source.

The amount of quantum randomness that can be extracted from the homodyne measurement of vacuum fluctuations is given by the leftover hash lemma [23, 24]

$$\ell \leq NH_{\min}(X|E) - \log \frac{1}{\epsilon_{\text{hash}}^2} . \quad (1)$$

Here $H_{\min}(X|E)$ is the min-entropy of a single measurement outcome drawn from a random variable X conditioned on the quantum side information E held by an adversary, N is the number of aggregated samples and ϵ_{hash} is the distance between a perfectly uniform random string and the string produced by a randomness extractor. It is therefore clear that we need to find the min-entropy of our practical – thus imperfect – realization in order to bound the amount of randomness. We achieve this in a two-step approach: First we theoretically derive a bound for the min-entropy using a realistic security model and express it in terms of experimentally accessible parameters. Second, we experimentally deduce these parameters through a metrology-grade characterization [19]. Using such an approach, we find the worst-case min-entropy compatible with the confidence intervals of our characterization and calibration measurements, thereby obtaining a string of ϵ -random bits that are trustworthy with the same level of confidence.

II. SECURITY ANALYSIS

The security analysis of the QRNG is made under the assumption that the quantum noise is Gaussian and stationary. Therefore, the QRNG follows a device-dependent security model, i.e. an adversary cannot change the system after system characterization has been performed. Stationarity is formally expressed by a Wigner function that takes the form of a stationary Gaussian distribution in phase-space. As we are dealing with Gaussian states, the Wigner function is completely characterized by the first and second moments of the field quadratures.

In our analysis we assume that homodyne detection is defined on a single optical mode, which at a given time is characterized by the field quadratures \hat{q} and \hat{p} . We also make a further assumption that the state is symmetric under a rotation of the quadratures

$$\hat{q} \rightarrow \hat{q} \cos \theta + \hat{p} \sin \theta , \quad (2)$$

$$\hat{p} \rightarrow \hat{p} \cos \theta - \hat{q} \sin \theta , \quad (3)$$

for all $\theta \in [0, 2\pi]$.

In the following Sections, we first assess the security of a source emitting i.i.d. (independent and identically distributed) signals, i.e., a source of infinite bandwidth. We then extend the security analysis to a source with finite bandwidth that emits correlated (non-i.i.d.) signals at different times.

A. IID limit

First consider an ideal i.i.d. scenario with unlimited bandwidth. As explained above, we assume that the source emits a Gaussian state of light that is symmetric in phase space. In the i.i.d. limit this corresponds to a source of thermal light. We therefore assume a thermal source characterized by a mean photon number per mode of n . Each mode is measured independently by homodyne detection. In other words, our security analysis holds under the assumption that a potential eavesdropper performs a collective Gaussian attack that preserves the above symmetry in the phase space.

For a thermal state ρ , the first moments of the field quadratures vanish, and the covariance matrix (CM) is

$$V_{\text{thermal}} = \begin{pmatrix} \langle \hat{q}^2 \rangle & \frac{1}{2} \langle \hat{q}\hat{p} + \hat{p}\hat{q} \rangle \\ \frac{1}{2} \langle \hat{p}\hat{q} + \hat{q}\hat{p} \rangle & \langle \hat{p}^2 \rangle \end{pmatrix} \quad (4)$$

$$= \begin{pmatrix} 1 + 2n & 0 \\ 0 & 1 + 2n \end{pmatrix} , \quad (5)$$

where we, as a matter of convention, put the variance of the vacuum equal to 1. In the equation above we use $\langle \hat{O} \rangle := \text{tr}(\rho \hat{O})$ for operator \hat{O} .

For such a state, the output X of ideal homodyne detection is a continuous (real-valued) variable whose probability distribution is

$$p_X(x) = G(x; 0, g^2(1 + 2n)) , \quad (6)$$

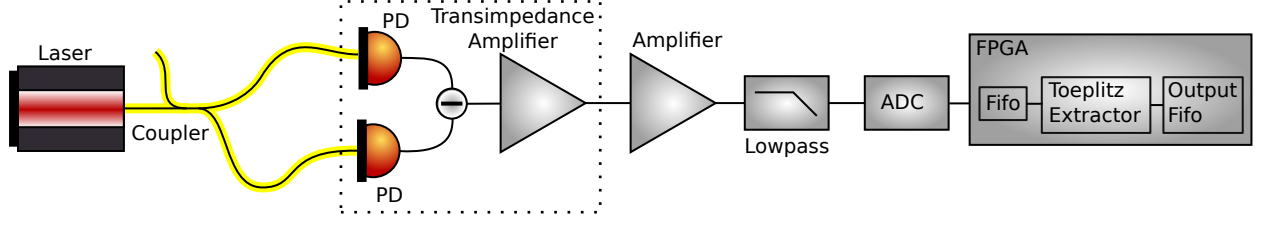


FIG. 1. Schematic of the QRNG. A 1.6 mW 1550 nm laser beam was split into two by a 3 dB fiber coupler and detected by a home-made homodyne detector based on a MAR-6 microwave amplifier from Minicircuits and two 120 μm InGaAs photo diodes (PD). The output of the detector was amplified with another microwave amplifier, lowpass filtered at 400 MHz and digitized with a 16 bit 1 GS/s analog-to-digital (ADC) converter. The ADC output was read by a Xilinx Kintex UltraScale FPGA. The ADC and FPGA were hosted by a PCI Express card from 4DSP (Abaco). The FPGA was used for real-time randomness extraction based on Toeplitz hashing. Fifo: First-in-first-out buffer.

where g is a gain factor and

$$G(x; \mu, v^2) = \frac{1}{\sqrt{2\pi}v} e^{-\frac{(x-\mu)^2}{2v^2}} \quad (7)$$

denotes a Gaussian in the variable x , with mean μ and variance v^2 .

To account for quantum side information we assume that a malicious adversary holds a purification of the

thermal state emitted by the source. It is well known that the purification of a thermal state can be assumed to be a two-mode squeezed vacuum (TMSV) state without loss of generality (more precisely up to isometries over the environment). Thereby one optical mode of this TMSV state, characterized by the field quadratures \hat{q}_e and \hat{p}_e , is controlled by the adversary. The TMSV state is a Gaussian state with zero mean and CM [22]

$$V = \begin{pmatrix} \langle \hat{q}^2 \rangle & \frac{1}{2} \langle \hat{q}\hat{p} + \hat{p}\hat{q} \rangle & \langle \hat{q}\hat{q}_e \rangle & \langle \hat{q}\hat{p}_e \rangle \\ \frac{1}{2} \langle \hat{p}\hat{q} + \hat{q}\hat{p} \rangle & \langle \hat{p}^2 \rangle & \langle \hat{p}\hat{q}_e \rangle & \langle \hat{p}\hat{p}_e \rangle \\ \langle \hat{q}_e\hat{q} \rangle & \langle \hat{q}_e\hat{p} \rangle & \langle \hat{q}_e^2 \rangle & \frac{1}{2} \langle \hat{q}_e\hat{p}_e + \hat{p}_e\hat{q}_e \rangle \\ \langle \hat{p}_e\hat{q} \rangle & \langle \hat{p}_e\hat{p} \rangle & \frac{1}{2} \langle \hat{p}_e\hat{q}_e + \hat{q}_e\hat{p}_e \rangle & \langle \hat{p}_e^2 \rangle \end{pmatrix} \quad (8)$$

$$= \begin{pmatrix} 1+2n & 0 & 2\sqrt{n(n+1)} & 0 \\ 0 & 1+2n & 0 & -2\sqrt{n(n+1)} \\ 2\sqrt{n(n+1)} & 0 & 1+2n & 0 \\ 0 & -2\sqrt{n(n+1)} & 0 & 1+2n \end{pmatrix}. \quad (9)$$

The correlations between the outcome X of ideal homodyne detection and the quantum side information held by the eavesdropper are described by the classical-quantum (CQ) state

$$\rho_{XE} = \int dx p_X(x) |x\rangle\langle x| \otimes \rho_E^x, \quad (10)$$

where $|x\rangle$ are orthogonal states used to represent the possible outcomes of homodyne detection, and the integral in Eq. (10) extends over the real line. The state ρ_E^x is the conditional state of the eavesdropper for a given measurement output value x . Assume, without loss of generality, that the quadrature \hat{q} is measured. It is then straightforward to compute the first moment of the field quadratures of ρ_E^x :

$$\begin{pmatrix} \langle \hat{q}_e \rangle \\ \langle \hat{p}_e \rangle \end{pmatrix} = \begin{pmatrix} \frac{2\sqrt{n(n+1)}}{g(1+2n)} x \\ 0 \end{pmatrix}, \quad (11)$$

as well as the CM

$$\begin{pmatrix} \langle \hat{q}_e^2 \rangle & \frac{1}{2} \langle \hat{q}_e\hat{p}_e + \hat{p}_e\hat{q}_e \rangle \\ \frac{1}{2} \langle \hat{p}_e\hat{q}_e + \hat{q}_e\hat{p}_e \rangle & \langle \hat{p}_e^2 \rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{1+2n} & 0 \\ 0 & 1+2n \end{pmatrix}. \quad (12)$$

In our QRNG the continuous variable X is mapped into a discrete and bounded variable \bar{X} due to the use of an analog-to-digital converter (ADC). We therefore consider a model in which X is replaced by a discrete variable \bar{X} such that

$$p_{\bar{X}}(k) = \int_{I_k} dx p_X(x), \quad (13)$$

where I_k 's are d intervals that discretize the outcome of homodyne detection. In a typical setting, these d non-overlapping intervals I_k are of the form

$$I_1 = (-\infty, -R], \quad (14)$$

$$I_d = (R, \infty), \quad (15)$$

and for $k = 2, \dots, d-1$

$$I_k = (a_k - \Delta x/2, a_k + \Delta x/2], \quad (16)$$

with $a_k = -R + (k-1)\Delta x/2$ and $\Delta x = 2R/(d-2)$. This choice of the intervals reflects the way in which an ideal ADC with range R and bin size Δx operates in mapping a continuous variable into a discrete one. However ADCs are not ideal devices, and in the Supplemental Material we show how the digitization error of an ADC reduces the min-entropy.

In terms of the discrete variable \bar{X} , the correlations with the eavesdropper are then described by the state

$$\rho_{\bar{X}E} = \sum_k p_{\bar{X}}(k) |k\rangle\langle k| \otimes \rho_E^{(k)}, \quad (17)$$

with

$$\rho_E^{(k)} = \frac{1}{p_{\bar{X}}(k)} \int_{I_k} dx p_X(x) \rho_E^x. \quad (18)$$

We are now ready to quantify the secure rate of the QRNG in terms of the conditional min-entropy. Given the state $\rho_{\bar{X}E}$ in Eq. (17), the min-entropy of \bar{X} conditioned on the eavesdropper (denoted with the letter E) reads [25] [26]

$$H_{\min}(\bar{X}|E)_\rho = \sup_\gamma \left[-\log \|\gamma_E^{-1/2} \rho_{\bar{X}E} \gamma_E^{-1/2}\|_\infty \right], \quad (19)$$

where $\|\cdot\|_\infty$ denotes the operator norm (equal to the value of the maximum eigenvalue).

Since a direct computation of the min-entropy is not feasible as it requires an optimization over all density operators γ in an infinite-dimensional Hilbert space, we instead focus on finding a computable and tight lower bound. A first lower bound on the min-entropy is obtained by computing $\|\gamma_E^{-1/2} \rho_{\bar{X}E} \gamma_E^{-1/2}\|_\infty$ for a given choice of the state γ , so that we have

$$H_{\min}(\bar{X}|E)_\rho \geq -\log \|\gamma_E^{-1/2} \rho_{\bar{X}E} \gamma_E^{-1/2}\|_\infty \quad (20)$$

$$= -\log \left[\sup_k p_{\bar{X}}(k) \|\gamma_E^{-1/2} \rho_E^{(k)} \gamma_E^{-1/2}\|_\infty \right], \quad (21)$$

where the last equality holds because the eigenstates $|k\rangle$ of $\rho_{\bar{X}E}$ in Eq. (17) are mutually orthogonal. Here we set γ equal to a Gaussian state with zero mean and CM

$$\begin{pmatrix} 1 + 2(n + \delta) & 0 \\ 0 & 1 + 2(n + \delta) \end{pmatrix}, \quad (22)$$

where the parameter δ will be optimized *a posteriori* to make the bound as tight as possible. This choice for the CM is somewhat arbitrary but, as we show in the Supplemental Material, it yields a tight bound on the min-entropy.

A second lower bound is obtained by applying the triangular inequality,

$$\begin{aligned} p_{\bar{X}}(k) \|\gamma_E^{-1/2} \rho_E^{(k)} \gamma_E^{-1/2}\|_\infty \\ = \|\gamma_E^{-1/2} \int_{I_k} dx p_X(x) \rho_E^x \gamma_E^{-1/2}\|_\infty \end{aligned} \quad (23)$$

$$\leq \int_{I_k} dx p_X(x) \|\gamma_E^{-1/2} \rho_E^x \gamma_E^{-1/2}\|_\infty, \quad (24)$$

which implies

$$H_{\min}(\bar{X}|E) \geq -\log \left[\sup_k \int_{I_k} dx p_X(x) \|\gamma_E^{-1/2} \rho_E^x \gamma_E^{-1/2}\|_\infty \right]. \quad (25)$$

Since ρ_E^x and γ_E are both Gaussian states, the above lower bound can be computed using the Gibbs-representation techniques developed in Ref. [27]. Employing these techniques and additional tools, Ref. [28] derived a formula for the min-entropy. By applying this result we obtain [29]

$$\int_{I_k} dx p_X(x) \|\gamma_E^{-1/2} \rho_E^x \gamma_E^{-1/2}\|_\infty = \frac{1}{g} \frac{(n + \delta)(1 + n + \delta)}{\sqrt{2\pi\delta(2n(n + 1 + \delta) + \delta)}} \int_{I_k} dx \exp \left[\frac{-x^2}{2g^2} \frac{\delta}{2n(n + 1 + \delta) + \delta} \right]. \quad (26)$$

The supremum over k can be computed for any given

collection of intervals I_k 's. For intervals as in Eqs. (14)-(16) we obtain

$$H_{\min}(\bar{X}|E) \geq -\log \left[\frac{(n+\delta)(1+n+\delta)}{\delta} \max \left\{ \operatorname{erf} \left(\sqrt{\frac{\delta}{4n(n+1+\delta)+2\delta}} \frac{\Delta x}{2g} \right), \frac{1}{2} \operatorname{erfc} \left(\sqrt{\frac{\delta}{4n(n+1+\delta)+2\delta}} \frac{R}{g} \right) \right\} \right]. \quad (27)$$

We remark that this is in fact a family of lower bounds parameterized by δ . One should then find the optimal value of δ for which the bound is tighter.

So far we have considered quantum noise, which was quantified by the mean photon number n . If the variance also includes an additive term ζ due to classical noise, $v^2 = g^2(1+2n) + \zeta^2$, then this can be included in the above analysis by treating it as quantum noise and re-defining $n \rightarrow n + \frac{\zeta^2}{2g^2}$.

B. Beyond IID: stationary Gaussian process

Going beyond i.i.d., we now consider a more realistic scenario where the measured signal has a finite bandwidth. The security analysis of this section holds under the assumption that a potential adversary performs a Gaussian attack. This attack is not necessarily i.i.d. (unlike the collective Gaussian attack considered in the previous section), but it is assumed to be stationary, i.e., it induces a stationary Gaussian process. The assumption of stationarity is thereby in accordance with the device-dependent nature of our scheme.

In this section we first build an i.i.d. model for the non-i.i.d. process. Then we apply the results of Section II A to compute a lower bound on the min-entropy with quantum side information.

The analysis deals with two processes. One is the signal X , i.e. the homodyne measurement of the quantum state including all additive noise processes. The second is the excess noise U , i.e. all noise sources present in the measurement apart from the pure vacuum fluctuations, for instance electronic noise of the detector and intensity noise of the local oscillator laser. When a measurement is performed at a given time t , the measured signal is denoted X_t . Similarly, we denote as U_t the excess noise at time t (which we need as a theoretical tool, even though this quantity is not accessible experimentally). We assume that both X and U are stationary Gaussian processes.

Let us first consider the output X of ideal homodyne detection, and denote as σ^2 its variance. From the power spectrum $f_X(\lambda)$ we can estimate the entropy rate of the signal [30]

$$h(X) = \frac{1}{2} \log(2\pi e \sigma_X^2), \quad (28)$$

where

$$\sigma_X^2 = \frac{1}{2\pi e} 2 \int_0^{2\pi} \frac{d\lambda}{2\pi} \log[2\pi e f_X(\lambda)] \quad (29)$$

is the conditional variance. Similarly, from the power spectrum of the excess noise $f_U(\lambda)$ we obtain the entropy rate of the noise

$$h(U) = \frac{1}{2} \log(2\pi e \sigma_U^2), \quad (30)$$

where

$$\sigma_U^2 = \frac{1}{2\pi e} 2 \int_0^{2\pi} \frac{d\lambda}{2\pi} \log[2\pi e f_U(\lambda)] \quad (31)$$

is the conditional variance of the excess noise.

Because of the finite bandwidth of the measuring apparatus, both the signal X_t and excess noise U_t at a given time t are correlated with their values at previous times. To filter out the effects of these correlations we consider the probability density distribution of X_t conditioned on all past signal values,

$$p_{X_t}(x_t|x_{<t}) = G(x_t; \mu_t, \sigma_X^2), \quad (32)$$

where x_t denotes the possible values of the variable X_t at time t , $x_{<t}$ denotes the collection of values of all signals at times $t' < t$, and μ_t is the mean value. Note that μ_t depends on $x_{<t}$, but the conditional variance σ_X^2 does not depend on time. This description is consistent with the assumption of a stationary Gaussian process [30]. By definition, the conditioned variable at time t is independent on previous signal values. Therefore, we can formally describe it as the outcome of a measurement applied on an i.i.d. quantum state with the same variance. We identify (using the notation of the Section II A):

$$\sigma_X^2 \equiv g^2(1+2n). \quad (33)$$

We can then write the (unconditional) signal variance σ^2 as

$$\sigma^2 = g^2(1+2n) + \zeta, \quad (34)$$

where $\zeta = \sigma^2 - \sigma_X^2$, and the term ζ is interpreted as classical noise due to the fluctuations of the mean value μ_t , whose variance is ζ .

In summary, we have defined an effective i.i.d. model for the non-i.i.d. signal. This i.i.d. model has been obtained using the notion of conditional variance. The i.i.d. model is characterized by the parameters n and g . Therefore, we need a second equation in addition to Eq. (34) to determine the model parameters n and g as function of the measured quantities σ^2 , σ_X^2 , and σ_U^2 . This is obtained through the conditional variance of the excess noise.

For the excess noise U_t , we can similarly write the probability density distribution conditioned on past values, i.e.,

$$p_{U_t}(u_t|u_{<t}) = G(u_t; \nu_t, \sigma_U^2), \quad (35)$$

where u_t denotes the possible values of the variable U_t at time t , $u_{<t}$ denotes its past values, and ν_t is the mean value of U_t . The quantity of interest is the conditional excess noise variance σ_U^2 , which represents the variance of the excess noise that is virtually independent of previous noise values. We identify this variance with the variance of the excess noise in the i.i.d. model of Section II A:

$$\sigma_U^2 \equiv 2g^2n. \quad (36)$$

Endowed with Eq. (34) and (36), we are now in the position of determining the parameters n and g of the i.i.d. model of the non-i.i.d. process. Solving for n and g the coupled equations,

$$\sigma_X^2 = g^2(1 + 2n), \quad (37)$$

$$\sigma_U^2 = 2g^2n, \quad (38)$$

we obtain

$$g = \sqrt{\sigma_X^2 - \sigma_U^2}, \quad (39)$$

$$n = \frac{1}{2} \frac{\sigma_U^2}{\sigma_X^2 - \sigma_U^2}. \quad (40)$$

Finally, we need to account for the classical noise variance ζ . As discussed in Section II A, we incorporate it in the quantum noise by re-defining

$$n \rightarrow n + \frac{\zeta}{2g^2} \quad (41)$$

$$= \frac{1}{2} \frac{\sigma_U^2}{\sigma_X^2 - \sigma_U^2} + \frac{1}{2} \frac{\sigma^2 - \sigma_X^2}{\sigma_X^2 - \sigma_U^2} \quad (42)$$

$$= \frac{1}{2} \frac{\sigma^2}{\sigma_X^2 - \sigma_U^2} - \frac{1}{2}. \quad (43)$$

In conclusion, this model allows us to compute a lower bound for the min-entropy of the non-i.i.d. process by inserting the above values for g [in Eq. (39)] and n [in Eq. (43)] in the min-entropy formula of Eq. (27). This is plotted in Fig. 2 for varying excess noise, ADC resolution and temporal correlations. The x-axis of the plot is the ratio of the conditional variance of the vacuum fluctuations and the excess noise, i.e. the quantum noise to excess noise ratio of the virtual i.i.d. process. Assuming that the measured homodyne signal and the excess noise have similar temporal correlations, this ratio is independent of the amount of correlations. The amount of correlations present in the system is instead characterized by the ratio σ_X^2/σ^2 which takes the value of 1 for an i.i.d. process and becomes smaller for increasing temporal correlations. For each ADC resolution the upper traces in the figure show the extractable min-entropy when almost no correlations

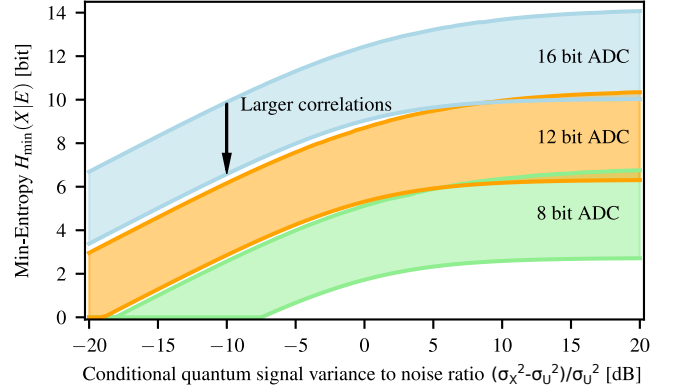


FIG. 2. Min-entropy for 8, 12 and 16 bit ADC resolution versus the ratio of conditional variance of the vacuum fluctuations and the conditional variance of the excess noise, $(\sigma_X^2 - \sigma_U^2)/\sigma_U^2$. The shaded areas indicate the regions between low correlations ($\sigma_X^2/\sigma^2 = 0.99$), upper trace, and high correlations ($\sigma_X^2/\sigma^2 = 0.1$), lower trace. The signal variance has been optimized to obtain the highest min-entropy.

are present. Obviously, stronger correlations yield lower randomness.

Similar to the result for classical side-information [12], we show that random numbers can be generated for noise treated as quantum side-information as well; and even in the large excess noise regime. This is due to the fact that relatively small vacuum fluctuations can give a substantial contribution to the entropy if the ADC resolution is sufficiently high. This property is preserved even when a large amount of temporal correlations is present in the recorded data (lower traces).

III. SYSTEM CHARACTERIZATION

Using the above results, we are now in a position to estimate the min-entropy through a metrology-grade characterization of our setup. According to the security analysis, the min-entropy can be found by determining the variance σ^2 as well as the conditional variances of the homodyne signal σ_X^2 and the excess noise σ_U^2 . To obtain a conservative, and thus reliable, estimate of the min-entropy, it is important that the determination of these parameters does not rely on any ideality assumptions of the homodyne detector. In previous studies on homodyne based QRNG, the sole presence of shot noise has been verified by characterizing its scaling with optical power. However, imperfect common-mode rejection, large intensity noise of the laser or stray light coupling into the signal port – likely to be an issue with integrated photonic chips – may unnecessarily constrain the extraction of random numbers. Furthermore this characterization method is not directly compatible with metrology-grade characterization as it is difficult to bound the estimation error on the shot noise level with a confidence interval. To circumvent these assumptions and issues we perform

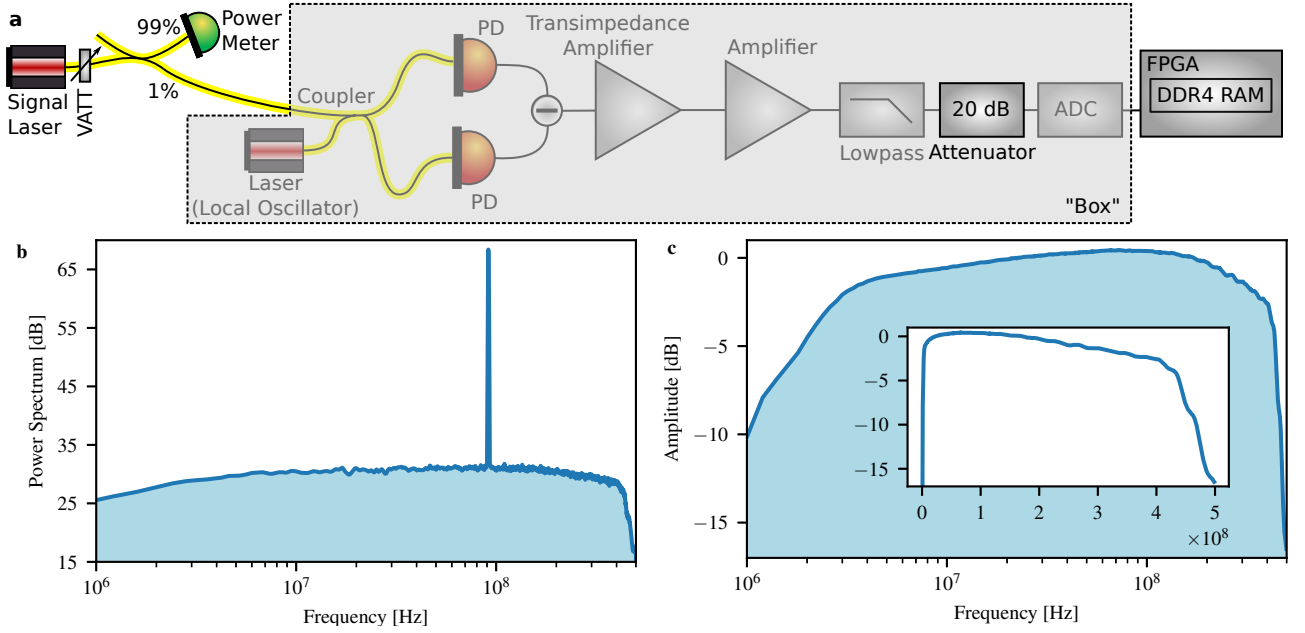


FIG. 3. Characterization of the transfer function of the detection system to obtain the vacuum fluctuations noise level. a) Experimental setup for the characterization. b) Power spectrum from a typical measurement. The transfer function is determined by the amplitude of the beat node. c) Transfer function of the homodyne detector and the electronics including the analog-to-digital converter. Inset: Transfer function with linear frequency scale.

an independent, reliable and metrology-grade characterization of the measuring device.

We basically consider the homodyne detector as a black box with an input and an output and minimal assumptions on its internal workings (see Fig 3a). Our strategy is thus to measure the transfer function of the box and to use this result to conservatively calibrate the power spectral density (PSD) of the vacuum fluctuations. This conservatively estimated result is then compared to the PSD of an actual noise measurement from which we deduce the conditional variances of the signal noise and the excess noise, and finally the min-entropy.

The transfer function of the box is measured by injecting a coherent state in form of a second laser beam (independent of the local oscillator laser) with low power P_{sig} into the signal port of the beam splitter as displayed in Fig. 3a. A typical beat signal is shown in Fig. 3b obtained by computing an averaged periodogram from the sampled signal. We record the transfer function $\text{TF}(\nu)$ by scanning the frequency of the signal laser. At each difference frequency ν we determine the power of the beat signal and normalize it to P_{sig} . At high signal-to-noise ratio the root-mean-square power of the beat signal is purely a function of the coherent state amplitude (i.e. the signal laser power) and independent of the noise properties of the two lasers and the detector.

The transfer function includes the efficiency of the interference, optical loss and the quantum efficiency of the photodiodes, as well as the frequency dependent gain of all amplifiers, the lowpass filter, and the analog bandwidth of the ADC. Since the vacuum noise was amplified

to optimally fill the range of the ADC, we used a 20 dB electrical attenuator with flat attenuation over the frequency band of interest to avoid saturation, see Fig. 3a. The result of the transfer function characterization, normalized to a maximum gain of 1 is shown in Fig. 3c. Assuming linearity of the detector we obtain the PSD of the vacuum fluctuations by multiplying the transfer function $\text{TF}(\nu)$ with the shot noise energy $\hbar\omega_L$ contained in 1 Hz bandwidth, where \hbar is Planck's constant and ω_L is the angular frequency of the local oscillator laser. By modelling the inner workings of the black box, we confirm in the Supplemental Material that with this procedure we indeed obtain a lower bound on the PSD of the vacuum fluctuations.

The conservatively estimated PSD of the vacuum fluctuations is shown in Fig. 4a together with the actually measured PSD of the signal. The spectra are clearly “colored” which indicates that the data samples are correlated and therefore non-i.i.d. This is further corroborated in Fig. 4b, where the autocorrelation of the signal is plotted. It justifies the importance of using the min-entropy relation associated with non-i.i.d. samples.

From the PSDs we calculate the three parameters for obtaining the min-entropy which are summarized in Table I. By minimizing the min-entropy over the confidence set of the estimated parameters, we obtain 10.74 bit per 16 bit sample with a failure probability of $\epsilon_{\text{PE}} = 10^{-10}$ (i.e. the probability that the actual parameters are outside the confidence intervals). To verify the Gaussian assumption in our security proof, we calculated the probability quantiles of the measured samples and compared

Parameter	Mean	Confidence interval
σ^2	3.96×10^7	0.09×10^7
σ_X^2	3.29×10^7	0.07×10^7
σ_U^2	2.49×10^7	0.06×10^7
Conditional quantum to excess noise ratio	-4.9 dB	
Temporal correlations σ_X^2/σ^2	0.83	
Min-entropy	10.74 bit	
Reduction due to ADC digitization error	1.77 bit	
Calculated secure length	2079 bit	
Extracted length	2048 bit	

TABLE I. Summary of parameters determined by the metrological characterization with their confidence intervals for $\epsilon_{PE} = 10^{-10}$: signal variance σ^2 , conditional signal variance σ_X^2 and conditional excess noise variance σ_U^2 . The calculated min-entropy minimized over the confidence intervals, the secure length according to the leftover hash lemma and the length of the extracted random sequence in the experiment.

those to the theoretical quantiles of a Gaussian distribution, see Fig. 4c.

Finally, we characterized the digitization error of our ADC which is shown in Fig. 4d. The measurement protocol is described in the Supplemental Material. Using confidence intervals, the worst case estimate of the reduction of the min-entropy due to the digitization error is 1.77 bit, thus yielding a total min-entropy of 8.97 bit. This relatively large reduction is due to the fact that our ADC is 4-way interleaved.

IV. REAL-TIME RANDOMNESS EXTRACTION

Having calculated the min-entropy, the next step is to extract random numbers. This is done by using a strong extractor based on a Toeplitz matrix hashing algorithm in which the seed can be reused [31]. We chose matrix dimensions of $n = 4096$ bits and $m = 2048$ bits, which corresponds to 256 input samples with a depth of 16 bit and an output length $m < l$, chosen such that Eq. (1) was fulfilled with $H_{\min} = 8.97$ bit and $\epsilon_{\text{hash}} < 10^{-33}$. The 16 bit samples provided by the ADC at a rate of 1 GHz are received by the FPGA in chunks of 64 bits at a rate of 250 MHz. For the algorithm implementing the Toeplitz hashing we followed the approach of Ref. [18]. Every clock cycle the 64 bits were stored in a block until n -bits were accepted, after which the next block started receiving data. For each full block, we carried out the hashing multiplication with bit-wise AND and subsequent XOR operations on the Toeplitz matrix by first splitting up the matrix into submatrices of width 16 bit, and then shifting the data through the operations. When the hashing was completed, the m -bit wide output data was stored in a register, and the next block was processed. The achieved throughput was 8 Gbit/s.

V. CONCLUSION

In conclusion, we have demonstrated a QRNG based on the measurement of vacuum fluctuations with a real-time extraction at a rate of 8 GBit/s. Our QRNG has a strong security guarantee with a failure probability of $N' \cdot \epsilon_{\text{hash}} + \epsilon_{PE} + \epsilon_{\text{seed}} = N' \cdot 10^{-33} + 10^{-10} + \epsilon_{\text{seed}}$, where N' is the number of QRNG runs in the past, ϵ_{hash} is the security parameter related to the removal of side information (see Eq. 1), $\epsilon_{PE} = 10^{-10}$ is the security parameter of the metrological grade parameter estimation and ϵ_{seed} describes the security of the random bits used for seeding the randomness extractor. Since an adversary may have access to all quantum side information from the past, ϵ_{hash} grows with time [1]. We chose a value of 10^{-33} to be able to generate Gaussian random numbers with security $\epsilon = 10^{-9}$ for a single execution of a continuous variable quantum key distribution (QKD) [5] protocol with 10^{10} transmitted quantum states even after 10 years of continuous operation of the QRNG. See the Supplemental Material for details. In our experiment the seed bits were chosen with a pseudo-random number generator, which did not allow us to give a security guarantee for ϵ_{seed} . The generated random numbers passed both the dieharder [32] and the NIST [33] statistical batteries of randomness tests.

Due to the choice of a very small ϵ_{hash} , the real-time speed of our QRNG was limited to 8 GBit/s by the input size of the Toeplitz extractor required by our FPGA implementation. Without limitations to the matrix size a speed of 8.9 GBit/s could be reached. The main limitations are the very conservative estimates of the min-entropy reduction due to the ADC digitization error and the shot noise calibration.

Nevertheless, our QRNG is perfectly suited for use in high-speed QKD links, for instance in GHz clocked discrete variable [34] as well as in high-speed continuous-variable QKD [35]. For Gaussian-modulated CVQKD the uniform random number distribution has to be converted to a Gaussian distribution which requires a larger random number generation rate. Furthermore, QKD requires composable security and a guarantee of privacy of the random numbers as provided by our system.

Further developments to guarantee reliable operation over a long time and to fulfill requirements by certification authorities would need to include power-on self-tests and online testing of the parameters in the security proof as well as the generated random numbers.

ACKNOWLEDGEMENTS

The authors acknowledge support from the Innovation Fund Denmark through the Quantum Innovation Center, Qubiz. TG, AK, DSN, NJ and ULA acknowledge support from the Danish National Research Foundation, Center for Macroscopic Quantum States (bigQ, DNRF142). TG, NJ, SP and ULA acknowledge the EU project CiViQ

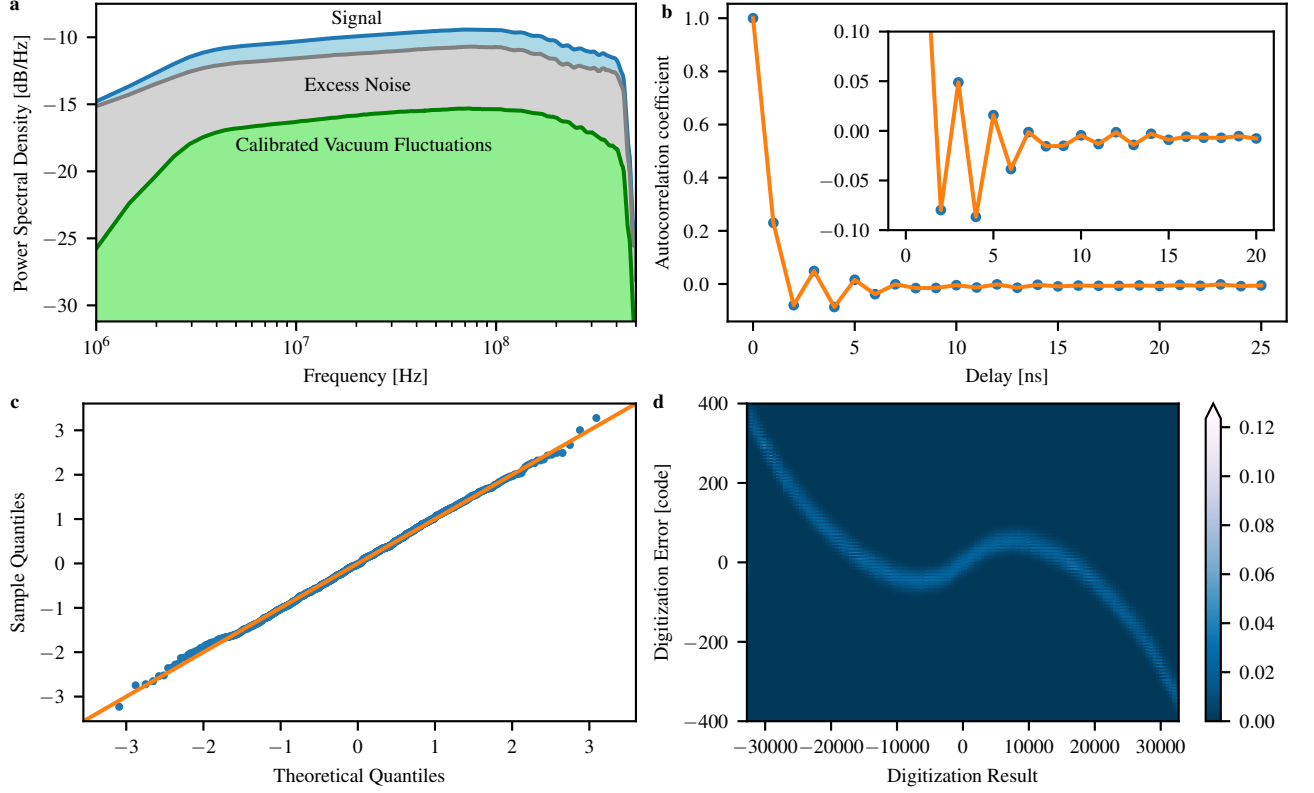


FIG. 4. Experimental results. a) The figure shows the power spectral densities of the signal, the calibrated vacuum fluctuations (obtained by the characterization) and the excess noise (obtained by subtracting the PSD of the vacuum fluctuations from the PSD of the signal). b) Autocorrelation coefficients calculated from the measured samples and averaged 1000 times. The inset shows a zoom. c) Q-Q plot indicating the Gaussianity of the measured samples. The variance of the samples has been normalized to 1. The limited ADC range truncates the tails of the Gaussian distribution which results in slight deviations from the theoretical quantiles towards the ends. d) Digitization error of the ADC with respect to the digitization results. The non-linearity and digitization noise of the ADC leads to a reduction of the min-entropy.

(grant agreement no. 820466). The authors thank Alberto Nannarelli for valuable discussions.

Appendix A: ADC Characterization

1. Min-entropy correction due to digitization error

Consider a model of ADC noise where the state

$$\rho_{\bar{X}E} = \sum_k p_{\bar{X}}(k) |k\rangle\langle k| \otimes \rho_E^{(k)}, \quad (\text{A1})$$

is replaced by its noisy version

$$\rho_{\bar{X}'\bar{X}E} = \sum_{k,k'} p_{\bar{X}'|\bar{X}}(k'|k) p_{\bar{X}}(k) |k'\rangle\langle k'| \otimes |k\rangle\langle k| \otimes \rho_E^{(k)}. \quad (\text{A2})$$

Consider for now a specific measurement that maps the quantum state on E into a random variable Z . Therefore, we can write the joint probability distribution of \bar{X}' , \bar{X} , \bar{Z} :

$$p_{\bar{X}'\bar{X}Z}(k', k, z) = p_{\bar{X}'|\bar{X}}(k'|k) p_{Z|\bar{X}}(z|k) p_{\bar{X}}(k) \quad (\text{A3})$$

$$= p_{\bar{X}'|\bar{X}}(k'|k) p_{\bar{X}|Z}(k|z) p_Z(z), \quad (\text{A4})$$

where the second equality follows from the Bayes rule.

Note that the variable \bar{X} is not observed by the user. Furthermore, we assume that it is not observed nor controlled by the adversary. This means that we are interested, for a given measurement performed by the adversary, on the conditional probability:

$$p_{\bar{X}'|Z}(k'|z) = \sum_k p_{\bar{X}'|\bar{X}}(k'|k) p_{\bar{X}|Z}(k|z).$$

The following holds:

$$p_{\bar{X}'|Z}(k'|z) = \sum_k p_{\bar{X}'|\bar{X}}(k'|k) p_{\bar{X}|Z}(k|z) \quad (\text{A5})$$

$$\leq \sum_k p_{\bar{X}'|\bar{X}}(k'|k) \max_h p_{\bar{X}|Z}(h|z). \quad (\text{A6})$$

From this, we obtain an upper bound on the guessing probability:

$$P_{\text{guess}}(\bar{X}'|Z) = \sum_z \max_{k'} p_{\bar{X}'|Z}(k'|z) p_Z(z) \quad (\text{A7})$$

$$\leq \max_{k'} \sum_k p_{\bar{X}'|\bar{X}}(k'|k) \sum_z \max_h p_{\bar{X}|Z}(h|z) p_Z(z) \quad (\text{A8})$$

$$= \max_{k'} \sum_k p_{\bar{X}'|\bar{X}}(k'|k) P_{\text{guess}}(\bar{X}|Z). \quad (\text{A9})$$

The above inequality holds for any measurement $E \rightarrow Z$. Therefore it also holds for the optimal measurement. This implies a bound for the probability of guessing with quantum side information:

$$P_{\text{guess}}(\bar{X}'|E) \leq \max_{k'} \sum_k p_{\bar{X}'|\bar{X}}(k'|k) P_{\text{guess}}(\bar{X}|E). \quad (\text{A10})$$

In terms of the min-entropy, this reads

$$H_{\min}(\bar{X}'|E) \geq H_{\min}(\bar{X}|E) - \log \left[\max_{k'} \sum_k p_{\bar{X}'|\bar{X}}(k'|k) \right]. \quad (\text{A11})$$

2. Estimation of the min-entropy correction

Note that the quantity in the argument of the logarithm in Eq. (A11) is proportional to a probability:

$$\max_{k'} \sum_k p_{\bar{X}'|\bar{X}}(k'|k) = N \max_{k'} \sum_k p_{\bar{X}'|\bar{X}}(k'|k) \frac{1}{N} = N \max_{k'} p_{\bar{X}'}(k'), \quad (\text{A12})$$

where N is the cardinality of \bar{X}' and \bar{X} , and $p_{\bar{X}'}(k') = \sum_k p_{\bar{X}'|\bar{X}}(k'|k) \omega_{\bar{X}}(k)$, where $\omega_{\bar{X}}(k) = \frac{1}{N}$ is the flat distribution for \bar{X} .

The probability $p_{\max} := \max_{k'} p_{\bar{X}'}(k')$ can be estimated experimentally. If using a sample of size S the most common value of k' is obtained S_{\max} times, then our best guess for p_{\max} is the relative frequency $\nu_{\max} = \frac{S_{\max}}{S}$. A confidence interval can then be obtained from the Hoeffding inequality:

$$\Pr \{p_{\max} \geq \nu_{\max} + \delta\} \leq e^{-2\delta^2 S}, \quad (\text{A13})$$

that is,

$$\Pr \left\{ p_{\max} \geq \nu_{\max} + \sqrt{\frac{1}{2S} \ln \frac{1}{\epsilon}} \right\} \leq \epsilon. \quad (\text{A14})$$

In conclusions, we obtain that the following min-entropy bound holds up to a probability smaller than ϵ :

$$H_{\min}(\bar{X}'|E) \geq H_{\min}(\bar{X}|E) - \log \left(N \nu_{\max} + N \sqrt{\frac{1}{2S} \ln \frac{1}{\epsilon}} \right). \quad (\text{A15})$$

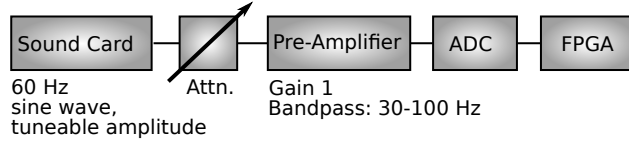


FIG. 5. Measurement setup for ADC characterization. The sound card has a high impedance output which was converted to 50 Ohm by the pre-amplifier.

For the experimental characterisation of the ADC, we have used a sample of size $S = 2^{18} \times 500000$, and the cardinality was $N = 2^{16}$. The relative frequency of the maximum value was such that $N\nu_{\max} \simeq 2.79$. Putting $\epsilon = 10^{-10}$, we then obtain the correction term

$$N\sqrt{\frac{1}{2S} \ln \frac{1}{\epsilon}} = 0.128\sqrt{\ln \frac{1}{\epsilon}} = 0.128\sqrt{10 \ln 10} \simeq 0.6. \quad (\text{A16})$$

3. Measurement

The setup for the ADC characterization is shown in 5. A sine wave of 60 Hz frequency is generated by a 24 bit sound card. The maximum amplitude of the sine is adjusted by a variable attenuator to match the input range of the ADC. A pre-amplifier (Standard Research Systems SR560) is used to bandpass filter the signal and as buffer to convert the high impedance output of the sound card to a 50 Ohm impedance as required by the ADC.

The measurement was performed as follows: We swept the amplitude of the sine wave from 0 to $2^{23} - 1$ in steps of 64. This yields 4 measurements with different voltage levels for each bin of the 16 bit ADC. For each amplitude setting 10 periods of the sine wave were recorded by the ADC with a sampling rate of 1 GS/s and the data was transferred to a computer via the FPGA. We analysed the data by determining the location of the maxima and minima and for each we calculated a histogram from 50,000 samples around the maximum and minimum, respectively. The 10 histograms were summed so that we obtained a histogram from 500,000 samples. To obtain $p_{\bar{X}|\bar{X}}(k'|k)$ we combined the 4 measurements per ADC bin and normalized the probability distribution.

Appendix B: Comparison with classical side information

In this Section we compare our lower bound on the min-entropy with quantum side information with an upper bound obtained under the assumption that the eavesdropper performs ideal homodyne detection.

In the main body of the paper we have obtained the following lower bound on the min-entropy with quantum side information:

$$H_{\min}(\bar{X}|E) \geq -\log \left[\frac{(n+\delta)(1+n+\delta)}{\delta} \max \left\{ \text{erf} \left(\sqrt{\frac{\delta}{4n(n+1+\delta)+2\delta}} \frac{\Delta x}{2g} \right), \frac{1}{2} \text{erfc} \left(\sqrt{\frac{\delta}{4n(n+1+\delta)+2\delta}} \frac{R}{g} \right) \right\} \right]. \quad (\text{B1})$$

For the sake of comparison, we simplify this expression by using an optimal choice for the ADC range R . This yields

$$H_{\min}(\bar{X}|E)_{\rho} \geq -\log \left[\frac{(n+\delta)(1+n+\delta)}{\delta} \text{erf} \left(\sqrt{\frac{\delta}{4n(n+1+\delta)+2\delta}} \frac{\Delta x}{2g} \right) \right]. \quad (\text{B2})$$

At the lowest order in Δx this in turn becomes

$$H_{\min}(\bar{X}|E)_{\rho} \gtrsim -\log \left(\frac{\Delta x}{g} \right) - \log \left(\frac{(n+\delta)(1+n+\delta)}{\sqrt{2\pi\delta 2n(n+1+\delta)+\delta}} \right). \quad (\text{B3})$$

This bound can be made tighter by using an optimal value for δ . For example, putting $\delta = n$ we obtain

$$H_{\min}(\bar{X}|E)_{\rho} \gtrsim -\log \left(\frac{\Delta x}{g} \right) - \log \left(\frac{\sqrt{2}(2n+1)}{\sqrt{\pi(4n+3)}} \right). \quad (\text{B4})$$

Let us now compute an upper bound for an eavesdropper measuring by homodyne detection. If user and eavesdropper both apply homodyne detection, then they generate a pair of correlated Gaussian variables X and Y such that

$$p_Y(y) = G(y; 0, 1 + 2n), \quad (\text{B5})$$

$$p_X(x|y) = G\left(x; \frac{2g\sqrt{n(n+1)}}{1+2n}y, \frac{g^2}{1+2n}\right). \quad (\text{B6})$$

The variable X is then mapped into a discrete and bounded variable \bar{X} as described in the main body of the paper. Using an optimal choice of the range R of the ADC, the min-entropy of \bar{X} conditioned on the Y is

$$H_{\min}(\bar{X}|C) = -\log \operatorname{erf}\left(\frac{\Delta x}{g} \frac{\sqrt{2+4n}}{4}\right), \quad (\text{B7})$$

or, up to correction of order higher than Δx ,

$$H_{\min}(\hat{X}|C) \simeq -\log\left(\frac{\Delta x}{g}\right) - \log\left(\sqrt{\frac{1+2n}{2\pi}}\right). \quad (\text{B8})$$

Figure 6 shows, for $\Delta x = n/1000$, the homodyne upper bound in Eq. (B7) and the min-entropy lower bound in Eq. (B4), as function of the mean photon number n . The latter is plotted for $\delta = n$. Figure 7 shows instead the difference between the lower bound in Eq. (B4) and the homodyne upper bound in Eq. (B8), again for $\delta = n$ (notice that the difference is independent of Δx). This plots show that our lower bound is tight and in fact homodyne is close to be the optimal measurement for Eve. Note that the difference between the upper and lower bounds is as small as a fraction of a bit.

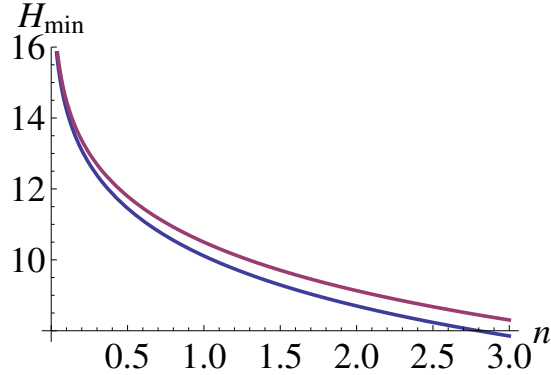


FIG. 6. Homodyne upper bound $H_{\min}(\bar{X}|C)$ as in Eq. (B7) [red line] and the min-entropy lower bound in Eq. (B4), vs the mean photon number n , for $\delta = n$ and $\Delta x = n/1000$.

Appendix C: Estimation of variances and the entropy rate

In this appendix we discuss the estimation of the variance, entropy rate and conditional variance of the noise and signal. To make things more concrete, we focus on the estimation of the signal variance σ^2 , entropy rate $h(X)$ and the conditional signal variance σ_X^2 . Assume that T is the runtime of the experiment, and n signal measurements are performed at regular time intervals of $\delta t = T/n$. The spectral density computed from these data is a function of n discrete frequencies, denoted as ω_j 's, taking values between $2\pi/T$ and $2\pi n/T$. Below we work with the discrete variable λ_j defined as $\lambda_j \equiv T\omega_j/n$, which can be approximated by the continuous variable λ taking values with domain $[0, 2\pi]$.

We estimate the spectral density $f(\lambda)$ by applying the Welch's method, according to which the data are first divided in M (possibly overlapping) blocks, and then in each block the periodogram is computed, i.e., the discrete Fourier transform of the data contained in that very block. The spectral density is then estimated by taking the average over the periodograms. We assume that the periodograms, as random variables, are independent and identically

distributed, and that each periodogram is distributed as the square of a Gaussian variable. Then the Welch's estimate of the spectral density is distributed as a (rescaled) $\chi^2(k)$ variable with M degrees of freedom. Denoting as $f_0(\lambda_j)$ the Welch's estimate for the spectral density and as $f(\lambda_j)$ its real value, then we can obtain a confidence interval by applying a tail bound of a $\chi^2(k)$ variable. For example we can exploit the tail bounds (see e.g. [36])

$$\Pr \left\{ f(\lambda_j) < \frac{f_0(\lambda_j)}{1+t} \right\} \leq e^{-Mt^2/8}, \quad (\text{C1})$$

$$\Pr \left\{ f(\lambda_j) > \frac{f_0(\lambda_j)}{1-t} \right\} \leq e^{-Mt^2/8}. \quad (\text{C2})$$

For $t \ll 1$ this yields, up to higher order terms,

$$\Pr \{f(\lambda_j) \notin [(1-t)f_0(\lambda_j), (1+t)f_0(\lambda_j)]\} = P(t) \quad (\text{C3})$$

with

$$P(t) \leq 2e^{-Mt^2/8}. \quad (\text{C4})$$

Let us first discuss the estimation of the entropy rate

$$h(X) = \frac{1}{2} \int_0^{2\pi} \frac{d\lambda}{2\pi} \log [2\pi e f(\lambda)], \quad (\text{C5})$$

as approximated by the finite sum

$$h(X) \simeq \frac{1}{2} \sum_{j=1}^n \frac{1}{n} \log [2\pi e f(\lambda_j)]. \quad (\text{C6})$$

For each given j , $1 - P(t)$ is the probability that $f(\lambda_j) \in [(1-t)f_0(\lambda_j), (1+t)f_0(\lambda_j)]$, then it follows (from an application of the union bound) that

$$\Pr \{\exists j \mid f(\lambda_j) \notin [(1-t)f_0(\lambda_j), (1+t)f_0(\lambda_j)]\} \leq nP(t). \quad (\text{C7})$$

This is equivalent to say that, with probability larger than $1 - nP(t)$, $f(\lambda_j)$ lays between $(1-t)f_0(\lambda_j)$ and $(1+t)f_0(\lambda_j)$ for all $j = 1, \dots, n$. Therefore

$$h(X) \in \left[\frac{1}{2} \sum_{j=1}^n \frac{1}{n} \log [2\pi e f_0(\lambda_j)] + \frac{1}{2} \log (1-t), \frac{1}{2} \sum_{j=1}^n \frac{1}{n} \log [2\pi e f_0(\lambda_j)] + \frac{1}{2} \log (1+t) \right] \quad (\text{C8})$$

with probability at least equal to $1 - nP(t) = 1 - 2ne^{-Mt^2/8}$. A further linear approximation for $t \ll 1$ yields the confidence interval

$$h(X) \in \left[\frac{1}{2} \sum_{j=1}^n \frac{1}{n} \log [2\pi e f_0(\lambda_j)] - \frac{\log e}{2} t, \frac{1}{2} \sum_{j=1}^n \frac{1}{n} \log [2\pi e f_0(\lambda_j)] + \frac{\log e}{2} t \right]. \quad (\text{C9})$$

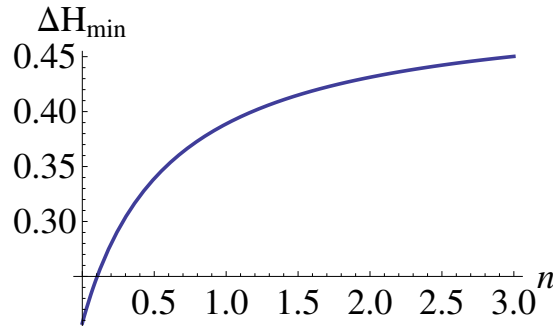


FIG. 7. The difference between the Homodyne upper bound in Eq. (B8) and the min-entropy lower bound in Eq. (B4), vs the mean photon number n , for $\delta = n$.

Finally, to take into account the overlap between adjacent periodograms, we replace $M \rightarrow \gamma M$, for $\gamma < 1$. For example, if the periodogram have a 50% overlap we put $\gamma = 1/2$. In conclusion, with an overlap of 50%, we obtain that for any given $\epsilon > 0$, the entropy rate lies within the interval

$$h(X) \simeq \frac{1}{2} \sum_{j=1}^n \frac{1}{n} \log [2\pi e f_0(\lambda_j)] \pm 2 \log e \sqrt{\frac{1}{M} \ln \left(\frac{2n}{\epsilon} \right)}, \quad (\text{C10})$$

up to a probability not larger than ϵ .

From the entropy rate we obtain a confidence interval for the conditional variance, $\sigma_X \in [\sigma_X^-, \sigma_X^+]$, where

$$\sigma_X^\pm = \frac{1}{2\pi e} 2^{\sum_{j=1}^n \frac{1}{n} \log [2\pi e f_0(\lambda_j)]} 2^{\pm 4 \log e \sqrt{\frac{1}{M} \ln \frac{2n}{\epsilon}}}. \quad (\text{C11})$$

Similarly, we obtain an estimate of the signal variance σ^2 by exploiting the relation

$$\sigma^2 = \int_0^{2\pi} \frac{d\lambda}{2\pi} f(\lambda), \quad (\text{C12})$$

from which we derive a confidence interval

$$\sigma^2 \simeq \left(1 \pm 4 \sqrt{\frac{1}{M} \ln \frac{2n}{\epsilon}} \right) \sum_{j=1}^n \frac{1}{n} f_0(\lambda_j). \quad (\text{C13})$$

Along the same lines we obtain a confidence interval for the conditional noise variance, $\sigma_U \in [\sigma_U^-, \sigma_U^+]$ (this must additionally includes systematic errors). To obtain a worst-case estimate of the min-entropy we consider the smaller value for the signal variance, σ_X^- , and the larger one for the noise, σ_U^+ .

Appendix D: Characterization of vacuum fluctuations power spectral density

Here, we open up the homodyne detector black box and show by including imperfections that the bound given in the main text is indeed a lower bound on the vacuum fluctuations. As described in the main text we beat two lasers, the local oscillator with power P_{LO} and an auxiliary signal laser with power P_{sig} which is frequency detuned with respect to the local oscillator by ν . The beams interfere at a beam splitter with splitting ratio $R(\nu) : 1 - R(\nu)$, where the frequency dependence ν accounts for a frequency dependent common mode rejection of the homodyne electronics. We furthermore take into consideration the visibility of the interference $\chi \in (0, 1]$ and the quantum efficiencies η_1 and $\eta_2 \in (0, 1]$ of the two photo diodes.

After photo detection and current subtraction the beat signal current at time t reads

$$i_{\text{beat}}(t) = 2\chi^2(\eta_1 + \eta_2) \sqrt{R(\nu)(1 - R(\nu))} \frac{e}{\hbar\omega} \sqrt{P_{\text{LO}} P_{\text{sig}}} \cos(2\pi\nu t). \quad (\text{D1})$$

Here ω is the absolute angular frequency of the local oscillator laser. The square of the root mean square (RMS) amplitude of the beat signal digitized by an analog-to-digital (ADC) converter as obtained by a power spectrum of acquired samples is then given by

$$\widetilde{\text{TF}}(\nu) := \left(\sqrt{2}\chi^2(\eta_1 + \eta_2) \sqrt{R(\nu)(1 - R(\nu))} \frac{e}{\hbar\omega} \right)^2 P_{\text{LO}} P_{\text{sig}} G(\nu), \quad (\text{D2})$$

where $G(\nu)$ describes the overall gain of homodyne detector, possible filters and ADC analog input as well as includes the digitization into integers. We call $\text{TF} := \widetilde{\text{TF}}/P_{\text{sig}}$ the transfer function.

The power spectral density (PSD) of the vacuum fluctuations after photo detection and digitization reads

$$\text{PSD}_{\text{vac}} = 2e(i_{\text{dc1}} + i_{\text{dc2}})G(\nu) = 2 \frac{e^2}{\hbar\omega} (\eta_1(1 - R(\nu)) + \eta_2 R(\nu)) P_{\text{LO}} G(\nu), \quad (\text{D3})$$

where i_{dc1} and i_{dc2} are the direct photo currents generated by the photo diodes. Using the characterization of the transfer function from Eq. (D2) yields

$$\text{PSD}_{\text{vac}} = \hbar\omega \frac{1}{\chi^2} \frac{\eta_1(1 - R(\nu)) + \eta_2 R(\nu)}{(\eta_1 + \eta_2)^2 R(\nu)(1 - R(\nu))} \frac{\widetilde{\text{TF}}(\nu)}{P_{\text{sig}}} \geq \hbar\omega \frac{\widetilde{\text{TF}}(\nu)}{P_{\text{sig}}}. \quad (\text{D4})$$

In the last step we lower bounded the PSD of the vacuum fluctuations by using $1/\chi \geq 1$ and

$$(\eta_1(1 - R(\nu)) + \eta_2 R(\nu)) / ((\eta_1 + \eta_2)^2 R(\nu)(1 - R(\nu))) \geq 1, \quad (\text{D5})$$

where equality holds for $\eta_1 = \eta_2 = 1$, $R = 0.5$.

Appendix E: Continuous-variable quantum key distribution application example

Continuous-variable quantum key distribution uses a Gaussian modulation of the coherent state excitation, i.e. of both the amplitude and phase quadrature components. For a single execution of the protocol a large amount of coherent quantum states, say 10^{10} are generated and transmitted. For each quantum state 2 random numbers are required, the amplitude and the phase quadrature values. In practise these are discretized and for our example we choose an 8 bit resolution. Thus, 16 bit are required to generate 1 coherent state. Since our QRNG delivers 2048 random bits per single execution of the randomness extraction, $N = \frac{16\text{bit} \cdot 10^{10}}{2048\text{bit}} \approx 7.8 \times 10^7$ runs are needed.

We now assume that the QRNG ran N_1 times prior to the generation of the random numbers for QKD. Then the random number string for QKD has an epsilon stemming from the hashing of

$$\epsilon_{\text{hash}}^{\text{QKD}} = (N_1 + 1)\epsilon_{\text{hash}} + (N_1 + 2)\epsilon_{\text{hash}} + \dots + (N_1 + N)\epsilon_{\text{hash}} = \left(N \cdot N_1 + \frac{(N \cdot (N + 1))}{2} \right) \epsilon_{\text{hash}} .$$

In our example we require that the random numbers used in the quantum key distribution system have an epsilon security parameter of less than 10^{-9} . Assuming that the parameter estimation epsilon is constant over time and noticing that ϵ_{PE} is one order of magnitude smaller, the series of random numbers must have an epsilon of $\epsilon_{\text{hash}}^{\text{QKD}} \leq 10^{-9}$ from hashing.

After continuously running the QRNG for 10 years,

$$N_1 = \frac{10 \text{ years}}{\frac{256}{1\text{GS/s}}} \approx 1.2 \times 10^{15} ,$$

where 256 is the number of acquired samples from the ADC per extraction run.

Solving the above equation for ϵ_{hash} and plugging in numbers for N_1 and N yields $\epsilon_{\text{hash}} \lesssim 10^{-32}$. Thus we chose $\epsilon_{\text{hash}} = 10^{-33}$ to be on the safe side.

-
- [1] D. Frauchiger, R. Renner, and M. Troyer, arXiv , 1311.4547 (2013).
 - [2] A. Acin and L. Masanes, Nature **540**, 213 (2016).
 - [3] X. Ma, Z. Cao, and X. Yuan, Quantum Information **2**, 16021 (2016).
 - [4] M. Herrero-Collantes and J. C. Garcia-Escartin, Review of Modern Physics **89**, 015004 (2017).
 - [5] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. Pereira, M. Razavi, J. S. Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden, arXiv , 1906.01645 (2019).
 - [6] C. Gabriel, C. Wittmann, D. Sych, R. Dong, W. Mauerer, U. L. Andersen, C. Marquardt, and G. Leuchs, Nature Photonics **4**, 711 (2010).
 - [7] T. Symul, S. M. Assad, and P. K. Lam, Applied Physics Letters **98**, 231103 (2011).
 - [8] M. Avesani, D. G. Marangon, G. Vallone, and P. Villoresi, Nature Communications **9**, 5365 (2018).
 - [9] D. G. Marangon, G. Vallone, and P. Villoresi, Physical Review Letters **118**, 060503 (2017).
 - [10] M. Fuerst, H. Weier, S. Nauwerth, D. G. Marangon, C. Kurtsiefer, and H. Weinfurter, Optics Express **18**, 13029 (2010).
 - [11] F. Xu, B. Qi, X. Ma, H. Xu, H. Zheng, and H.-K. Lo, Optics Express **20**, 12366 (2012).
 - [12] J. Y. Haw, S. M. Assad, A. M. Lance, N. H. Y. Ng, V. Sharma, P. K. Lam, and T. Symul, Physical Review Applied **3**, 054004 (2015).
 - [13] Y.-Q. Nie, L. Huang, Y. Liu, F. Payne, J. Zhang, and J.-W. Pan, Review of Scientific Instruments **86**, 063105 (2015).
 - [14] Y. Shi, B. Chng, and C. Kurtsiefer, Applied Physics Letters **109**, 041101 (2016).
 - [15] C. Abellán, W. Amaya, M. Jofre, M. Curty, A. Acín, J. Capmany, V. Pruneri, and M. W. Mitchell, Optics Express **22**, 1645 (2014).
 - [16] X.-G. Zhang, Y.-Q. Nie, H. Zhou, H. Liang, X. Ma, J. Zhang, and J.-W. Pan, Review of Scientific Instruments **87**, 076102 (2016).
 - [17] L. Huang and H. Zhou, Journal of the Optical Society of America B **36**, 130 (2019).
 - [18] Z. Zheng, Y. Zhang, W. Huang, S. Yu, and H. Guo, Review of Scientific Instruments **90**, 043105 (2019).
 - [19] M. W. Mitchell, C. Abellan, and W. Amaya, Physical Review A **91**, 012314 (2015).
 - [20] X. Zhang, Y. Q. Nie, H. Liang, and J. Zhang, IEEE-NPSS Real Time Conference, RT 2016 , 1 (2016).
 - [21] J. H. Shapiro, IEEE Journal of Quantum Electronics **QE-21**, 237 (1985).
 - [22] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Reviews of Modern Physics **84**, 621 (2012).
 - [23] R. Renner, *Security of Quantum Key Distribution*, Ph.D. thesis, ETH Zürich (2005).

- [24] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, IEEE Transactions on Information Theory **57**, 5524 (2011).
- [25] M. Tomamichel, *A framework for non-asymptotic quantum information theory*, Ph.D. thesis, ETH Zurich (2012).
- [26] Here \log stands for the logarithm in base 2 and \ln for the natural logarithm.
- [27] L. Bianchi, S. L. Braunstein, and S. Pirandola, Physical Review Letters **115**, 260501 (2015).
- [28] K. P. Seshadreesan, L. Lami, and M. M. Wilde, Journal of Mathematical Physics **59**, 072204 (2018).
- [29] Note that Theorem 24 of Ref. [28] gives a formula to compute the norm of the operator $\rho^{1/2}\gamma^{-1}\rho^{1/2}$. However, it is easy to show that this is the same as the norm of $\gamma^{-1/2}\rho\gamma^{-1/2}$. In fact, consider the operator $\rho^{1/2}\gamma^{-1/2}$, and its singular value decomposition $\rho^{1/2}\gamma^{-1/2} = U\Delta V$, where U and V are unitary and Δ is diagonal. We then have $\rho^{1/2}\gamma^{-1}\rho^{1/2} = U\Delta^2U^\dagger$. Analogously, we also have $\gamma^{-1/2}\rho\gamma^{-1/2} = V^\dagger\Delta^2V$. As the operators $\rho^{1/2}\gamma^{-1}\rho^{1/2}$ and $\gamma^{-1/2}\rho\gamma^{-1/2}$ are unitary equivalent, they also have the same norm.
- [30] R. M. Gray, Foundations and Trends in Communications and Information Theory **2**, 155 (2006).
- [31] M. N. Wegman and J. L. Carter, Journal of Computer and System Sciences **22**, 265 (1981).
- [32] R. G. Brown, “<http://www.phy.duke.edu/~rgb/General/dieharder.php>,” (2018).
- [33] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Heckert, J. Dray, and S. Vo, NIST Special Publication **800-22** (2001).
- [34] A. R. Dixon, Z. L. Yuan, J. F. Dynes, A. W. Sharpe, and A. J. Shields, Optics Express **16**, 18790 (2008).
- [35] D. Huang, P. Huang, D. Lin, C. Wang, and G. Zeng, Optics Letters **40**, 3695 (2015).
- [36] C. Lupo, C. Ottaviani, P. Papanastasiou, and S. Pirandola, Physical Review A **97**, 052327 (2018).