Research Article | Vol. 24, No. 14 | 11 Jul 2016 | OPTICS EXPRESS 15822

Optics EXPRESS

Brownian motion properties of optoelectronic random bit generators based on laser chaos

PU LI,[1,2] XIAOGANG YI,[1,2] XIANGLIAN LIU,[1,2] YUNCAI WANG,[1,2,4] AND YONGGE WANG[3,5]

[1]*Key Laboratory of Advanced Transducers and Intelligent Control System, Ministry of Education of China, Taiyuan, 030024, China*
[2]*Institute of Optoelectronic Engineering, Taiyuan University of Technology, Taiyuan, 030024, China*
[3]*University of North Carolina at Charlotte, Charlotte, NC 28223-0001, USA*
[4]*wangyc@tyut.edu.cn*
[5]*Yongge.Wang@uncc.edu*

Abstract: The nondeterministic property of the optoelectronic random bit generator (RBG) based on laser chaos are experimentally analyzed from two aspects of the central limit theorem and law of iterated logarithm. The random bits are extracted from an optical feedback chaotic laser diode using a multi-bit extraction technique in the electrical domain. Our experimental results demonstrate that the generated random bits have no statistical distance from the Brownian motion, besides that they can pass the state-of-the-art industry-benchmark statistical test suite (NIST SP800-22). All of them give a mathematically provable evidence that the ultrafast random bit generator based on laser chaos can be used as a nondeterministic random bit source.

OCIS codes: (190.3100) Instabilities and chaos; (140.5960) Semiconductor lasers; (140.1540) Chaos; (060.4785) Optical security and encryption.

References and links

1. A. Rukhin, J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Heckert, J. Dray, S. Vo, and L. E. Bassham III, "A statistical test suite for random and pseudorandom number generators for cryptographic applications," National Institute of Standards and Technology, Special Publication 800–22, Revision 1a, April 2010, NIST Statistical Tests Suite, [Online]. Available: http://csrc.nist.gov/groups/ST/toolkit/rng/documentation_software.html
2. J. Walker, "Hotbits: Genuine random numbers generated by radioactive decay," [Online]. Available: http://www.fourmilab. ch/hotbits.
3. RANDOM.ORG, [Online]. Available: http://www.random.org/.
4. B. Jun and P. Kocher, "The Intel random number generator," White Paper Prepared for Intel Corporation, Cryptography Research Inc., 1999, http://www.cryptography.com/resources/whitepapers/IntelRNG.pdf.
5. A. Uchida, K. Amano, M. Inoue, K. Hirano, S. Naito, H. Someya, I. Oowada, T. Kurashige, M. Shiki, S. Yoshimori, K. Yoshimura, and P. Davis, "Fast physical random bit generation with chaotic semiconductor lasers," Nat. Photonics **2**(12), 728–732 (2008).
6. I. Reidler, Y. Aviad, M. Rosenbluh, and I. Kanter, "Ultrahigh-speed random number generation based on a chaotic semiconductor laser," Phys. Rev. Lett. **103**(2), 024102 (2009).
7. I. Kanter, Y. Aviad, I. Reidler, E. Cohen, and M. Rosenbluh, "An optical ultrafast random bit generator," Nat. Photonics **4**(1), 58–61 (2010).
8. A. Wang, P. Li, J. Zhang, J. Zhang, L. Li, and Y. Wang, "4.5 Gbps high-speed real-time physical random bit generator," Opt. Express **21**(17), 20452–20462 (2013).
9. P. Li, Y. C. Wang, and J. Z. Zhang, "All-optical fast random number generator," Opt. Express **18**(19), 20360–20369 (2010).
10. N. Oliver, M. C. Soriano, D. W. Sukow, and I. Fischer, "Fast random bit generation using a chaotic laser: approaching the information theoretic limit," IEEE J. Quantum Electron. **49**(11), 910–918 (2013).
11. R. M. Nguimdo, G. Verschaffelt, J. Danckaert, X. Leijtens, J. Bolk, and G. Van der Sande, "Fast random bits generation based on a single chaotic semiconductor ring laser," Opt. Express **20**(27), 28603–28613 (2012).
12. A. Argyris, S. Deligiannidis, E. Pikasis, A. Bogris, and D. Syvridis, "Implementation of 140 Gb/s true random bit generator based on a chaotic photonic integrated circuit," Opt. Express **18**(18), 18763–18768 (2010).
13. X. Z. Li and S. C. Chan, "Heterodyne random bit generation using an optically injected semiconductor laser in chaos," IEEE J. Quantum Electron. **49**(10), 829–838 (2013).
14. N. Li, B. Kim, V. N. Chizhevsky, A. Locquet, M. Bloch, D. S. Citrin, and W. Pan, "Two approaches for ultrafast random bit generation based on the chaotic dynamics of a semiconductor laser," Opt. Express **22**(6), 6634–6646 (2014).

#263841 http://dx.doi.org/10.1364/OE.24.015822
Journal © 2016 Received 23 Apr 2016; revised 25 Jun 2016; accepted 29 Jun 2016; published 5 Jul 2016

Optics EXPRESS

15. X. Tang, Z. M. Wu, J. G. Wu, T. Deng, J. J. Chen, L. Fan, Z. Q. Zhong, and G. Q. Xia, "Tbits/s physical random bit generation based on mutually coupled semiconductor laser chaotic entropy source," Opt. Express **23**(26), 33130–33141 (2015).
16. M. Sciamanna and K. A. Shore, "Physics and applications of laser diode chaos," Nat. Photonics **9**(3), 151–162 (2015).
17. M. Virte, E. Mercier, H. Thienpont, K. Panajotov, and M. Sciamanna, "Physical random bit generation from chaotic solitary laser diode," Opt. Express **22**(14), 17271–17280 (2014).
18. Diehard Test Suite, [Online]. Available: http://www.stat.fsu.edu/pub/diehard/.
19. J. Walker, Ent: A Pseudorandom Number Sequence Test Program [Online]. Available: http://www.fourmilab.ch/random/
20. Y. Wang and T. Nicol, "Statistical properties of pseudo random sequences and experiments with PHP and Debian OpenSSL," in Proc. ESORICS 2014, Wroclaw, Poland, Sept. 2014, pp. 454–471.
21. Y. Wang, "Resource bounded randomness and computational complexity," Theor. Comput. Sci. **237**(1-2), 33–55 (2000).
22. W. Feller, "The fundamental limit theorems in probability," Bulletin of AMS **51**(11), 800–833 (1945).
23. P. Erdös and M. M. Kac, "On certain limit theorems of the theory of probability," Bulletin of AMS **52**(4), 292–303 (1946).
24. A. Y. Khinchin, "On a theorem of probability calculation," Fundam. Math. **6**, 9–20 (1924).
25. W. Feller, *Introduction to Probability Theory and its Applications* (John Wiley & Sons Inc., 1968).
26. J. A. Clarkson and C. R. Adams, "On definitions of bounded variation for functions of two variables," Tran. AMS **35**, 824–854(1933).

1. Introduction

In communications, the ultrafast and large-amount generation of nondeterministic random bit sequences is a mandatory requirement for the information-theoretically secure cipher, 'one-time pad'.

One can easily generate random bits utilizing some certain algorithms and seeds, but they are deterministic and commonly called as pseudo-random bits. For instance, cryptographic hash functions (SHA1, SHA2, and SHA3) and symmetric key block ciphers (AES and TDES) have been widely used to design pseudo-random bit generators . However, they may not satisfy the properties of randomness and thus be distinguished from a uniformly chosen sequence, even though these pseudo-random bit generators success to pass the existing statistical tests (*e.g.*, NIST SP800-22 [1]).

Employing stochastic physical processes, one can generate nondeterministic random strings. For example, the radioactive decay [2], atmospheric noise [3] and thermal noise [4] have been extensively used to construct random bit generator (RBG) in practice. However, these conventional nondeterministic RBGs are limited to a very slow bit rate at Mb/s. It is far below the current communication rates at Gb/s.

In recent years, the optoelectronic RBG based on laser chaos attract considerable attention due to its brilliant bit rate [5–17]. For instance, Uchida *et al.* initiated the study of fast random bit generations using the physical chaos in laser diodes and produced experimentally 1.7 Gb/s random bit outputs in 2008 [5]. Then, Reidler *et al.* enhanced chaotic the generation rates to 12.5 Gb/s and 300 Gb/s [6,7]. Our previous studies also reported that random bits can be generated in a real-time rate of 4.5 Gb/s using optoelectronic techniques [8,9]. More recently, Tang *et al.* even analyzed the feasibility of Tb/s random bit generation using chaotic lasers [15].

However, there has been a dispute whether the outputs of the RBGs based on laser chaos are nondeterministic. Empirical statistical test suites [1, 18,19] are only can be used in determining whether or not a generator produces high-quality random bits, but not in ensuring the non-deterministic property. For example, NIST SP800-22 proposed a state of the art statistical test technique to determine whether a RBG is suitable for a particular cryptographic application [1]. For an in-depth analysis, NIST SP800-22 recommends that some additional statistical tests such as χ^2-test should be performed. But, the NIST SP800-22 test suite still has several drawbacks. Barker, one of the chief designers of the NIST SP800-22 test suite, pointed out that we are not really sure how good these tests are. Moreover, Yongge Wang *et al.* also demonstrated that the NIST SP800-22 test suite has inherent limitations, which mean the straightforward Type II errors [20,21].

Optics EXPRESS

In fact, Feller proved that a nondeterministic random binary string can be rigorously and mathematically convinced by two fundamental limit theorems. One is the central limit theorem and the other is the law of iterated logarithm (LIL) [22]. The NIST SP800-22 test suite includes several frequency related tests, which cover the central limit theorem. "the cumulative sums test" in the test suite even covers the limit distribution of the maximum of the absolute values of the partial sums $\overline{S}_n$. That is firstly obtained by Erdös and Kac [23]. But the NIST SP800-22 test suite does not include any test for the LIL. The LIL says that, for a nondeterministic random sequence, the value S_{lil} should stay in [-1, 1] range and reach both ends infinitely often when the sequence length n goes to infinite. In other words, the LIL gives a complete characterization of the Wiener process, often called as Brownian motion.

In this paper, we experimentally analyze the nondeterministic property of the RBG based on laser chaos from two aspects of the central limit theorem and the LIL. In particular, we investigate the Brownian motion property of the RBG according to the LIL, considering the NIST SP800-22 testing suite has covered the central limit theorem. Specifically, 100 GBytes (= 800 Gbits) random bits are used, which is extracted from a chaotic laser diode using a multi-bit extraction technique. The results demonstrate that the random bits from the chaotic RBG can simulate the Brownian motion perfectly. All our experiments give a mathematically proved evidence that the output of the RBG based laser chaos is nondeterministic.

2. Experimental setup

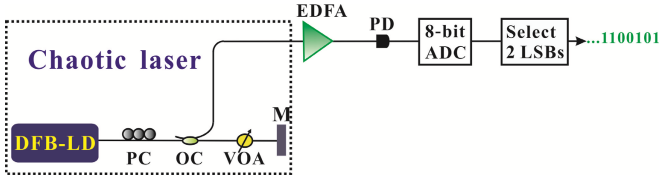

Fig. 1. Experimental setup. DFB-LD, distributed feedback laser diode; PC, polarization controller; OC, optical coupler; VOA, variable optical attenuator; M, fiber mirror; EDFA, erbium-doped fiber amplifier; PD, photodetector; 8-bit ADC, 8-bit analog-to-digital converter; LSBs, least significant bits.

The schematic of the optoelectronic RBG based on laser chaos is illustrated in Fig. 1. As plotted in the dash box, the chaotic laser consists of a DFB laser diode (DFB-LD) with optical feedback, working at 1.6 times its threshold current (22 mA) and in 1.55 μm band. The DFB-LD connects to an 60:40 optical coupler (OC) whose secondary output passes through a variable optical attenuator (VOA) onto a fiber mirror (M), forming a feedback cavity with a time-delay of 105.453 ns and a feedback strength of 1.32%. Such an optical feedback drives the DFB-LD into a chaotic oscillation, whose polarization is adjusted through a polarization controller (PC). The chaotic light exiting the 60% port of the OC passes through an erbium-doped fiber amplifier (EDFA) onto a 50 GHz bandwidth of photoelectric detector (PD). The PD converts the chaotic light into the corresponding chaotic electrical signal. Then we use this chaotic electrical signal to generate high-speed random bit sequences in the following way. The detected chaotic signal is sampled and quantized by an electrical 8-bit analog-to-digital converter (8-bit ADC). Every sample point will correspond to 8-bit Boolean values. Through retaining just the 2 least significant bits (LSBs), we can harvest random bit sequences.

3. Chaos characterisitcs and random bit generation

Figure 2 shows the measured temporal waveform of the chaotic signal and its associated RF spectrum detected by the 50 GHz PD (u2T XPDV2120RA). Comparing with the background noise floor obtained by extinguishing the DFB-LD (WTD VDM5S752), the chaotic signal possesses a large intensity fluctuation and high bandwidth. This high signal level may make the random bit extraction not very susceptible to bias from small nonrandom external perturbations including temperature fluctuations.

Optics EXPRESS

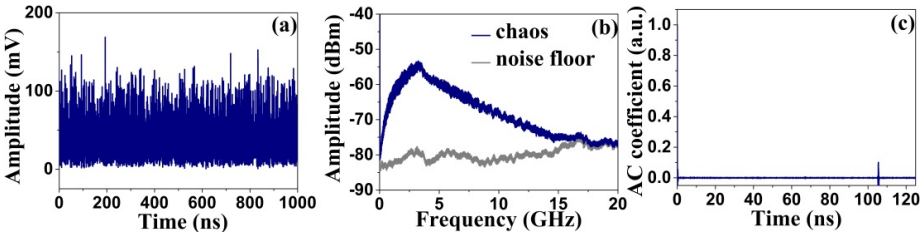

Fig. 2. (a) Measured temporal waveform, (b) RF spectrum and (c) AC function of the chaotic signal.

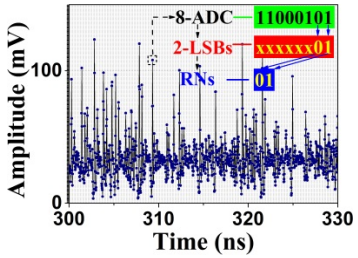

Fig. 3. Random bit extraction procedure from the chaotic signal. Blue dots denote the sample points.

Figure 3 is the partial enlargement of Fig. 2(a) from 300 to 330 ns. The extraction of random bits is illustrated through a flowchart at the top right corner. Every sample point [blue dot] is firstly converted into an 8-bit binary Boolean sequence by the 8-bit ADC in a 36 GHz bandwidth oscilloscope (Lecory LabMaster10-36Z). The sampling rate of the 8-bit ADC is set at 40 GSa/s. Through discarding the first 6 bits and only retaining the 1-st and 2-ed LSBs, the ultrafast random bits can be harnessed. Benefit from the high sampling rate of 40 GSa/s and 2 LSBs remained in the experiment, our RBG possesses a bit rate of 80 Gb/s (= 40 GSa/s × 2). Note, there is no strict requirement for the sampling rate, except that it should not be in proportion to the time-delay (TD) of the laser external cavity.

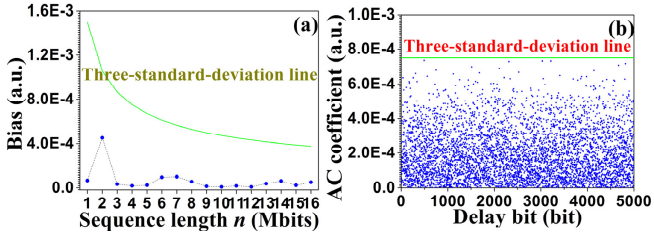

Fig. 4. (a) Bias of the generated 80 Gb/s random bit sequence with different lengths n = 1, …, 16 Mbits, where the solid line is its three-standard-deviation line, $3\sigma_{bias} = (3n^{-1/2})/2$. (b) AC coefficients as a function of the delay bit for the random bit sequence with a length n = 16 Mbits where the solid line is its three-standard-deviation line, $3\sigma_{AC} = 3n^{-1/2}$.

Figures 4(a) and 4(b) give preliminary measurements of the bias and the normalized autocorrelation (AC) coefficient of the obtained 80 Gb/s random bit stream, respectively. Both of them are then estimated using the Gaussian distribution estimation $\boldsymbol{N}(0, \sigma^2)$, where the standard deviation σ for the bias $\sigma_{bias} = (n^{-1/2})/2$ but for the AC coefficient $\sigma_{AC} = n^{-1/2}$. Note, the number n denotes the length of the used random bit stream. The red solid lines in Figs. 4(a) and 4(b) are their corresponding three-standard-deviation lines, $3\sigma_{bias}$ and $3\sigma_{AC}$. It is obvious that both the bias and AC coefficient keep below their own three-standard-deviations, so the generated random bit sequence can be considered to be statistically unbiased and independent.

Optics EXPRESS

4. Non-deterministic property analysis

As aforementioned, Feller has proved that a nondeterministic random binary string can be rigorously and mathematically convinced by the two fundamental limit theorems: the central limit theorem and the LIL [22]. Therefore, our nondeterministic property analysis on the RBG will be executed in two steps.

In the first step, we use the state-of-the-art NIST SP800-22 test suite to qualify the statistical randomness of the RBG, because its frequency-related tests and the cumulative sums test cover the central limit theorem. The NIST SP800-22 test suite contains 15 types of statistical tests. Here, 1000 samples of 1 Mb random bits are applied and the significant level α is chosen to be 0.01. In this case, the passing criterion is that the uniformity of the p-values (the P-value) is larger than 0.0001 and the proportion of the sequences satisfying p-value > α is in the range of 0.99 ± 0.0094392. Figure 5 is a typical result of the NIST SP800-22 tests, where the left and right diagram shows the P-value and passed proportion of each tests, respectively. The numbers on the horizontal axis represent 15 different statistical tests in NIST testing suite, which are named as 'Frequency', 'Block frequency', 'Cumulative sums', 'Runs', 'Longest-run', 'Rank', 'FFT', 'Non-periodic templates', 'Overlapping templates', 'Universal', 'Approximate entropy', 'Random excursions', 'Random excursions variant', 'Serial' and 'Linear Complexity', respectively. This result indicates that all NIST tests have been passed. Especially, the success of all frequency-related tests and the cumulative sums test demonstrates that the random bits from our RBG satisfy the central limit theorem very well.

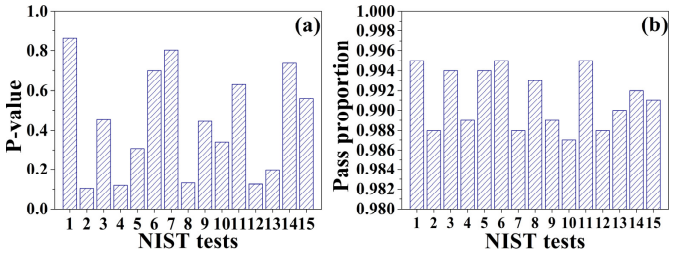

Fig. 5. Typical results of NIST statistical tests. Using 1000 samples of 1-Mb data and significance level $\mathcal{D} = 0.01$, for "Success," the P-value (uniformity of p-values) should be larger than 0.0001 and the proportion should be greater than 0.9805608. The numbers on the horizontal axis represent 15 different statistical tests in the NIST test suite, which are named as 'Frequency', 'Block frequency', 'Cumulative sums', 'Runs', 'Longest-run', 'Rank', 'FFT', 'Non-periodic templates', 'Overlapping templates', 'Universal', 'Approximate entropy', 'Random excursions', 'Random excursions variant', 'Serial' and 'Linear Complexity', respectively.

In the second step, we introduce the LIL to analyze the Brownian motion property of the RBG. The LIL describes the exact □uctuation scales of a random walk. Use a random binary sequence ξ to describe a random walk of a particle and assume that the particle starts from position 0 at time 0. The particle moves one position up if the next bit is 1 and one position down if the next bit is 0. Then, the particle position at time n can be characterized by the value $S(\xi[0..n\text{-}1])$ and the value $n\text{-}S(\xi[0..n\text{-}1])$. Herein, $S(\xi[0..n\text{-}1])$ denotes the number of 1s within the string $\xi[0..n\text{-}1]$, whereas $n\text{-}S(\xi[0..n\text{-}1])$ represents the number of 0s. For a random walk, a particle could reach any positive and any negative position when n is large enough. That is, $n\text{-}2 \times S(\xi[0..n\text{-}1])$ tends to infinity when n tends to infinity. In order to exactly characterize the path of a random walk, one commonly uses the reduced number $S^*(\xi[0..n\text{-}1])$ of 1s within $\xi[0..n\text{-}1]$.

Specifically, the following values are defined firstly for a nonempty string $x \in \{0, 1\}^*$:

$$S(x) = \sum_{i=0}^{|x|-1} x[i] \quad \text{and} \quad S^*(x) = \frac{2 \cdot S(x) - |x|}{\sqrt{|x|}} \tag{1}$$

Optics EXPRESS

, where $S(x)$ denotes the number of 1s in x and $S^*(x)$ denotes the reduced number of 1s in x. $S^*(x)$ amounts to measuring the deviations of $S(x)$ from $|x|/2$ in units of $\sqrt{|x|}/2$. The law of large numbers says that, for a random sequence ξ, the limit of $S(\xi[0..n\text{-}1])/n$ is 1/2, which corresponds to the frequency test in NIST SP800-22. But it says nothing about the reduced deviation $S^*(\xi[0..n\text{-}1])$. It is intuitively clear that $S^*(\xi[0..n\text{-}1])$ for a random sequence ξ will sooner or later take on arbitrary large values (though slowly). The law of LIL, which was first discovered by Khintchine [24], gives an optimal upper bound $\sqrt{2\ln\ln n}$ for the fluctuations of $S^*(\xi[0..n\text{-}1])$. Based on this fact, for a sequence $\xi \in \{0, 1\}^\infty$, we can define:

$$S_{lil}(\xi \upharpoonright n) = \frac{2\sum_{i=0}^{n-1}\xi[i] - n}{\sqrt{2n\ln\ln n}} \tag{2}$$

Then, for each random sequence $\xi \in \{0, 1\}^\infty$, we have both $\limsup_{n\to\infty} S_{lil}(\xi \upharpoonright n) = 1$ and $\liminf_{n\to\infty} S_{lil}(\xi \upharpoonright n) = -1$.

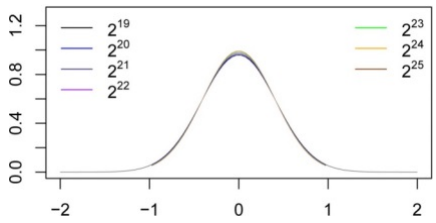

Fig. 6. Density functions for distributions μ_n^U with $n = 2^{19}, \ldots, 2^{25}$.

We will use this function $S_{lil}(\xi \upharpoonright n)$ to execute the LIL technique . The distribution induced by the function $S_{lil}(\xi \upharpoonright n)$ defines a probability measure on the real line R. Let $R \subset \{0,1\}^n$ be a set of m sequences with a uniform probability definition. That is, for each $x_0 \in R$, let $Prob[x = x_0] = 1/m$. Then each set $R \subset \{0,1\}^n$ induces a probability measure $\mu_n^R(I) = Prob[S_{lil}(x) \in I, x \in R]$ for each Lebesgue measurable set I on R. For $U = \{0,1\}^n$, we use μ_n^U to denote the corresponding probability measure. If R_n is the collection of all length n sequences generated by a nondeterministic RBG, the difference between μ_n^U and $\mu_n^{R_n}$ is negligible. For a uniformly chosen ξ, the distribution of $S^*(\xi \upharpoonright n)$ could be approximated by a Gaussian distribution of mean 0 and variance 1 with error bounded by $1/n$ [25]. In other words, the measure μ_n^U can be calculated as $\mu_n^U((-\infty, x]) \approx \Phi(x\sqrt{2\ln\ln n}) = \sqrt{2\ln\ln n}\int_{-\infty}^{x}\phi\left(y\sqrt{2\ln\ln n}\,dy\right)$. Figure 6 shows the distributions μ_n^U for $n = 2^{19}, \cdots, 2^{25}$. For the reason of convenience, we will use $\mathcal{B}$ as the discrete partition of the real line R defined by $\{(\infty,1),[1,\infty)]\} \cup \{[0.05x-1, 0.05x-0.95) : 0 \le x \le 39\}$.

Optics EXPRESS

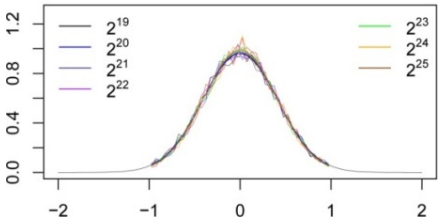

Fig. 7. Density functions for distributions μ_n^{Rlaser} for $n = 2^{19}, \ldots, 2^{25}$ with 6000 bit strings from our generator

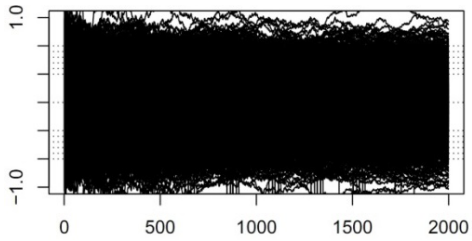

Fig. 8. LIL plot for the laser chaos based RBG with 6000× 17 MB strings

There are various definitions of statistical distances for probability measures. In our analysis, we will consider the total variation distance $d(\mu_n^{R_{laser}}, \mu_n^U) = \sup_{A \subseteq B} \left| \mu_n^{R_{laser}}(A) - \mu_n^U(A) \right|$ [26]. For nondeterministic m sequences, the expected distance $d(\mu_n^{Rlaser}, \mu_n^U)$ should be smaller than $4.6983 \times m^{-0.57}$. In our experiment, 100 GB (= 800 Gb) random bits from the RBG are divided into 6000 sequences so that each sequence consists of 17 MB (= 143165576 bits) data. After calculating, we get that the total variation distance for our generated 6000 sequences is 0.029, smaller than the expected total variation distance $d(\mu_n^{Rlaser}, \mu_n^U) = 4.6983 \times 6000^{-0.57} = 0.033$.

Our analysis results can be also visually illustrated in Figs. 7 and 8. Figure 7 is the associated distributions of μ_n^{Rlaser} for $n = 2^{19}, \ldots, 2^{25}$. These curves are plotted on top of the expected distribution μ_n^U in Fig. 6. From Fig. 7, it can be clearly observed that our results have nearly no distance from the expected distribution for an ideal nondeterministic random bit generator. As aforementioned, we can interpret each sequence ξ as a particle. That is, it locates at the position 0 at time 0 and at the position $S^*(\xi[0..n\text{-}1])$ at time n. By using x-axis to denote the time and y-axis to denote the particle's position, we can get the position curves of the 6000 particles corresponding to the 6000 sequences, as shown in Fig. 8. If the 6000 sequences are nondeterministic, the corresponding 6000 particles position curves should follow the Brownian motions curves. That is, they should mainly move between the y-axis interval [-1, 1] and reach both the line $y = -1$ and $y = 1$ infinitely often. The curves in Fig. 8 visually show that the output of our RBG can simulate the Brownian motion perfectly. All of them demonstrate that the laser chaos based RBG also satisfies the LIL.

5. Discussions

5.1 Justification why the bit truncation is used and why just 2-LSBs are retained

The reason why the bit truncation is used in our random bit generator (RBG) comes from two aspects. One is the asymmetric distribution of chaotic amplitudes, which can be roughly observed from Fig. 2(a). The other is the so-called "time-delay (TD) signature" which is an inherent weakness of the chaotic laser with optical feedback. Specifically, the TD signature is induced by the external feedback cavity. It can be located through the autocorrelation (AC) function in Fig. 2(c). From Fig. 2(c), one can see that a residual correlation peak with an AC coefficient about 0.100 appears at the delay time of 105.453 ns. The residual peak

Optics EXPRESS

corresponds to the TD signature of the chaotic laser adopted in our experiment and the associated AC coefficient is a quantitative indicator for the TD signature.

The bit truncation is a common post-processing method for improving the distribution uniformity and for destroying the TD signature in the chaotic dynamics [6,7, 10–17]. Figure 9 shows the normalized distributions of the decimal quantization values (in integer representation) generated by retaining *m*-LSBs: (a) *m* = 1; (b) *m* = 2; (c) *m* = 3; (d) *m* = 4; (e) *m* = 5; (f) *m* = 6; (g) *m* = 7; (h) *m* = 8. From Fig. 9(h), one can see clearly that the 8-LSBs completely inherit the feature of asymmetric distribution from the original chaotic signal [Fig. 2]. With the decrease of the number *m*, the distribution uniformity gets gradually improved from Fig. 9(g) to Fig. 9(d). In particular, the distributions become totally uniform when only 3-LSBs or lower are retained as shown from Fig. 9(c) to Fig. 9(a). At the same time, it can be verified that the TD signatures within the associated binary sequences have also been substantially reduced through retaining 3-LSBs or lower. Figure 10(a-i), 10(a-ii) and 10(a-iii) depict the AC functions of the binary sequences by retaining 1-LSB, 2-LSBs and 3-LSBs, respectively. Note, the *m*-LSBs binary sequences are obtained by interleaving 1-st LSB to *m*-th LSB together, as commonly done in Refs [6,7, 10–17].

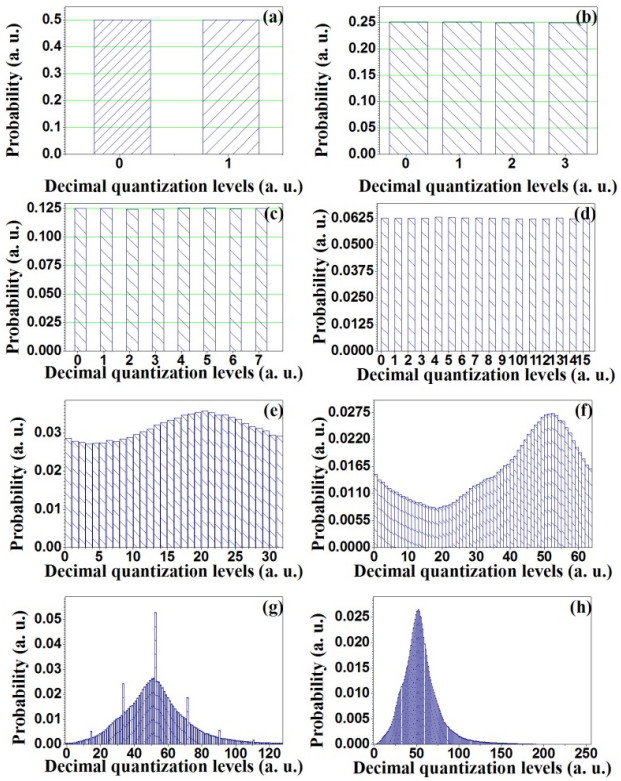

Fig. 9. Normalized distributions for the decimal quantization values (in integer representation) generated by retaining *m*-LSBs for cases: (a) *m* = 1, (b) *m* = 2, (c) *m* = 3, (d) *m* = 4, (e) *m* = 5, (f) *m* = 6, (g) *m* = 7 and (h) *m* = 8.

However, it must be pointed that the 3-LSBs binary sequences fail to simulate the Brownian motion in our experiment, although they can successfully pass all the NIST tests. Only when we retain 2-LSBs and 1-LSB, the corresponding binary sequences can simultaneously satisfy the NIST test suite and the Brownian motion. Therefore, only 2-LSBs (not 3-LSBs) are selected to generate random bit in our system. To figure out the reason why the 3-LSBs fails, we calculate the AC functions of the independent (not interleaved) binary sequences by retaining 1-st LSB, 2-nd LSB and 3-rd LSB, as shown from Fig. 10(b-i) to Fig. 10(b-iii). It can be observed clearly that there is a great difference between the AC functions

Optics EXPRESS

of the 3-LSBs [Fig. 10(a-iii)] and 3-th LSB [Fig. 10(b-iii)]. That is, the TD signature in 3-LSBs actually has not been thoroughly eliminated when 3-th LSB is observed alone.

In addition, we point that the generated bit rate is determined by the product of the selected number m of LSBs and the sampling rate of the adopted ADC. Only 2-LSBs are retained and the 8-bit ADC works at a sampling rate of 40 GSa/s in our experiment, so the associated bit rate of our RBG is 80 Gb/s (= 2 × 40 GSa/s). The more the alternative LSBs are, the faster the generated bit rate will be.

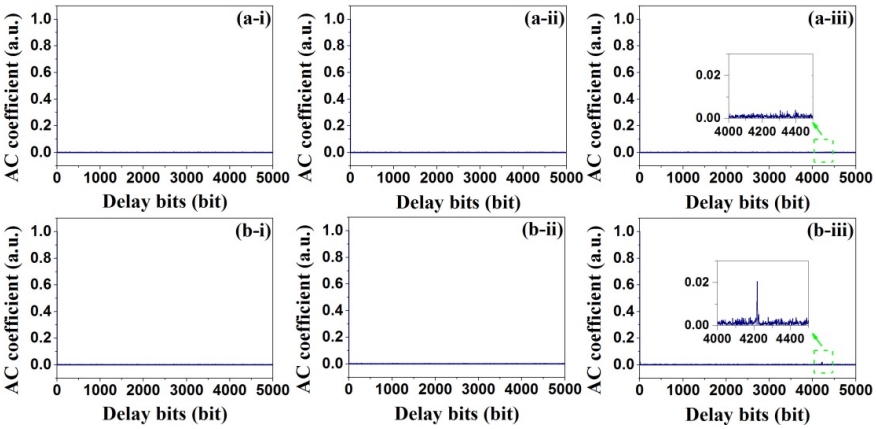

Fig. 10. AC functions of the digitized binary sequences from (a) interleaved m-LSBs and (b) independent m-th LSB: (a-i) 1-LSBs, (a-ii) 2-LSBs, (a-iii) 3-LSBs, (b-i) 1-st LSB, (b-ii) 2-nd LSB and (b-iii) 3-rd LSB. The insets in (a-iii) and (b-iii) are the magnified ranges corresponding to the TD signature of the chaotic dynamics.

5.2 Requirement on the physical side to guarantee the non-determinism property

From the discussions on Figs. 9 and 10 aforementioned, we can get at least two necessary conditions for a random bit generator (RBG) based on the laser chaos and multi-bit extraction technique to simulate Brownian motion perfectly. One is that all decimal quantization windows should be equiprobable. The other is that every independent (not interleaving) binary bit sequences by retaining m-th LSB (not m-LSBs) must have no residual correlations inherited from the original chaotic laser (*i.e.*, the TD signature). In the scenario using the bit truncation to generate random bits, the latter necessary condition is more crucial than the former.

Ideally, it is the favorite that the TD signature level of chaotic signal should be 0. Note, the TD signature level denotes the AC coefficient at the TD location. However, it is very difficult to completely eliminate the TD signature of a chaotic laser with optical feedback from the root, especially when the external cavity is long. The operation state in Fig. 2 is the optimum obtained in our experiment after we carefully adjusted the external parameters such as the injection current, external cavity time-delay, feedback strength and polarization. In this optimized state, we only get a TD signature level about 0.1.

Nevertheless, there is still an experience worth to be mentioned on the TD selection of the chaotic laser. In the process of optimizing the operation state of our chaotic laser, we find that an incommensurable external cavity is the key, after the other external parameters are optimal. That is, the external-cavity TD τ should satisfy the criterion $l \times \tau \neq t \times \tau_{sampling}$ for any low order integers l and t, where $\tau_{sampling}$ is the sampling interval of the adopted ADC. This criterion is empirical, but it can be convinced through a simulation based on the Lang-Kobayashi (LK) equations. The chaotic system with time-delayed optical feedback from an external mirror can be modeled by the following two rate equations.

Optics EXPRESS

$$\frac{dE(t)}{dt}=\frac{1+i\alpha}{2}\left[\frac{G_n[N(t)-N_0]}{1+\varepsilon|E(t)|^2}-\frac{1}{\tau_p}\right]E(t)+\kappa_f E(t-\tau)\exp(-iw\tau) \quad (3)$$

$$\frac{dN(t)}{dt}=\frac{J}{e}-\frac{N(t)}{\tau_e}-\frac{G_n[N(t)-N_0]}{1+\varepsilon|E(t)|^2}|E(t)|^2 \quad (4)$$

Herein, $E(t)$ and $N(t)$ denote the slowly varying complex electric field and carrier density of the chaotic laser, respectively. To approach the operation state in experiment, these adopted parameters are listed as follows: external feedback strength κ_f = 0.01, transparency carrier density $N_0 = 1.5 \times 10^8$ μm^{-3}, differential gain $G_n = 1.414 \times 10^{-3}$ μm^{-3}ns^{-1}, carrier lifetime τ_e = 2.5 ns, photon lifetime τ_p = 1.2 ps, line-width enhancement factor α = 5, gain saturation parameter $\varepsilon = 5 \times 10^{-5}$ μm^3 and injection current $J = 1.6J_{th}$ (where threshold current J_{th} = 22 mA). The time-delay τ induced by the external feedback cavity will be specified later. In the simulation, we use the fourth-order Runge-Kutta algorithm with a 1 ps time-step to perform Eqs. (3) and (4). The simulated chaotic signals are recorded with a fixed sampling interval of 25 ps, corresponding to the 40 GHz sampling rate of the 8-bit ADC adopted in the experiment. Figure 11 show the RF spectrums and AC functions for two simulated chaotic signal. One has a commensurable external cavity with the sampling rate of 40 GHz, whose time-delay τ equals to 100.000 ns. Figure 11(a-i) and 11(a-ii) is its simulated RF spectrum and AC function. The other has an incommensurable external cavity which has the same time-delay τ = 105.453 ns as the adopted chaotic laser in the experiment. Figure 11(b-i) and 11(b-ii) is the associated RF spectrum and AC function for the latter chaotic signal. The TD signature level of 0.098 in Fig. 11(b-ii) is obviously lower than that in Fig. 11(a-ii). This verifies our empirical criterion for the selection of an incommensurable external cavity length.

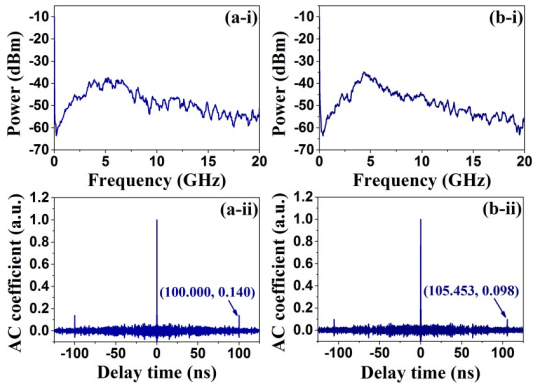

Fig. 11. Simulated RF spectrum and AC function for two different time-delays of chaotic signals. (a-i) and (a-ii) correspond to the RF spectrum and AC function for the chaotic signal with a time-delay of τ = 100.000 ns, whereas (b-i) and (b-ii) are the counterparts for the other chaotic signal with the same incommensurable time-delay of τ = 105.453 ns as the chaotic laser in experiment.

The introduction of the bit truncation can substantially reduce the level of the TD signature in the original chaotic signal further. However, it must be kept in mind that the bit truncation is not omnipotent. Figure 12 depicts all the calculated AC functions of independent m-th LSB from the aforementioned two simulated chaotic signals. Figure 12(a) illustrates these AC functions corresponding to the chaotic signal with a TD signature level of 0.140 [Fig. 11(a)]. Figure 12(b) illustrates these AC functions corresponding to the chaotic signal with a TD signature level of 0.098 [Fig. 11(a)]. These plots marked with Roman numerals i, ii, iii, iv, v, vi, vii and viii corresponds to the AC functions of the associated 1-th, 2-nd, 3-rd, 4-th, 5-th, 6-th, 7-th and 8-th LSB binary sequence. Through extracting every TD signature

Optics EXPRESS

levels, we can illustrate a variation trend of the TD signature level as a function of the *m*-th LSB in Fig. 13. From Fig. 13, it can be seen clearly that the TD signature level is gradually reduced with the decrease of the number *m*. Moreover, when the level of the TD signature of the 8-th LSB binary sequence is too high, the bit truncation will fail even if only the 1-LSB is retained.

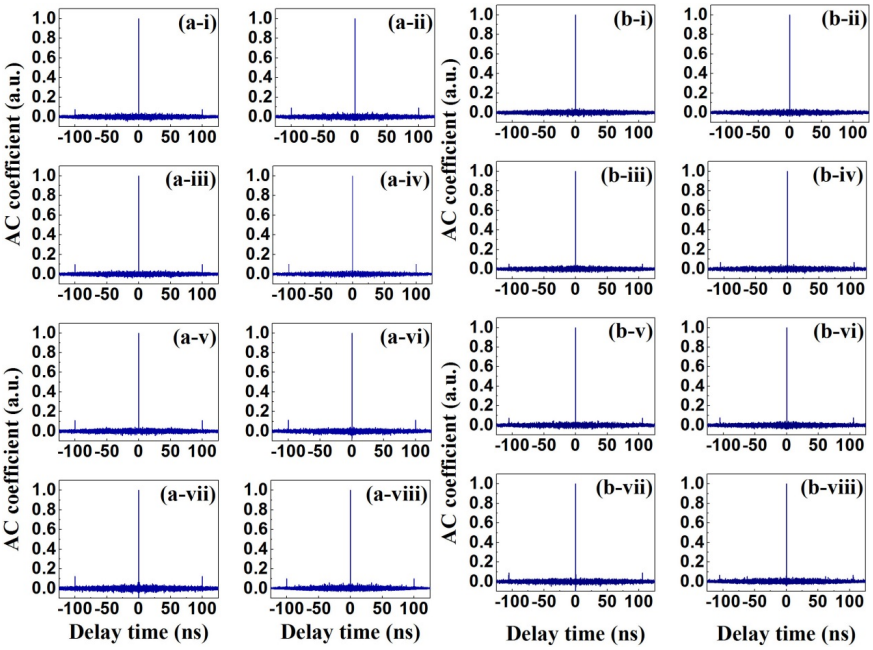

Fig. 12. AC functions of the independent *m*-th LSB binary sequences from the two simulated chaotic signal with different TD signature levels, where *m* = 1, 2, 3, 4, 5, 6, 7 and 8. Note, the first and second columns correspond to the chaotic signal with a high TD signature level of 0.140 in Fig. 8(a), while the third and fourth columns correspond to the other chaotic signal with a low TD signature level of 0.098 in Fig. 8(b). Inset numbers from i to viii in the plot corresponds to *m* value from 1 to 8 in order.

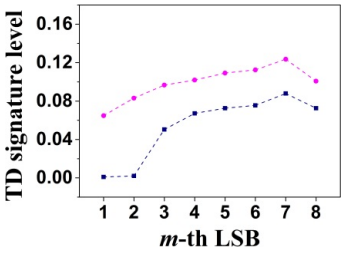

Fig. 13. Extracted TD signature levels from Fig. 12 as a function of *m*-th LSB, where *m* = 1, 2, 3, 4, 5, 6, 7 and 8. Note, the dots and squares corresponds to the TD signature levels in Fig. 12(a) and Fig. 12(b), respectively.

Next, we attempt to figure out a quantitative requirement for the TD signature (a crucial physical quantity) of the chaotic laser. Our investigation is based on the second necessary condition for the RBG to simulate Brownian motion perfectly mentioned at the beginning. Through checking the simulation results in Fig. 13 carefully, we can quantitatively determine that 2 LSBs can be retained when the TD signature level of the 8-th LSB binary sequence is 0.06. In other words, when the TD signature level of an 8-th LSB binary sequence is lower than 0.06, it can be guaranteed that at least 1 LSB have no residual correlation of the TD signature. Therefore, a tolerable TD signature level of the 8-th LSB binary sequence can be determined below the value of 0.06.

Optics EXPRESS

We notice that there is a direct relevance between the AC function of the extracted 8-th LSB (also called as 1-st most significant bit (MSB) in an 8-bit ADC) binary sequence and that of the associated chaotic signal. The 8-th LSB sequence is obtained by comparing the time series of the chaotic signal with its median value. Therefore, the 8-th LSB binary sequence is equivalent to the output of a 1-bit ADC with a threshold, which is equal to the median value of the chaotic signal. Perfect 1-bit analog-to-digital conversion suggests that the digitized binary sequence should inherit all the information from the analog signal. That means that the AC function of 8-th LSB binary sequence should be appropriately equal to that of the signal to be digitized in theory. However, some of the information will inevitably be lost during the conversion, due to the quantization error. So the AC function of 8-th LSB binary sequence is usually lower than that of the signal to be digitized in practice. Based on this fact, we can conservatively give a requirement on the physical side to guarantee the non-determinism property of the RBG. This requirement is that the TD signature level of the original chaotic laser should be lower than 0.06, although a higher level may be also tolerated actually.

6. Conclusions

In conclusion, we have experimentally analyzed the stochastic properties of the optoelectronic RBG based on laser chaos from the aspects of the central limit theorem and LIL. The Brownian motion performance of the RBG is mainly investigated. The experimental results show that the random bits extracted from the chaotic laser diode have no statistical distance from the Brownian motion. All of our results demonstrate that the ultrafast RBG based on laser chaos can be used as a nondeterministic randomness source.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant Nos. 61227016, 61505137 and 51404165), Natural Science Foundation of Shanxi (Project No. 2015021088), Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (Grant No. 2015122) and the Special/Youth Foundation of Taiyuan University of Technology (Grant No. 2014QN029).