Truly random number generation based on measurement of phase noise of a laser

Hong Guo,[*] Wenzhuo Tang, Yu Liu, and Wei Wei
CREAM Group, State Key Laboratory of Advanced Optical Communication Systems and Networks (Peking University) and Institute of Quantum Electronics, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China

(Received 21 August 2009; revised manuscript received 8 March 2010; published 27 May 2010)

We present a simple approach to realize truly random number generator based on measuring the phase noise of a single-mode vertical cavity surface emitting laser. The true randomness of the quantum phase noise originates from the spontaneous emission of photons and the random bit generation rate is ultimately limited only by the laser linewidth. With the final bit generation rate of 20 Mbit/s, the truly random bit sequence guaranteed by the uncertainty principle of quantum mechanics passes the three standard randomness tests (ENT, Diehard, and NIST Statistical Test Suites). Moreover, a *continuously* generated random bit sequence, with length up to 14 Gbit, is verified by two additional criteria for its true randomness.

DOI: 10.1103/PhysRevE.81.051137 PACS number(s): 05.40.−a, 42.55.Ah, 42.55.Px

Random number generator (RNG) has wide applications in statistical sampling [1], computer simulations [2], randomized algorithm [3], and cryptography [4]. Traditionally, pseudorandom number generator (PRNG) based on computational algorithms is adopted to generate random bits and is competent in many fields. However, it cannot produce intrinsically unpredictable and irreproducible bit sequence and so may result in potential dangers in security-related applications, say, in quantum cryptography [5]. Actually, the unconditional security of quantum key distribution can only be guaranteed when a truly random number generator (TRNG), based on quantum-mechanical process rather than the intractability assumption of classical algorithms [6], is available.

Distinct from PRNG, the TRNG can only be realized by quantum-mechanical process instead of by algorithm or deterministic physical process (e.g., chaotic lasers). The quantum-mechanical processes, such as radioactive decay [7] and those based on laser (photon) emission or detection [8–11], can ensure the inability of pre-estimation on random numbers and so can be adopted as candidates to implement TRNG. In particular, those based on the detection of laser field has attracted tremendous interests in recent decade. Recently, the chaotic lasers were utilized for GHz random bit generation [12–14]. However, although the quantum noise is also amplified therein, the observed signal is mainly due to the chaotic behavior of the laser(s) rather than the quantum noise. Therefore, the chaotic laser-based RNG is not inherently random owing to its deterministic nature [15,16]. On the other hand, the abovementioned TRNGs [7–10] cannot offer the high generation rate as the PRNG based on chaotic laser(s) [12–14]. The typically maximal generation rate of recent TRNGs is around 4 Mbit/s for photon detection scheme [9]. Moreover, the statistical bias and correlation for long random bit sequence were not investigated in those schemes.

In this paper, we propose a simple TRNG scheme based on measuring the quantum phase noise, which is a Gaussian random variable [17,18], of a single-mode vertical cavity surface emitting laser (VCSEL). The true randomness of the quantum phase noise is originated from the random nature of spontaneous emission. In the following, it shows that the generation rate of this TRNG is ultimately limited only by the laser linewidth. In our experiment, the final generation rate reached 20 Mbit/s; and further, the true randomness is not only guaranteed by quantum-mechanical uncertainty principle and three standard randomness tests, but is also verified by two additional criteria (statistical bias and correlation coefficient) for the long (up to 14 Gbit) random bit sequence.

The schematic setup is shown in Fig. 1 and the delayed self-homodyne method is used to measure the phase noise of the VCSEL. In this case, the output alternative current (AC) voltage of the avalanche photodetector (APD) detecting the beat signal is $V_{ph} \propto \mathrm{AC}[I_{beat}] = 2\mathcal{E}(t)\mathcal{E}(t+\tau)\cos[\phi(t)-\phi(t+\tau)]$, where the amplitude fluctuations of $\mathcal{E}(t)$ and $\mathcal{E}(t+\tau)$ are negligible compared to the phase fluctuation corresponding to $\cos[\phi(t)-\phi(t+\tau)]$ [17,18]. When the delay time is much longer than the coherence time of laser (i.e., $\tau \gg \tau_{coh}$), the phase difference $\Delta\phi(t)=\phi(t)-\phi(t+\tau)$ is a Gaussian random variable [17] and then

$$\langle E^*(t)E(t+\tau)\rangle \propto \exp(-|\tau|/\tau_{coh}) \to 0, \tag{1}$$

where $\tau_{coh}=(\pi\Delta\nu_{laser})^{-1}$ [18] and $\Delta\nu_{laser}$ is the laser linewidth. This indicates that the electric field amplitudes of the laser at different time are mutually independent if the time interval (i.e., delay time) is much longer than the coherence time of the laser. Further, similar calculation procedure can

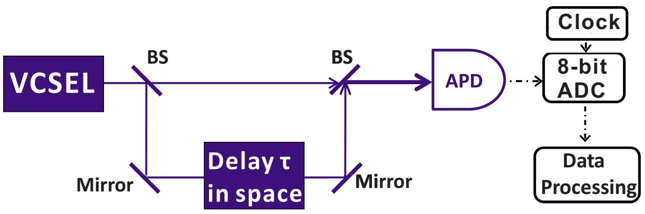

FIG. 1. (Color online) Schematic setup of TRNG based on the phase noise measurement using delayed self-homodyne method. BS, beam splitter; APD, avalanche photodetector with the low (high) cutoff frequency of 50 kHz (1 GHz). ADC, 8-bit binary analog-digital converter working at 40 MHz.

[*] Corresponding author; hongguo@pku.edu.cn

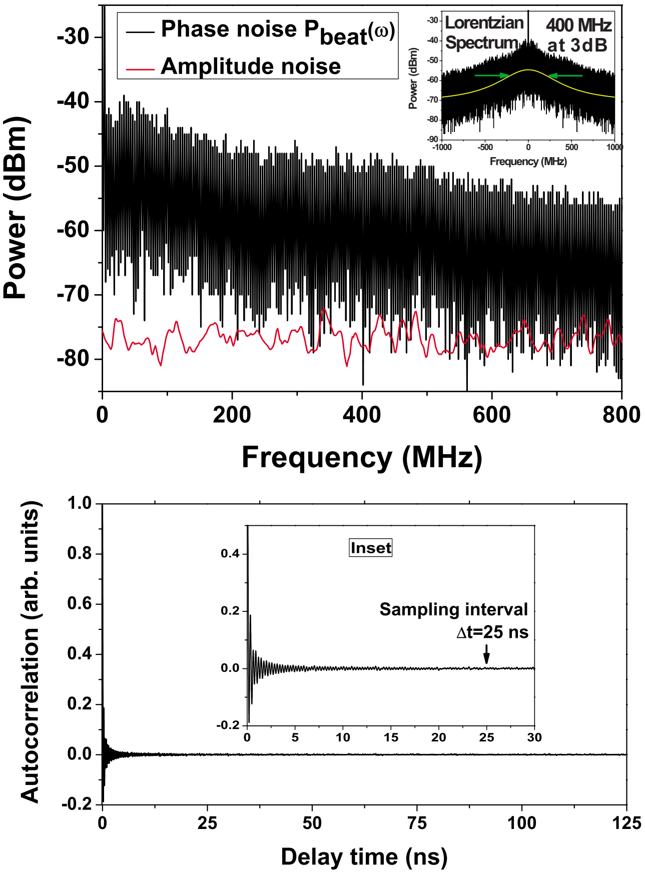

FIG. 2. (Color online) (a) The quantum phase (classical amplitude) noise of the laser field is observed with (without) the beat signal. Inset is the power spectral density of the beat signal. (b) Autocorrelation function of the beat signal vs time interval. In our experiment, the sampling interval of 25 ns (40 MHz sampling rate) is chosen.

be applied to obtain the autocorrelation function of the beat signal $[E_{beat}(t)]$ as $\langle E^{*}_{beat}(t)E_{beat}(t+\Delta t)\rangle$, where Δt is the sampled time interval of voltages for original random bit generation. Using $E_{beat}(t)=E(t)+E(t+\tau)$ and Eq. (1), it is evident that when the sampling interval Δt meets $\Delta t \gg \tau + \tau_{coh}$, no autocorrelation of the beat signal will be observed. Thus, the bits extracted from the beat signal are mutually independent and can be adopted to generate truly random numbers.

As shown in Fig. 1, in experiment, a 795 nm VCSEL laser works at 1.5 mA, a little above the threshold current 1.0 mA. The laser linewidth $\Delta\nu_{laser}=200$ MHz ($\tau_{coh}=1.59$ ns) is inversely proportional to the laser power, while the classical noises (e.g., occupation fluctuation and $1/f$ noise) are independent of the laser power [19,20]. Therefore, the quantum phase noise of laser dominates over its classical amplitude noise to ensure the true randomness of generated numbers. The delay time τ is set to be about 10 ns (corresponds to 3.0 m space delay) in order to fulfill $\tau \gg \tau_{coh}$. So, the self-homodyne method with delay time τ is used to obtain the beat signal with 3 dB linewidth of 400 MHz (detected by an APD) and its power spectral density is shown in the inset of Fig. 2(a). From Fig. 2(a), it can be seen that the classical

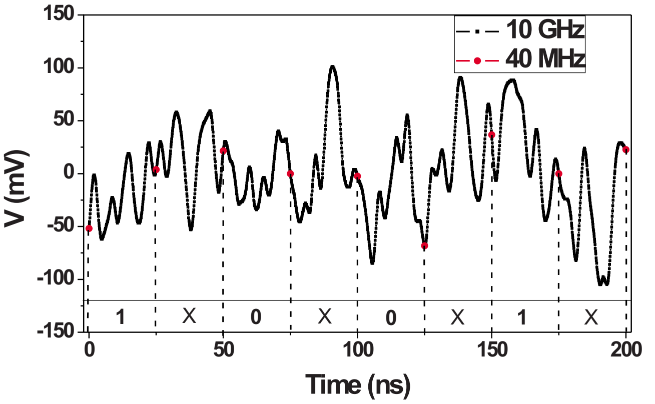

FIG. 3. (Color online) A 200 ns trace of the APD-detected voltages of the beat signal (small black dots) is recorded at 10 GHz, while the random signal (big red dots) is sampled at 40 MHz rate (25 ns interval). The final random bit is obtained from the least significant bit (LSB, i.e., its parity) of a sequence of 8-bit binary derivatives obtained by performing subtraction between two consecutive sampled voltages (shown in the bottom strip).

amplitude fluctuation is negligible compared to the quantum phase fluctuation within 200 MHz (the gap between them is about 20 dB). Using Wiener-Khintchine theorem [21,22], i.e.,

TABLE I. Results of Diehard statistical test suite. Data sample containing 100 Mbits is used for the Diehard test. For the cases of multiple p values, a Kolmogorov-Smirnov (KS) test is used to obtain a final P value, which measures the uniformity of the multiple p values. The test is considered successful if all final P values satisfy $0.01 \le P \le 0.99$.

Statistical test	P value	Result
Birthday spacings	0.910531[KS]	Success
Overlapping permutations	0.294899	Success
Ranks of 31×31 matrices	0.322213	Success
Ranks of 32×32 matrices	0.482575	Success
Ranks of 6×8 matrices	0.749427[KS]	Success
Monkey tests on 20-bit words	0.019887[KS]	Success
Monkey test OPSO	0.079864[KS]	Success
Monkey test OQSO	0.725649[KS]	Success
Monkey test DNA	0.293543[KS]	Success
Count 1's in stream of bytes	0.244463	Success
Count 1's in specific bytes	0.062188[KS]	Success
Parking lot test	0.806898[KS]	Success
Minimum distance test	0.326209[KS]	Success
Random spheres test	0.902946[KS]	Success
Squeeze test	0.815876[KS]	Success
Overlapping sums test	0.806025[KS]	Success
Runs test (up)	0.817356	Success
Runs test (down)	0.805323	Success
Craps test No. of wins	0.502035	Success
Craps test throws/game	0.403322	Success

TABLE II. Results of NIST statistical test suite. Using 1000 samples of 1 Mbits data and significance level $a=0.01$, for "Success," the P value (uniformity of p values) should be larger than 0.0001 and the proportion should be greater than 0.980 560 8 [25]. For the tests which produce multiple P values and proportions, the worst case is shown. As advised by NIST, the fast Fourier transform test is disregarded [26].

Statistical test	P value	Proportion	Result
Frequency	0.679846	0.9916	Success
Block frequency	0.248571	0.9897	Success
Cumulative sums	0.858032	0.9888	Success
Runs	0.816029	0.9907	Success
Longest run	0.648795	0.9935	Success
Rank	0.609895	0.9860	Success
Nonperiodic	0.569334	0.9823	Success
Overlapping	0.565500	0.9916	Success
Universal	0.143336	0.9888	Success
Approximate	0.590520	0.9879	Success
Random excursions	0.016388	0.9880	Success
Random variant	0.029796	0.9865	Success
Serial	0.946683	0.9916	Success
Linear complexity	0.732979	0.9915	Success

$$R_{beat}(t) = \int_{-\infty}^{+\infty} P_{beat}(\omega)\exp(-i\omega t)d\omega, \qquad (2)$$

the autocorrelation function $[R_{beat}(t)]$ of the beat signal is obtained from the power spectral density of the phase noise $[P_{beat}(\omega)$ in Fig. 2(a)] and is illustrated in Fig. 2(b). It can be seen from Fig. 2(b) that no correlation of the sampled voltages is observed when the sampling interval is set to meet $\Delta t \gg \tau + \tau_{coh}$. Accordingly, the sampling rate is chosen as 40 MHz in our experiment, i.e., $\Delta t=25$ ns, so the bits extracted from these sampled voltages are mutually independent.

These sampled voltages are digitized by an 8-bit analog-digital converter (ADC), which is shown as the red dots in Fig. 3. We take the least-significant bit (LSB) of each sampled 8-bit voltage as the original random bit, i.e., the parity of this 8-bit binary number, which represents whether the voltage falls in an even or odd bin of the total 256 bins. Since the probabilities of "even" and "odd" bins of the total 256 bins are not perfectly equal, the bit sequence shows a statistical bias δ which is smaller than 10^{-2} for 1 Gbit random bit sequence. To lower the bias, we perform a subtraction between two consecutive sampled voltages to obtain a sequence of $N/2$ 8-bit binary derivatives as $V_2-V_1, V_4-V_3, \ldots, V_N-V_{N-1}$, where N is the total number of the original sampled voltages. During this process, each voltage is used only once, so no correlation is introduced and the statistical bias is lowered to the magnitude of δ^2 (smaller than 10^{-4}). After that, we adopt the LSB of the 8-bit binary derivatives to generate the final random bits. Therefore, we obtain the final random bit at generation rate of 20 Mbit/s with a software-based processing. Note that, based on the independence of original sampled voltages, this processing enhances the performance of the random bits sequence by lowering the statistical bias while not introducing any additional correlations.

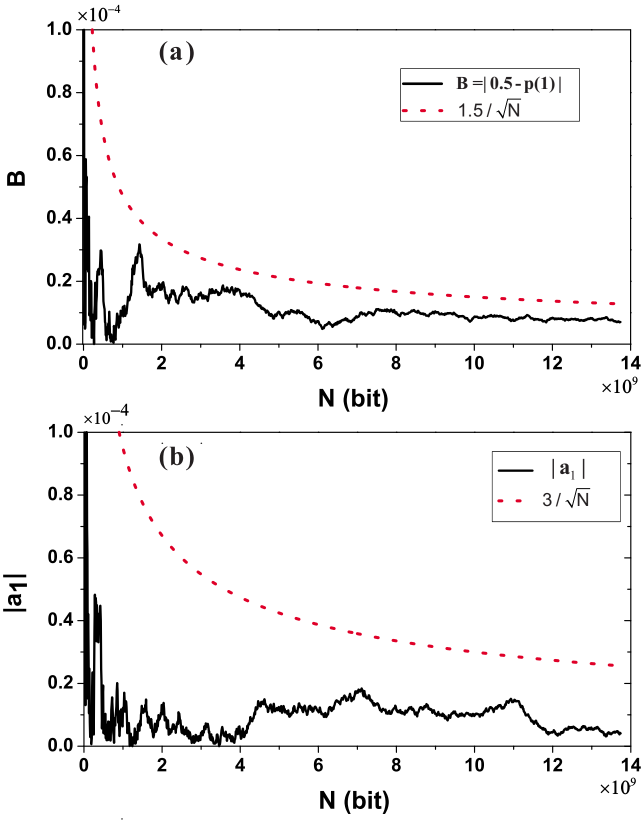

FIG. 4. (Color online) (a) The statistical bias (B) of the final random bit sequence. It can be seen that $B<1.5/\sqrt{N}$ always holds and converges to zero for large bit sequence, where $p(1)$ is the probability of ones in sequence. (b) The absolute value of the first-order correlation coefficient $|a_1|$ of the final random bit sequence. It can be seen that $|a_1|<3/\sqrt{N}$ always holds and $|a_1|$ converges to zero for large bit sequence.

We continuously record a final random bit sequence of 1 Gbit, which passes three standard randomness tests, i.e., a pseudorandom number sequence test program (ENT) [23], Diehard [24], and the National Institute of Standards and Technology-Statistical Test Suite (NIST-STS) [25]. The ENT results are entropy=1.000 000 bit per bit (the optimum compression would reduce the bit file by 0%). χ^2 distribution is 0.53 (randomly would exceed this value by 46.62% of the times). Arithmetic mean value of data bits is 0.5000. Monte Carlo value for π is 3.141 725 650. Serial correlation coefficient is −0.000 017. The Diehard and NIST-STS test results are shown in Tables I and II, respectively. Additionally, it should be noted that, for a TRNG, both the statistical bias and the absolute value of the first-order correlation coefficient of the final random bit sequence are expected to be smaller than three standard deviations $(3\sigma_1=1.5/\sqrt{N}$ for statistical bias [Fig. 4(a)] and $3\sigma_2=3/\sqrt{N}$ for correlation coefficient [Fig. 4(b)]) with the probability of 99.7%. In our case, both criteria are well satisfied for the continuously recorded final random bit sequence up to 14 Gbit. Here, we comment that, both for applications and tests for TRNG, the long-bit sequence with true randomness and desired length is rather crucial in practice.

It should be noted that, for a nonuniform distribution of the probability of 256 8-bit binary derivatives, if more than 1 bit are extracted from each 8-bit binary derivatives in order to improve the random bit generation rate (see, e.g., five LSBs are adopted in [13]), an additive correlation in the final random bit sequence will be introduced, even though this additive correlation is not so significant to fail the standard randomness tests. Taking five LSBs for an instance, due to the nonuniform distribution, every set of the five LSBs possesses a different probability and thus these five bits from the same set are correlated with some extent. Nevertheless, with this additive correlation within the same set, both the random bit sequence of extracting five LSBs (at sampling rate of 2.5 GHz in [13]) and four LSBs (at sampling rate of 40 MHz in our scheme) from an 8-bit binary number both successfully pass the three standard randomness tests. This fact also indicates that the standard randomness tests are only a way to examine whether the random bit stream is "sufficiently" random, but not to judge whether it is truly random.

We propose a simple approach to realize a high-speed TRNG, which is compact and convenient for implementation. The true randomness of our TRNG is physically guaranteed by the intrinsically random nature of the quantum phase noise originated from the spontaneous emission of photons. Moreover, here, the true randomness is verified by both the statistical bias and the correlation coefficient for long random bit sequence up to 14 Gbit. It is worth noting that this long random bit sequence possesses significant values for applications (even more important than the speed) because it is the length of the random bit sequence that is required in most applications and, essentially, it is a metric for qualifying the true randomness. Compared to the chaotic laser, the true randomness of intrinsic phase noise of a free-running laser is confirmed by two additional criteria, besides three standard tests. Further, this true randomness only depends on its inherently quantum-mechanical process and does not need the external optical feedback (which introduces a photon round trip period). Although the random bit generation rate is not as high as the PRNG based on chaotic laser(s) [12–14], its physically guaranteed true randomness, together with its simplicity and compactness, is attractive for the applications which need true randomness. Moreover, a much higher generation rate is attainable when both larger laser linewidth and faster data-acquisition hardware are applicable.

This work is supported by the Key Project of National Natural Science Foundation of China (NSFC) (Grant No. 60837004). We acknowledge the support from W. Jiang, K. Deng, X. X. Liu, C. Zhou, and G. D. Xie and the helpful discussions with B. Luo and J. B. Chen.

[1] S. L. Lohr, *Sampling: Design and Analysis* (Duxbury, Pacific Grove, 1999).
[2] J. E. Gentle, *Random Number Generation and Monte Carlo Methods (Statistics & Computing)*, 2nd ed. (Springer-Verlag, New York, 2003).
[3] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis* (Cambridge University Press, London, 2005).
[4] A. J. Menezes *et al.*, *Handbook of Applied Cryptography* (CRC, Cleveland, 1997).
[5] C. H. Bennett *et al.*, J. Cryptology **5**, 3-28 (1992).
[6] L. Blum *et al.*, SIAM J. Comput. **15**, 364 (1986).
[7] H. Schmidt, J. Appl. Phys. **41**, 462 (1970).
[8] M. Stipčević and B. Medved Rogina, Rev. Sci. Instrum. **78**, 045104 (2007).
[9] J. F. Dynes *et al.*, Appl. Phys. Lett. **93**, 031109 (2008).
[10] W. Wei and H. Guo, Opt. Lett. **34**, 1876 (2009).
[11] B. Qi *et al.*, Opt. Lett. **35**, 312 (2010).
[12] A. Uchida *et al.*, Nat. Photonics **2**, 728 (2008).
[13] I. Reidler, Y. Aviad, M. Rosenbluh, and I. Kanter, Phys. Rev. Lett. **103**, 024102 (2009).
[14] I. Kanter *et al.*, Nat. Photonics **4**, 58 (2010).
[15] C. Werndl, Br. J. Philos. Sci. **60**, 195 (2009).
[16] T. Jennewein *et al.*, Rev. Sci. Instrum. **71**, 1675 (2000).
[17] M. Lax, Phys. Rev. **160**, 290 (1967).
[18] C. H. Henry, IEEE J. Quantum Electron. **18**, 259 (1982).
[19] K. Vahala and A. Yariv, Appl. Phys. Lett. **43**, 140 (1983).
[20] K. Kikuchi and T. Okoshi, Electron. Lett. **19**, 812 (1983).
[21] N. Wiener, Acta Math. **55**, 117 (1930).
[22] A. Khintchine, Math. Ann. **109**, 604 (1934).
[23] J. Walker, http://www.fourmilab.ch/random/
[24] G. Marsaglia, Diehard: A Battery of Tests of Randomness, 1995 http://www.stat.fsu.edu/pub/diehard/
[25] http://csrc.nist.gov/groups/ST/toolkit/rng/stats_tests.html
[26] http://csrc.nist.gov/groups/ST/toolkit/rng/documentation_software.html