IEICE Proceeding Series

Multi-bit sampling from chaotic time series in random number generation

Kenichi Arai, Takahisa Harayama, Peter Davis, Jun Muramatsu, Satoshi Sunada

Vol. 1 pp. 268-271
Publication Date: 2014/03/17
Online ISSN: 2188-5079

2012 International Symposium on Nonlinear Theory and its Applications
NOLTA2012, Palma, Majorca, Spain, October 22-26, 2012

Multi-bit sampling from chaotic time series in random number generation

Kenichi Arai[1], Takahisa Harayama[1,2], Peter Davis [3], Jun Muramatsu[1], and Satoshi Sunada[4]

[1] Communication Science Laboratories, Nippon Telegraph and Telephone Corporation
[2] School of Sciences and Engineering, Toyo University
[3] Telecognix Corporation
[4] College of Sciences and Engineering, Kanazawa University

Abstract—We discuss the methods to increase random number generation rates base on the theory which says that nondeterministic macroscopic states can be extracted from the amplification of intrinsic microscopic noise by a chaotic system. The theory guarantees to obtain nondeterministic random number sequences after microscopic noise distributions converge to an invariant density. Here, we focus on the convergence time needed by the methods such as the multi-bit samplings and the post processing. In addition, we consider the measurement noise which affects extracted sequences especially using the multi-bit samplings because the multi-bit samplings need more precise measurements. Taking account of the advantages and the disadvantages, we show that, in some cases, the multi-bit samplings and the post processing are efficient to generate random number sequences at high rates.

1. Introduction

Random sequence generators are key technologies for ciphers and numerical simulations. Various methods to generate random sequences have been proposed. Unpredictable random sequences are strongly required especially for high security applications. There is the possibility that physical devices can meet the expectations for nondeterministic random sequence generations. Indeed, methods to generate random sequences using chaotic semiconductor lasers have recently been proposed [1, 4, 3, 5, 8], which can also achieve very high generation rates.

There is competition aiming at high generation rate of random sequences using chaotic semiconductor lasers. Some methods employ multi-bit samplings and post processing [5, 8]. On the other hand, the quality of random sequences are important for some applications. Harayama et al. provided a theoretical description of how the chaotic dynamics in lasers with optical delayed feedback transduce microscopic quantum noise of spontaneous emission into random transitions between macroscopic states [4]. Recently Sunada et al. provided experimental evidence supporting the theory [9].

We would like to discuss multi-bit samplings from the viewpoints of the above theory. It would appear that there is a trade off between the number of bits per sample and the rate of sampling. A longer interval is thought to be needed for unpredictable more bits because the convergence to an invariant distribution is required in more precise resolutions. We examine the efficiency of the multi-bit sampling by numerical experiments. In addition, we would like to point out the effect of the measurements noise on obtained sequences, which should be considered especially for multi-bit sampling methods.

2. Theory

2.1. Amplification of microscopic noise

Let us suppose that we observe a variable $x(t)$ generated by a chaotic dynamics perturbed by the microscopic noise. The theory for random number generation is based on the property of the chaotic dynamics which amplifies intrinsic nondeterministic microscopic noise such as quantum noise. The strong chaos property implies that any smooth initial probability density of $x(t)$ converges to the invariant density $\rho(x)$ corresponding to the natural invariant density of this chaos system. We emphasize that the asymptotic invariant measure does not depend on the initial noise density. In principle the nondeterminism of the microscopic noise is the origin of the nondeterminism of $x(t)$, but the asymptotic invariant density of the large-amplitude $x(t)$ is a property of the chaotic dynamics. This convergence to the invariant density is a key fundamental point for the use of chaotic system to generate large-amplitude signals for robust nondeterministic random-bit generation.

2.2. Extraction of random sequence

Let us extract numbers from a chaotic time series by assigning the numbers $0, \cdots, N-1$ to extraction subdomains $S_0, \cdots, S_{N-1}$ of $x(t)$, where the disjoint subdomains S_i are defined by the invariant density $\rho(x)$ so that it satisfies

$$\int_{S_0} \rho(x)dx = \int_{S_1} \rho(x)dx = \cdots = \int_{S_{N-1}} \rho(x)dx. \qquad (1)$$

Then, we extract a number i when an observable $x(t)$ is found in a subdomain S_i. Each extraction subdomain S_i is not necessarily connected. The extraction subdomains can be constructed by dividing the domain of x into M ($> N$) small subdomains and uniting the small subdomains. This can be regarded as a so-called post processing based on the theory. We can optimize how to construct the extraction

subdomains. In addition, it is important to note that real systems cannot exactly achieve the above equality which assumes that the observation of $x(t)$ and the classifications of the subdomains are done with infinite precision. We will refer this problem later from viewpoints of the measurement noise.

In order to extract random numbers, we observe the time evolution of a variable $x(t)$ at discrete sampling times $t = 0, \tau, 2\tau, \ldots$. Let us suppose that a time series $x(t)$ is precisely x_j at $j\tau$, that is, $x(j\tau) = x_j$. However, just after the observation, the value of $x(t)$ is perturbed by intrinsic nondeterministic microscopic noise and has a probability density $\rho_0(x)$. Let $\rho_t(x)$ be a probability distribution of $x(j\tau + t)$, evolved from $\rho_0(x)$. If t is so long, $\rho_t(x)$ is practically the same as the invariant density and numbers depend only on the invariant density and the extraction subdomains. In this case, we obtain a number i with an even probability, in other words, all numbers appear unbiasedly and the successive number is independent due to the nondeterministic property of microscopic noise. We can determine the memory time τ_m with an appropriate criteria in terms of the distance of distributions such that we can not distinguish between the invariant density and the distribution $\rho_t(x)$ for $t > \tau_m$.

2.3. Criterion for memory time

We employ Kullback-Leibler divergence (KL divergence) to define the memory time. Let P_i be a probability of obtaining a number i from a chaotic time series and $Q_i = 1/N$ for all i, that is, all numbers can be obtained with equal probabilities. KL divergence between P_i and Q_i is described as

$$D(P\|Q) = \sum_i P_i \log \frac{P_i}{Q_i} = \log N - H_N(P), \qquad (2)$$

where $H_N(P)$ $(= -\sum_i P_i \log P_i)$ is the entropy of the distribution P. $D(P\|Q)$ can be regarded as a kind of a distance between distributions P and Q. For example, if P and Q are the same distributions, $D(P\|Q)$ is 0. All these things make it clear that the distribution P can not practically be distinguished from the uniform probability distribution if KL divergence $D(P\|Q)$ is small enough. Therefore, the memory time τ_m can be defined so that KL divergence $D(P\|Q)$ becomes smaller than a critical value D_m if $t > \tau_m$.

3. Noise in signal measurement system

In this section, we discuss the limitation of precisions of the variable $x(t)$ and thresholds Xi, which determine small subdomains. We consider the system to measure chaotic signals generated by chaotic semiconductor lasers. Figure 1 shows the measurement system with a photo detector and a digital oscilloscope. We investigate the intrinsic noise effects insides a high speed 4-channel oscilloscope (Tektronix DSA71254), which was used to generate fast random-bit using on-chip chaos laser in [4]. The

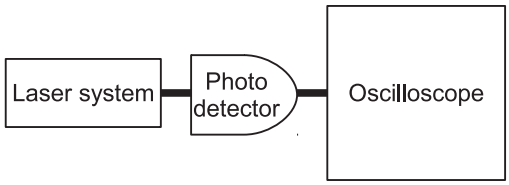

Figure 1: Measurement system

intrinsic noise effects can be examined by setting the common input signal to ground. Figure 2 shows an example of the distribution of samples over quantization levels, in the case of GND coupling and 100 mV full scale. The

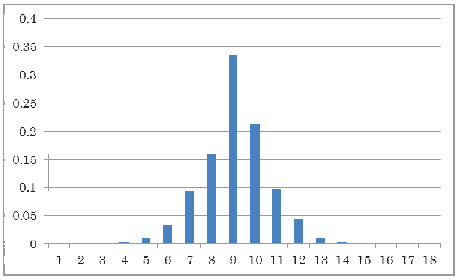

Figure 2: Code distribution 18 of the 256 bins are shown. Bin 9 corresponds to DC ground level zero volts.

so-called noise free code resolution (NFCR) is the number of bits which are not affected by the intrinsic noise of the ADC. In the above example, the distribution is spread over a wide range of 16 bins, so the truly noise-free bits in the data are just (8 − 4=) 4 bits. However, large fluctuations have low frequencies. The standard deviation corresponds to $(\pm \log_2(1.53) = \pm 0.61$ bits) 1.2 bits, giving NFCR (8 − 1.2 =) 6.8 bits.

4. Memory time for chaotic map system

4.1. Tent map

We employ the following asymmetric tent map as a simple chaotic system to investigate the memory times.

$$x_{n+1} = \begin{cases} \frac{x_n}{a} & x_n < a \\ \frac{x_n - 1}{a - 1} & a \le x_n \end{cases} \qquad (3)$$

Note that the invariant density of the map is a uniform distribution for any a value. The time evolutions of KL divergence $D(P\|Q)$'s were experimentally obtained as follows. First we prepared many narrow initial distributions, centers of which are chosen randomly from the set of all possible initial conditions. The invariant distribution was used as the probability distribution for the centers of initial distributions. The evolutions of each distribution determine that of KL divergence at every time step. At every time

step, we averaged $D(P||Q)$'s over all prepared initial distributions. Suppose that the critical value D_m is 0.01. In

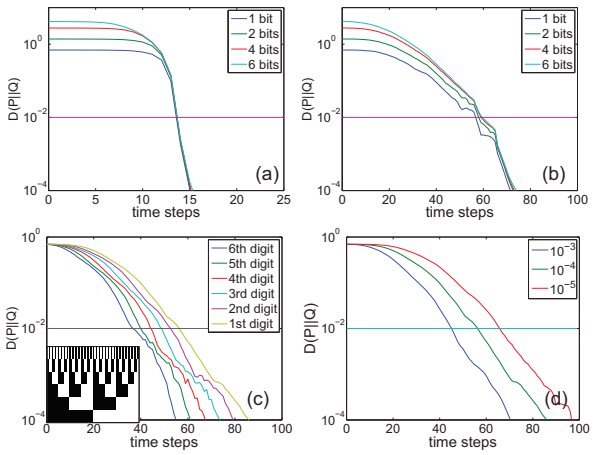

Figure 3: Time evaluation of $D(P||Q)$'s for tent maps. $D(P||Q)$'s of some number of bits are shown in (a) $a = 0.50$ and (b) $a = 0.90$. $D(P||Q)$'s for some extraction subdomain assignments are shown in (c). $D(P||Q)$'s for some initial noise amplitudes is shown in (d).

Figs. 3, we used 10000 initial distributions and 100000 initial points within each initial distribution. Figures 3 (a) and (b) show the time evolutions of $D(P||Q)$'s for n bit(s) samplings when the domain of x is divided into N $(= 2^n)$ extraction subdomains by equally-spaced thresholds. In the case of $a = 0.5$, $D(P||Q)$'s become smaller than D_m almost at the same time for all multi-bit samplings, that is to say, memory time is independent of the number of bits obtained in one sampling. On the other hand, for $a = 0.90$, the time evolutions of $D(P||Q)$'s make a difference and the memory times are slightly different. As a approaches 1, the difference of the memory times tends to become larger. When $a = 0.5$, the absolute value of the slop of the map is 2 everywhere. As a approaches 1, the slope of the right part of map becomes steeper and the slope of the left part of map becomes close to 1. The speeds of convergence depend on the position of initial distributions. This could make the difference of memory times among the number of bits. As far as this experiments, since the difference is not so large even for large a values, the multi-bit sampling is efficient to increase the generation rates of random numbers.

There are a variety of possible assignments of small subdomains to construct the same 2 bits sampling. Whole domain $[0, 1]$ is divided into 64 equi-spaced small subdomains $[k/64, (k + 1)/64)$, $k = 0, 1, \cdots, 63$, which are numbered by the binary number representation, e.g., 011010 for $k = 26$. If i-th significant digit is 0 (1), then 0 (1) is assigned to the subdomains. In inset of Fig. 3 (c), six patterns of assignments are shown; black (white) regions indicate a bit of 0 (1). This process can be regarded as the so-called post processing. We can see in Fig. 3 (c) that the assignment using the 6th digit has the smallest memory time. There are many other assignment patterns and how to find the smallest memory time is an optimization problem. To argue this point is out of our purpose.

We consider the effect of initial noise intensities, or the width of initial distributions, on the memory times. Figure 3 (d) shows the time evolutions of $D(P||Q)$'s for some initial noise intensities. We can see that the memory times are smaller as initial noise intensities are larger. The functional forms shift to right as initial noise intensities become larger, as reported by Mikami et al.[7]. At the early stage, narrow initial distributions are simply stretched by several iterations of this simple map. As a result, in the case of very small noise, the multi-bit sampling is efficient.

4.2. Modified Bernoulli map

Next, we consider the following modified Bernoulli map.

$$x_{n+1} = \begin{cases} x_n + 2^{B-1}(1-2b)x_n^B + b, & x_n < 1/2 \\ x_n + 2^{B-1}(1-2b)(1-x_n^B) + b, & 1/2 \le x_n \end{cases} \quad (4)$$

By solving Frobenius-Perron equation approximately, the map has the invariant measure as follows [2],

$$\rho(x) \sim x^{1-B} + (1-x)^{1-B}. \quad (5)$$

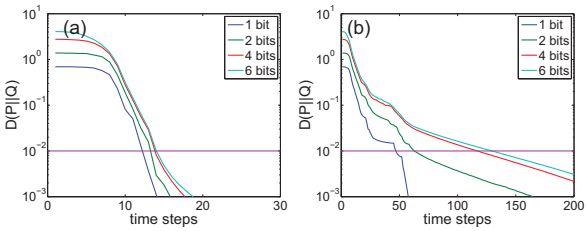

Figure 4: Time evaluation of $D(P||Q)$'s for modified Bernoulli map with $B = 1.3$ (a), and $B = 1.8$ (b).

Figure 4 shows the time evolutions of KL divergence $D(P||Q)$'s for some B's and $b = 0$. Although the differences of the memory times of the modified Bernoulli map are larger than the tent maps, the increased number of bits which can be obtained at one sampling more than make up for the extension of the memory times for $2 \sim 6$ bits in the case of $B = 1.3$, and for $4 \sim 6$ bits in the case of $B = 1.8$.

5. Laser with delayed optical feedback

We consider the random number generation from chaotic semiconductor lasers. In a single mode semiconductor laser with delayed optical feedback, the dynamics of the macroscopic variables of light field amplitude E and the carrier density N is described by the Lang-Kobayashi equations [6] as

$$\frac{dE}{dt} = \frac{1+i\alpha}{2}\left(G - \frac{1}{\tau_p}\right)E + \frac{\kappa}{\tau_{in}}E(t-\tau_D)e^{-i\theta} + \sqrt{\frac{C_s N}{\tau_s}}\xi,$$

$$\frac{dN}{dt} = J - \frac{1}{\tau_s}N - G|E|^2. \quad (6)$$

The following parameter values are employed; $\alpha = 5$, $G_0 = 10^{-12}\mathrm{m}^3\mathrm{s}^{-1}$, $\epsilon = 8.16 \times 10^{24}\mathrm{m}^3$, $\tau_{\mathrm{in}} = 14\mathrm{ps}$, $\tau_D = 0.182\mathrm{ns}$, $\theta = 0\mathrm{rad}$, $\tau_s = 2.04\mathrm{ns}$, $\kappa = 0.32$, $N_0 = 1.4 \times 10^{24}\mathrm{m}^{-3}$, and $C_s = 10^{-3}$. Figure 5 shows an invariant density and the time evolutions of KL divergence $D(P||Q)$'s. The invariant

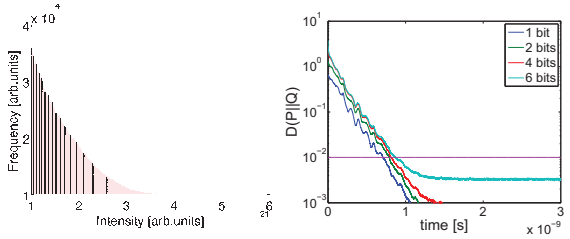

Figure 5: (left) Invariant intensity distribution of Lang-Kobayashi model and extraction thresholds for 16 divisions (4 bits). (right) Time evaluation of $D(P||Q)$'s for Lang-Kobayashi model.

density is obtained from a long term time series. Extraction thresholds determined from equal measure partitions of the invariant distribution of intensity are also shown in the right figure of Fig. 5. We can see that the memory times become longer as the number of bits sampled at one time is larger. However, the proportional increase of memory time is less than that of the increase of the number of bits.

6. Summary

We discuss the random number sequence generation using multi-bit sampling from chaotic time series from the viewpoint of the theory based on the property of a chaotic system and the intrinsic microscopic noise. In a strong chaotic system, any smooth density of the intrinsic microscopic noise converges to the invariant density. As the number of bits obtained at every sampling becomes larger, the interval of sampling should be longer, as measured by the KL divergence, because the more delicate convergence to invariant density should be needed. However, the examples that we studied numerically with up to 6-bit sampling, showed that it is possible for the memory time per bit to decrease. This suggests that larger rates of bit generation can be obtained with multi-bit sampling.

Another important point is the measurement noise which affects extracted bit sequences especially with multi-bit sampling. Our investigation shows that the oscilloscope (Tektronix DSA71254) with 8 bits resolution ADC have internal noise which corresponds to the 2 least significant bits.

For multi bit sampling, we should carefully consider the interval between samplings and the fluctuations of measurement systems. Our numerical experiments showed that multi-bit sampling can be an effective way to obtain random number sequences if we take account of the trade-off between the number of bits and sampling frequency.

References

[1]

[2] Y. Aizawa, Y. Kikuchi, T. Harayama, K. Yamamoto, M. Ota, and K. Tanaka. Stagnant motions in hamiltonian systems. *Prog. Theor. Phys. Suppl*, 98:36–82, 1989.

[3] 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. *Optics Express*, 18(18):18763–18768, 2010.

[4] Takahisa Harayama, Satoshi Sunada, Kazuyuki Yoshimura, Peter Davis, Ken Tsuzuki, and Atsushi Uchida. Fast nondeterministic random-bit generator using on-chip chaos lasers. *Physical Review A*, 83:031803(R), 2011.

[5] I. Kanter, Y. Aviad, I. Reidler, E. Cohen, and M. Rosenbluh. An optical ultrafast random bit generator. *Nature Photonics*, 4(1):58–61, 2009.

[6] R. Lang and K. Kobayashi. External optical feedback effects on semiconductor injection laser properties. *Quantum Electronics, IEEE Journal of*, 16(3):347–355, 1980.

[7] T. Mikami, K. Kanno, K. Aoyama, A. Uchida, T. Ikeguchi, T. Harayama, S. Sunada, K. Arai, K. Yoshimura, and P. Davis. Estimation of entropy rate in a fast physical random-bit generator using a chaotic semiconductor laser with intrinsic noise. *Physical Review E*, 85(1):016211, 2012.

[8] I. Reidler, Y. Aviad, M. Rosenbluh, and I. Kanter. Ultrahigh-speed random number generation based on a chaotic semiconductor laser. *Physical review letters*, 103(2):24102, 2009.

[9] S. Sunada T., Harayama, P. Davis, K. Arai, K. Yoshimura, K. Tsuzuki, and A. Uchida. Noise amplification in high dimensional chaotic laser systems and its application to nondeterministic physical random bit generation. *Chaos*. Submitted.