Experimental fast quantum random number generation using high-dimensional entanglement with semi-self-testing

Feihu Xu, Jeffrey H. Shapiro, and Franco N. C. Wong

Research Laboratory of Electronics, Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, USA

Abstract: A quantum random number generator (QRNG) generates genuine randomness from the intrinsic probabilistic nature of quantum mechanics. The central problems for most QRNGs are estimating the entropy of the genuine randomness and producing such randomness at high rates. Here we propose and demonstrate a fast semi-self-testing QRNG at a rate of 24 Mbit/s based on a high-dimensional entanglement system, in which the user monitors the entropy in real-time via the observation of a nonlocal quantum interference. Our work provides a practical approach to a robust QRNG with trusted but error-prone devices.

Introduction. Randomness is indispensable for information processing. Quantum random number generators (QRNGs) can generate true randomness by exploiting the fundamental indeterminism of quantum mechanics [1]. Most current QRNGs are photonic systems built with trusted and calibrated devices [2] and provide Gbit/s generation speeds at relatively low cost. However, a central issue in these QRNGs is how to certify and quantify the entropy of the genuine randomness, i.e., the randomness that originates from the intrinsic unpredictability of measurements in quantum mechanics. Entropy estimates for specific setups were recently proposed using sophisticated theoretical models [3]. Nevertheless, these techniques require parameter characterization which may be difficult to accurately assess in practice. An elegant solution to estimating the entropy is the device-independent (DI) or self-testing QRNG [1, 4, 5], but its practical implementation is challenging because it requires loophole-free violation of Bell's inequality, resulting in low generation rates of $\sim$1 bit/s [1, 5]. Recently, Lunghi *et al.* proposed a more practical solution without the need for Bell violation [6], which can be termed *semi-self-testing*, in which the randomness can be guaranteed based on a few general assumptions that do not require detailed device characterization. This scheme is highly desirable as it focuses on real-world implementations with trusted but error-prone devices, although its implementation to date still suffers from low generation rates of tens of bits/s [6].

In this work, based on high-dimensional entanglement, we propose and experimentally demonstrate a simple, practical and fast semi-self-testing QRNG at a rate over 24 Mbits/s. The amount of genuine quantum randomness is quantified directly from observation of a nonlocal interference and a pair of incompatible quantum measurements, and it is separated from other sources of randomness such as technical noise with a randomness extractor. We achieved the high generation rate by virtue of a number of experimental features: a high-dimensional time-energy entangled-photon source capable of producing multiple random bits per photon, a high-visibility Franson interferometer for evaluating entanglement, and high-efficiency superconducting-nanowire single photon detectors (SNSPDs).

Protocol. An entanglement source is used to generate high-dimensional entangled photon pairs. In the ideal case, the N_d-dimensional entangled state can be written as

$$|\psi\rangle = \frac{1}{\sqrt{N_d}} \sum_{i=0}^{N_d-1} |i\rangle_{\mathrm{A}} \bigotimes |i\rangle_{\mathrm{B}}. \tag{1}$$

This state is observed by two measurement systems **A** and **B**. The randomness is generated from **A**, while **B** is used to test the states. To model imperfections, we assume that the state ρ_A of system **A** is not pure and is correlated with environmental noise.

Using the same concept as the semi-self-testing QRNG [6], we assume the devices in the protocol are *not* deliberately designed to fool the user, but the implementation may be imperfect. The central task is to estimate the amount of genuine randomness based only on measurements. This is a nontrivial task as the observed randomness can have different origins. If the state ρ_A is a superposition of high-dimensional states, then the outcome R_i cannot be predicted with certainty, even if the internal state is known, thus resulting in genuine quantum randomness. On the other hand, the randomness may be due to technical imperfections such as

detector noise and temperature fluctuations, whose randomness clearly contains *no* quantum randomness, since the outcome R_i can be perfectly guessed if the imperfections were well quantified.

In our approach, the amount of genuine randomness is quantified via the uncertainty relation [7, 8] from a pair of incompatible quantum measurements, namely time measurement $\mathbb{T}$ and frequency measurement $\mathbb{W}$. $\mathbb{T}$ is realized by directly measuring the photon's arrival time, while $\mathbb{W}$ is performed by Franson interferometry [9]. In particular, the Franson visibility V is used to bound the randomness. Classical fields result in V that is no greater than 50%. For a maximally entangled state, V would be unity in the ideal case [9, 10]. Conceptually, $V > 50\%$ guarantees that the source's output is entangled and thus contains genuine randomness.

More rigorously, Ref. [11] has proved that V provides an explicit bound for the correlations in $\mathbb{W}$, under the assumption that the quantum state is Gaussian, which in turns bounds the conditional maximum entropy (given system **B**) in $\mathbb{W}$ via the theory developed in [8]. By using the entropic uncertainty relation for smooth entropies [7], we can determine the conditional min-entropy given the environmental noise and thus the guessing probability, i.e., the amount of genuine randomness. Our protocol provides self-monitoring because measurements of V directly quantify the amount of genuine randomness in the observed data. A threshold value V_0 is pre-selected and the randomness can be generated *only* when the observation satisfies $V > V_0$. Two particular advantages of this approach are: (i) the observation of V does not rely on detailed models of the devices that are employed; and (ii) no loophole-free Bell inequality violation is required.

Experiment. Figure 1 shows our experimental setup. It uses a high-dimensional time-energy entanglement source based on spontaneous parametric down-conversion (SPDC) in a PPKTP waveguide [12], which supports multiple spatial modes at telecom wavelengths. The 46.1 μm grating period was designed for type-II quasi-phase-matched wavelength-degenerate outputs at 1560 nm in the fundamental modes of the signal and idler fields. The phase-matching bandwidth is 1.6 nm with a corresponding biphoton correlation time of 2 ps. The pump was a 780-nm continuous-wave diode laser. We extracted the fundamental signal and idler modes using a dichroic mirror to remove the pump and a 10-nm band-pass filter to spectrally remove the higher-order SPDC spatial modes. We achieved ~81% waveguide-to-fiber coupling efficiency. A polarization beam splitter was used to separate the signal and idler photons and send them to systems **A** and **B**, respectively. Losses in the waveguide and from the waveguide to the fiber are ~15% and ~12%, respectively. Overall, we measured a heralding efficiency of ~50%.

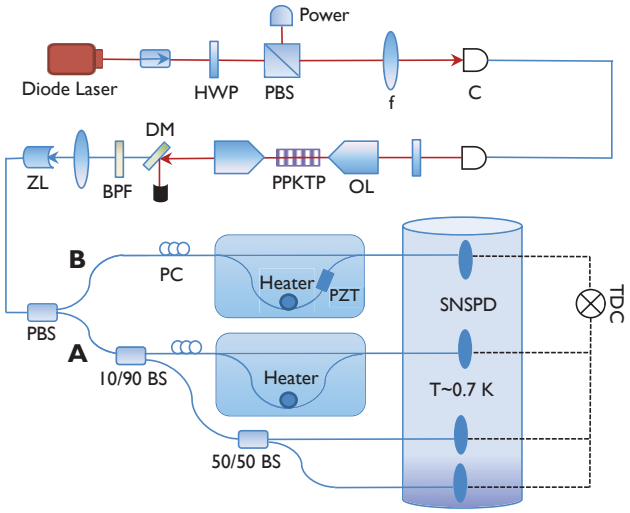

Fig. 1. Experimental setup. HWP: half wave plate; PBS: polarization beam splitter; C: coupler; OL: objective lens; DM: dichroic mirror; BPF: band-pass filter; ZL: zoom lens; PC: polarization controller; BS: beam splitter; PZT: piezoelectric transducer; SNSPD: superconducting-nanowire single photon detectors; TDC: time-digital converter.

We used a Franson interferometer with local dispersion cancellation to measure the source's entanglement quality. We set up two identical Mach-Zehnder interferometers (MZIs), in which the long arm was made of standard single-mode fiber and low-dispersion LEAF fiber, such that the differential group delay (due to dispersion) between the long and short arms was zero [10]. To achieve long-term stability, the MZIs were enclosed in a multilayered thermally insulated box, whose temperature was actively stabilized. The path mismatch of each MZI was measured to be $\Delta T = 3.2$ ns. We coiled the long-path fiber of each MZI on a closed-loop temperature-controlled heater to precisely match the ΔT of the two MZIs. The variable phase shift was set by a piezoelectric transducer fiber stretcher. By carefully fine-tuning the input polarizations and the temperatures, we measured an interference visibility V of $98.8 \pm 0.3\%$.

We implemented the random basis choice passively with a 10/90 beam splitter. The photon arrival times were measured by WSi SNSPDs with detection efficiencies of ~85%, dark-count rates of ~400/s, timing

jitters of ∼250 ps, and maximum count rates of ∼2 MHz without detector saturation. To mitigate the long reset times of the SNSPDs and to achieve a higher generation rate, system **A** used a passive 50/50 beam splitter to distribute incident photons equally between two WSi SNSPDs and their data were interleaved. Hence, a total of four WSi SNSPDs were used and their outputs were recorded by time-to-digital converters.

Results. In our proof-of-principle experiment, we recorded the data for a maximum duration of 60 s. We monitored the Franson visibility before and after the data recording to ensure that the experimental V exceeded V_0, where V_0 is a preset threshold value that is selected to be less than the experimentally observed V values. The amount of genuine randomness increases monotonically with increasing V_0, as shown in Fig. 2, and it is therefore desirable to use high-quality devices. Here, we selected $V_0 = 98.5\%$ as a conservative threshold, since it was the lower bound in most runs in our experiment.

We quantified the entropy of genuine randomness using a method that applies specifically to high Franson visibilities. Simulation shows that for $V_0 = 98.5\%$ the optimal number of bits per photon coincidence is 6.0 bits/coincidence at $N_d = 2048$. Figure 2 shows the experimental results for QRNG throughput for different total running times. A longer running time produces more data, thus minimizing the finite-data effect and yielding more genuine randomness from our raw data. The results show that a continuous running time of ∼60 s can already produce randomness that is close to the asymptotic case of an infinitely long operation.

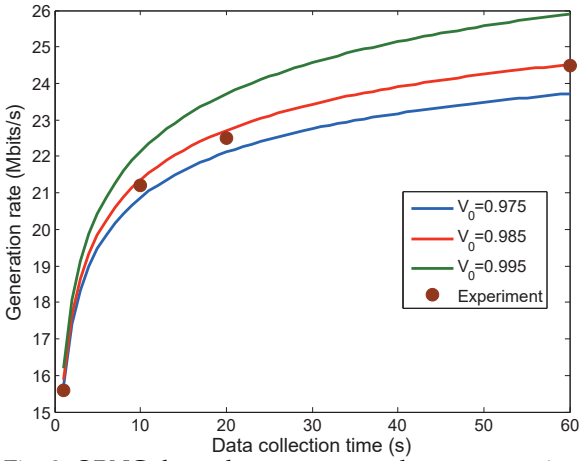

Fig. 2. QRNG throughput versus total system running time. Solid red circles are experimental rates with $V_0 = 98.5\%$. The three curves are simulated results with different V_0 values. A higher V_0 yields a higher throughput. For running times below 10 s, the finite-data effect reduces the QRNG throughput substantially.

We implemented a Toeplitz-hashing extractor [3] to extract genuine random numbers. The output random bits successfully pass all the tests of DIEHARD test suit. The count rates for the two SNSPDs were about 1.8 and 2.3 Msamples/s. Hence, the final QRNG rate is about $6 \times 4.1 = 24.6$ Mbit/s, which is several orders of magnitude higher than previous self-testing [1, 5] and semi-self-testing [6] experiments. This significantly faster rate benefits from the following factors: a high-dimensional entanglement system that can generate *multiple* bits per photon, highly efficient WSi SNSPDs, a high-quality PPKTP waveguide SPDC source with high fiber-coupling efficiency, and high-visibility Franson interferometry.

Conclusion. We have demonstrated a high-speed semi-self-testing QRNG based on high-dimensional entanglement with a rate over 24 Mbit/s. Compared to the standard device-dependent approaches with fully calibrated devices, our QRNG delivers a stronger form of security requiring less characterization of the physical implementation, while still maintaining a high generation rate comparable to those of commercial QRNGs. Though our approach offers a weaker form of security than fully self-testing QRNG, it focuses on a scenario with trusted but error-prone devices, which we believe is more relevant in practice.

References

1. S. Pironio *et al.*, *Nature* **464,** 1021 (2010).
2. B. Sanguinetti *et al.*, *Phys. Rev. X* **4,** 031056 (2014); C. Abellán *et al.*, *Phys. Rev. Lett.* **115,** 250403 (2015).
3. X. Ma *et al.*, *Phys. Rev. A* **87,** 062327 (2013); D. Frauchiger, R. Renner, and M. Troyer, *arXiv:1311.4547* (2013).
4. R. Colbeck. PhD thesis, University of Cambridge, 2006; C. Miller and Y. Shi, *arXiv: 1402.0489* (2014).
5. B. G. Christensen *et al.*, *Phys. Rev. Lett.* **111,** 130406 (2013).
6. T. Lunghi *et al.*, *Phys. Rev. Lett.* **114,** 150501 (2015).
7. M. Tomamichel and R. Renner, *Phys. Rev. Lett.* **106,** 110506 (2011); G. Vallone *et al.*, *Phys. Rev. A* **90,** 052327 (2014).
8. F. Furrer *et al.*, *Phys. Rev. Lett.* **109,** 100502 (2012).
9. J. D. Franson, *Phys. Rev. Lett.* **62,** 2205 (1989); J. D. Franson, *Phys. Rev. Lett.* **67,** 290 (1991).
10. T. Zhong and F. N. C. Wong, *Phys. Rev. A* **88,** 020103(R) (2013).
11. Z. Zhang, J. Mower, D. Englund, F. N. C. Wong, and J. H. Shapiro, *Phys. Rev. Lett.* **112,** 120506 (2014).
12. T. Zhong, F. N. C. Wong, T. D. Roberts, and P. Battle, *Opt. Express* **14,** 12019 (2009).