POLITECNICO DI TORINO
Repository ISTITUZIONALE

From Chaos to Pseudorandomness: A Case Study on the 2-D Coupled Map Lattice

Original
From Chaos to Pseudorandomness: A Case Study on the 2-D Coupled Map Lattice / Wang, Yong; Liu, Zhuo; Zhang, Leo Yu; Pareschi, Fabio; Setti, Gianluca; Chen, Guanrong. - In: IEEE TRANSACTIONS ON CYBERNETICS. - ISSN 2168-2267. - STAMPA. - 53:2(2023), pp. 1324-1334. [10.1109/TCYB.2021.3129808]

Availability:
This version is available at: 11583/2975908 since: 2023-02-10T13:28:20Z

Publisher:
IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC

Published
DOI:10.1109/TCYB.2021.3129808

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

12 October 2024

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808

From Chaos to Pseudo-Randomness: A Case Study on the 2D Coupled Map Lattice

Yong Wang, Zhuo Liu, Leo Yu Zhang, *Member, IEEE,* Fabio Pareschi, *Senior Member, IEEE,* Gianluca Setti, *Fellow, IEEE* and Guanrong Chen, *Life Fellow, IEEE*

***Abstract*—Applying chaos theory for secure digital communications is promising and it is well acknowledged that in such applications the underlying chaotic systems should be carefully chosen. However, the requirements imposed on the chaotic systems are usually heuristic, without theoretic guarantee for the resultant communication scheme. Among all the primitives for secure communications, it is well-accepted that (pseudo) random numbers are most essential. Taking the well-studied two-dimensional coupled map lattice (2D CML) as an example, this paper performs a theoretical study towards pseudo-random number generation with the 2D CML. In so doing, an analytical expression of the Lyapunov exponent (LE) spectrum of the 2D CML is first derived. Using the LEs, one can configure system parameters to ensure the 2D CML only exhibits complex dynamic behavior, and then collect pseudo-random numbers from the system orbits. Moreover, based on the observation that least significant bit distributes more evenly in the (pseudo) random distribution, an extraction algorithm $\mathbf{E}$ is developed with the property that, when applied to the orbits of the 2D CML, it can squeeze uniform bits. In implementation, if fixed-point arithmetic is used in binary format with a precision of z bits after the radix point, $\mathbf{E}$ can ensure that the deviation of the squeezed bits is bounded by 2^{-z}. Further simulation results demonstrate that the new method not only guide the 2D CML model to exhibit complex dynamic behavior, but also generate uniformly distributed independent bits with good efficiency. In particular, the squeezed pseudo-random bits can pass both NIST 800-22 and TestU01 test suites in various settings. This study thereby provides a theoretical basis for effectively applying the 2D CML to secure communications.**

***Index Terms*—Chaos, Lyapunov Exponent, Random Number Generator, Secure Communication, 2D Coupled Map Lattice.**

I. Introduction

This work was supported by the Natural Science Foundation of Chongqing, China [No. cstc2021jcyj-msxmX0557], the Science and Technology Foundation Project of Guizhou Province(QianKeHeJiChu[2020]1Y422), the MOE Layout Foundation of Humanities and Social Sciences, China [No. 20YJAZH102], and the National Natural Science Foundation of China [No. 71901045].

Y. Wang and Z. Liu are with the College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, China. Meanwhile, Z. Liu is a Lecturer with the School of Mathematics and Big Data, Guizhou Education University, Guiyang, China. (email: wangyong1@cqupt.edu.cn; liuzhuo1987@outlook.com).

L. Zhang is with the School of Information Technology, Deakin University, Australia (email: leo.zhang@deakin.edu.au).

F. Pareschi and G. Setti are with the Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Torino, Italy, and also with the Advanced Research Center on Electronic Systems (ARCES), University of Bologna, 40125 Bologna, Italy (e-mail: fabio.pareschi@polito.it; gianluca.setti@polito.it).

G. Chen is with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong SAR (email: eegchen@cityu.edu.hk).

IN the past two decades, chaotic systems have been widely used for secure communications, owing to their unique qualities including complexity, pseudo-randomness and ergodicity [1], [2]. These traits are related to confusion and diffusion, which are desired characteristics for good cryptographic primitives, according to [3]. For this reason, many encryption algorithms and secure communication schemes based on chaotic systems have been proposed, creating a great deal of research across nonlinear dynamics and information security [4]–[7]. Essentially, chaotic systems' inherent features play a significant part in maintaining the schemes' security. As a result, selecting or constructing a chaotic system from the perspective of cryptographic applications has become a critical challenge for chaos-based secure communications.

There are two different options on how to choose chaotic systems for secure communications [8]. The first one is to choose simple chaotic systems since simple chaotic systems generally have fewer arithmetic operations, and thus less computational complexity. Secure communication schemes based on simple systems typically have better runtime efficiency [7], [9]. However, if the structure of a simple chaotic system is not complicated enough, an attacker might be able to predict its chaotic orbits [10], which are used for secure communication directly or indirectly [11], [12]. This probably brings some potential vulnerability to the secure communication schemes.

To alleviate this problem, on the one hand, one may consider enhancing the dynamic complexity of simple chaotic systems [13]–[20]. In this way, a better trade-off between security and efficiency can be obtained if the enhancing method is still efficient. On the other hand, one may use higher-dimensional chaotic systems to design better secure communication schemes [21], [22]. Compared with simple chaotic systems, such as the logistic and the tent maps, a higher-dimensional chaotic system typically has more complicated dynamic behavior. Generally, it is more difficult, if not impossible, to predict the sequences generated from higher-dimensional chaotic systems.

However, this gain in security is not free, since more arithmetic operations of a higher-dimensional chaotic system will consume more computational capacity and incur inferior running speed. To decrease the number of iterations of a higher-dimensional system and achieve higher efficiency, it is common to extract more pseudo-random bits from a single iteration of the underlying higher-dimensional chaotic system [13], [23], [24]. But this is heuristic and depends on the pseudo-randomness of the underlying chaotic system. Another method is to iterate the higher-dimensional chaotic system in

Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808

parallel to improve the efficiency. However, it only applies to higher-dimensional chaotic systems that support parallelization implementation.

Besides the consideration about the trade-off between security and efficiency, another critical issue is the effect of finite-precision representation of chaotic orbits [25]. Generally, chaotic orbits are real numbers, and real numbers are then truncated and represented in either floating-point or fixed-point arithmetic on digital computer. When expressed with a certain precision under the fixed-point arithmetic, all chaotic systems will inevitably degenerate to become periodic. Worse yet, if chaotic systems are not configured correctly, their orbits, even represented by real numbers, could fall into periodic with a short period. Without deliberate design, therefore, these problems will degrade the security of chaos-based encryption algorithms. Based on these observations, to design a fully-fledged chaotic secure communication system, it is imperative and indispensable to have some theoretic guidelines for ensuring chaotic orbits to run in full chaotic state and the pseudo-random bits extracted from digitized orbits are evenly distributed with a sufficiently long period. The present work tries to address this issue by taking the two-dimensional coupled map lattice (2D CML) for a case study.

The coupled map lattice [26] is a classic model of spatiotemporal chaos. It is a complex two-directional chaotic system coupled with multiple identical simple chaotic maps, and it has excellent scalability. All the nodes in the CML have the same structure, so their arithmetic operations can be computed individually, which makes it possible to run the CML in parallel by certain special design [27]. Moreover, since the CML consists of multiple nodes, it will degenerate to a periodic system only if all the nodes are in periodic state simultaneously. From this view, even under finite precision representation, the period of CML model is longer than that of a single node (a simple chaotic map) [28], which can easily make the orbital period to be long enough for practical applications.

CML's potential has been utilised in recent years to create numerous secure communication algorithms [24], [29]–[35] due to the qualities described above. For example, by using the CML as the core of diffusion operations, a cryptographic model is proposed to guarantee better security of information processes [29]. The work [24] proposes a novel CML-based pseudo-random number generator (PRNG) with strong potential for cryptographic applications. The sequences generated from CML are used for constructing nonlinear substitution boxes (S-boxes) [30], which is a basic building block for many encryption algorithms. In [31], the way of designing S-boxes based on CML is assembled to an image encryption algorithm to enhance the image privacy. CML can also be used with other technologies, such as DNA coding, to create more efficient encryption methods [32]. From a theoretical perspective, to further enhance the complexity of dynamic behavior, the one-dimensional CML is extended to the two-dimensional form [33]. Some hash functions and encryption algorithms are heuristically developed based on the complexity, diffusion and randomness of the 2D CML [34], [35].

When different values for the CML parameters are specified, the state of a CML system (either 1D or 2D) can exhibit diverse patterns, such as frozen random pattern, competition intermittency, and fully developed chaos, according to [26], [36]. To our knowledge, however, there is little theoretical study on the dynamics of 2D CML, despite the fact that the 2D CML has previously been heuristically used to secure communications [33]–[35]. Furthermore, as previously stated, heuristically extracting pseudo-random bits from chaotic orbits would leave a security flaw for the integrated system, which is also the case with the 2D CML.

To address these challenges, this paper studies the dynamics of 2D CML as well as pseudo-random bit generation from orbits of 2D CML in a theoretical perspective. This work makes the following contributions:

- For a commonly used 2D CML system, the analytic formula between the Lyapunov exponents (LEs) and the parameters of the model is derived, which provides theoretical guarantee for ensuring the 2D CML to run in fully developed chaotic state.
- Focusing on the fixed-point arithmetic, an extraction algorithm $\mathbf{E}$ is introduced to process the digitized orbits of the 2D CML, which ensures the pseudo-random bits extracted from the orbits to be uniform under certain mild assumption. This extraction algorithm, concerning the acquisition of uniform bits from non-uniform random sources, is also of research interest in its own right.
- Extensive experiments are performed to verify the theoretical results. In particular, it is demonstrated that the extracted bits from the orbits of 2D CML successfully pass both NIST 800-22 and TestU01 tests [37], [38] under different settings. Moreover, efficiency comparison validates that the proposed PRNG is faster than the only known theoretic sound chaotic PRNG [39], and also faster than some recently designed heuristic PRNGs [18], [24].

The rest of the paper is organized as follows. Some preliminary knowledge is given in Sec. II. In Sec. III, theoretical Lyapunov exponent analysis of the 2D CML is presented and the practical question of how to extract uniform random bit from the orbits of the 2D CML is addressed. In Sec. IV, some numerical tests are presented to verify the theoretical results. Finally, conclusion is drawn in Sec. V.

II. Preliminaries

To prepare for the technical development of this work, some preliminary knowledge is provided in this section.

A. Two CML Models

The first CML model, as depicted in Fig. 1, is proposed by Kaneko in [26] and is one of the classic platform for studying spatiotemporal chaos.

Definition 1. *[40] The general one-dimensional nearest-neighboring CML model is described by*

$$x_{n+1}^{u} = (1-\varepsilon)\mathrm{F}(x_n^u) + \frac{\varepsilon}{2}\left[\mathrm{F}(x_n^{u+1}) + \mathrm{F}(x_n^{u-1})\right],$$

where $n = 1, 2, ...,$ is the time index, $u = 1, 2, ..., R$, is the space index, x_n^u is the state value of the u-th node at time n,

Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808

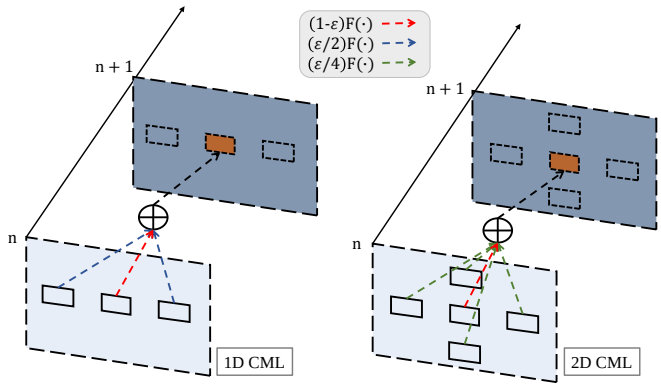


Fig. 1. The diagram of the 1D and 2D CML models.

$\varepsilon \in (0,1)$ is the coupling parameter, and F is a chaotic map. The periodic boundary condition is $x_n^0 = x_n^R$.

In the lattice, the coupling parameter ε between the nodes plays an important role in maintaining the complex dynamic behavior of the model. A tiny change in one node can affect other nodes and even be diffused to the whole lattice after some iterations. To further enhance the dyamic complexity, the one-dimensional CML model is extended to the two-dimensional version as follows.

Definition 2. *The two-dimensional CML model is defined by*

$$x_{n+1}^{u,v} = (1-\varepsilon)\mathrm{F}(x_n^{u,v}) + \frac{\varepsilon}{4}\left[\mathrm{F}(x_n^{u+1,v}) + \mathrm{F}(x_n^{u-1,v}) + \mathrm{F}(x_n^{u,v+1}) + \mathrm{F}(x_n^{u,v-1})\right], \quad (1)$$

where $u = 1,2,...,R$ and $v = 1,2,...,L$ are the row and column indexes of the nodes, respectively. The periodic boundary conditions are $x_n^{u+R,v} = x_n^{u,v}$ and $x_n^{u,v+L} = x_n^{u,v}$.

B. Lyapunov Exponent

Regarding nonlinear dynamic systems, LE is the key to measure chaotic behaviors. The maximum LE of the system $x_{n+1} = \mathrm{F}(x_n)$ is defined as [41]

$$\mathrm{LE} = \lim_{n\to\infty} \frac{1}{n} \ln \left| \prod_{m=0}^{n} \mathrm{F}'(x_m) \right|. \quad (2)$$

The value of LE is affected by the parameters in F and a positive LE indicates that the dynamic system is chaotic if its orbits are globally bounded. Moreover, the larger the LE is, the more complicated chaotic behavior the system has. For higher-dimensional systems or coupled systems like CML, it can have multiple LEs, and a strictly positive maximum LE indicates chaos.

For cryptographic confusion and diffusion suggested by Shannon [3], it is desirable to run a chaotic system with a large LE, so that any tiny change of the system parameters will spread out and be amplified gradually. For the requirement of pseudo-randomness [42], it is necessary to run the underlying chaotic system with a large LE so as to minimize the correlation between orbits produced by proximal parameters.

C. Non-uniform Randomness and Non-uniform Pseudo-Randomness

Let $x_{n+1} = \mathrm{F}(x_n)$ be an iterative chaotic system with value range $(0,1)$. Given the initial condition and system parameters, the orbits of this system are said to follow a pseudo-random distribution. Conceptually, by iterating F, the distribution function of $\{x_i\}_{i=0}^{\infty}$, $\mathrm{D}(\mathrm{F}, x_0)$, can be obtained.

Definition 3. $\mathrm{D}(\mathrm{F}, x_0)$ *is a pseudo-random distribution (PRD) if there exists a non-uniform true random distribution with the same density and there is no probabilistic polynomial-time (p.p.t.) algorithm that can distinguish them with a non-negligible probability.*

In the studies of true random number generation, it is common to employ certain entropy extractor to squeeze uniform true random numbers from non-uniform true random source [43]. While in the studies of pseudo-random number generation, there is no explicit tool that allows one to extract uniform pseudo-random bits (the functionality of PRNG) from a non-uniform pseudo-random source with an unknown distribution, and the introduction of PRD could be seen as one particular way to bridge this gap, as will be further discussed in Sec. III-B. In fact, it is easy to build a PRD by inversely sampling the output of a PRNG [44], but not vice versa. From this point of view, it is reasonable to believe that, like true random numbers obtained by squeezing non-uniform true random source, uniform pseudo-random numbers may be obtained by manipulating and compressing certain non-uniform PRD.

D. Independence Test

Definition 4. *Let x and y be two random variables that obey two random distributions, and x_i and y_i ($i \in [1,K]$) be K independent observations of x and y, respectively. The Pearson correlation coefficient k_{xy} between x and y is*

$$k_{xy} = \frac{\sum_{i=1}^{K} x_i y_i - K\bar{x}\bar{y}}{\sqrt{\left(\sum_{i=1}^{K} x_i^2 - K\bar{x}^2\right)\left(\sum_{i=1}^{K} y_i^2 - K\bar{y}^2\right)}},$$

where $\bar{x} = \frac{1}{K}\sum_{i=1}^{K} x_i$ and $\bar{y} = \frac{1}{K}\sum_{i=1}^{K} y_i$.

If x and y are independent of each other, the Fisher's transformation of k_{xy},

$$D = \frac{\sqrt{K-3}}{2} \ln \left| \frac{1+k_{xy}}{1-k_{xy}} \right|,$$

will approximately follow the standard normal distribution. By setting a significance level α, one can compare whether the empirical value D falls within the confidence interval $(\Phi^{-1}(\alpha/2), -\Phi^{-1}(\alpha/2))$, where Φ is the cumulative distribution function of the standard normal distribution. For a more reliable result, multiple tests should be applied and other tests, such as the chi-square test of independence and Kolmogorov-Smirnov test, should also be used. More importantly, one can apply this kind of tests between a PRD and a true random

Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808

distribution, provided that no p.p.t. algorithm can distinguish this PRD from a certain true random distribution. Also, as will be used later in Sec. III-B, the test can be applied to two PRDs if both of them cannot be distinguished from true random distributions.

III. From Chaos to Pseudo-Randomness

With the preliminary knowledge introduced above, a theoretic analysis of the 2D CML is presented in this section. Firstly, the analytical expression of all the LEs of the 2D CML is deduced in Sec. III-A, which will be used to enforce the system to run in fully developed chaotic state by using appropriate parameters. Secondly, it will discuss how to extract pseudo uniform bits with minimum bias from a pseudo-random distribution. Based on this result, an end-to-end extraction algorithm $\mathbf{E}$ will be designed to distill random bits from the orbits of 2D CML with theoretical support.

A. Lyapunov Exponent Formula of the 2D CML

As stated in Sec. I, for either designing chaotic ciphers or producing random numbers, one should carefully set the parameters to make the chaotic system to operate in a fully developed chaotic mode. As stated in Sec. II-B, this can be accomplished by carefully selecting the parameter(s) of the chaotic system. Here, since LE is defined asymptotically as a limit, the synchronization of node values will be used to derive an analytical expression of all LEs for the 2D CML.

To begin with, convert the 2D CML model with R rows and L columns to a one-dimensional model by rearranging the nodes according to the order from left to right and from top to bottom. Thus, (1) is converted to

$$\begin{aligned} x_{n+1}^{(u-1)L+v} =&(1-\varepsilon)\mathrm{F}(x_n^{(u-1)L+v}) + \frac{\varepsilon}{4}\Big[\mathrm{F}(x_n^{u\times L+v}) \\ &+ \mathrm{F}(x_n^{(u-2)L+v}) + \mathrm{F}(x_n^{(u-1)L+v+1}) \\ &+\mathrm{F}(x_n^{(u-1)L+v-1})\Big]. \end{aligned} \tag{3}$$

Correspondingly, the periodic boundary conditions will be changed to $x_n^{(u+R)L+v} = x_n^{u\times L+v}$ and $x_n^{(u-1)L+v+L} = x_n^{(u-1)L+v}$. All the node values in the converted model can be arranged as an $(R\times L)$-dimensional column vector, $\mathbf{z}_n = [x_n^1, x_n^2, \cdots, x_n^L, x_n^{L+1}, x_n^{L+2}, \cdots, x_n^{2L}, \cdots, x_n^{R\times L}]^T$. Similarly to the method used for the 1D CML [45], one can differentiate (3) and evaluate the derivatives along their synchronized trajectories. Upon synchronization, all the entries of $\mathbf{z}_n$ become equal, i.e., $x_n^1 = x_n^2 = \cdots = x_n^L = x_n^{L+1} = x_n^{L+2} = \cdots = x_n^{2L} = \cdots = x_n^{R\times L} = x_n$.

Along the synchronized trajectory, one has the derivatives of F as

$$\begin{aligned} &\mathrm{F}'(x_n^{(u-1)L+v}) = \mathrm{F}'(x_n^{u\times L+v}) = \mathrm{F}'(x_n^{(u-2)L+v}) \\ &= \mathrm{F}'(x_n^{(u-1)L+v+1}) = \mathrm{F}'(x_n^{(u-1)L+v-1}) = \mathrm{F}'(x_n), \end{aligned} \tag{4}$$

and the differentials of the 2D CML are

$$\begin{aligned} \delta(x_{n+1}^{(u-1)L+v}) =&(1-\varepsilon)\mathrm{F}'(x_n^{(u-1)L+v})\delta(x_n^{(u-1)L+v}) \\ &+ \frac{\varepsilon}{4}\Big[\mathrm{F}'(x_n^{u\times L+v})\delta(x_n^{u\times L+v}) \\ &+ \mathrm{F}'(x_n^{(u-2)L+v})\delta(x_n^{(u-2)L+v}) \\ &+ \mathrm{F}'(x_n^{(u-1)L+v+1})\delta(x_n^{(u-1)L+v+1}) \\ &+\mathrm{F}'(x_n^{(u-1)L+v-1})\delta(x_n^{(u-1)L+v-1})\Big]. \end{aligned} \tag{5}$$

By incorporating (4), (5) can be written as

$$\begin{aligned} \delta(x_{n+1}^{(u-1)L+v}) = \mathrm{F}'(x_n)\Big[&(1-\varepsilon)\delta(x_n^{(u-1)L+v}) \\ &+ \frac{\varepsilon}{4}(\delta(x_n^{u\times L+v}) + \delta(x_n^{(u-2)L+v}) \\ &+\delta(x_n^{(u-1)L+v+1}) + \delta(x_n^{(u-1)L+v-1}))\Big]. \end{aligned} \tag{6}$$

Applying (6) to all the $(R\times L)$ elements produced at time instants $(n+1)$ and n, i.e., $\mathbf{z}_{n+1}$ and $\mathbf{z}_n$, one can get a matrix form of (5), as,

$$\delta\mathbf{z}_{n+1} = \mathbf{J}_n\delta\mathbf{z}_n,$$

where $\mathbf{J}_n$ is the Jacobin matrix satisfying $\mathbf{J}_n = \mathrm{F}'(x_n)\mathbf{K}$ and $\mathbf{K}$ is an $(R\times L)\times(R\times L)$ circulant matrix in the form of

$$\mathbf{K} = \begin{bmatrix} \mathbf{A_1} & \mathbf{A_2} & \cdots & \mathbf{A_R} \\ \mathbf{A_R} & \mathbf{A_1} & \cdots & \mathbf{A_{R-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A_2} & \mathbf{A_3} & \cdots & \mathbf{A_1} \end{bmatrix}_{(R\times L)\times(R\times L)}.$$

Here, the matrices $\mathbf{A_1}, \mathbf{A_2}, \cdots, \mathbf{A_R}$ are given by

$$\mathbf{A_1} = \begin{bmatrix} 1-\varepsilon & \varepsilon/4 & 0 & \cdots & \varepsilon/4 \\ \varepsilon/4 & 1-\varepsilon & \varepsilon/4 & \ddots & 0 \\ 0 & \varepsilon/4 & 1-\varepsilon & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \varepsilon/4 \\ \varepsilon/4 & 0 & \cdots & \varepsilon/4 & 1-\varepsilon \end{bmatrix}_{L\times L},$$

$$\mathbf{A_2} = \mathbf{A_R} = \begin{bmatrix} \varepsilon/4 & & \\ & \ddots & \\ & & \varepsilon/4 \end{bmatrix}_{L\times L},$$

and

$$\mathbf{A_3} = \mathbf{A_4} = \cdots = \mathbf{A_{R-1}} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}_{L\times L}.$$

To calculate the LEs of the 2D CML, one needs to multiply all the Jacobin matrices. Let λ represent an eigenvalue of the matrix $\mathbf{K}$ and denote $\mathbf{G} = \mathbf{J}_1\times\mathbf{J}_2\times\cdots\times\mathbf{J}_n = \mathbf{K}^n\cdot\left(\prod_{m=1}^n \mathrm{F}'(x_m)\right)$. It is easy to verify that the eigenvalue of $\mathbf{G}$ is $\lambda^n\cdot\left(\prod_{m=1}^n \mathrm{F}'(x_m)\right)$. According to (2), the LEs of 2D CML are given by the following formula, parametrized by

Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808

λ:

$$\begin{aligned}\text{LEs} &= \lim_{n\to\infty}\frac{1}{n}\ln|\mathbf{J}_1\times\mathbf{J}_2\times\cdots\times\mathbf{J}_n| \\ &= \lim_{n\to\infty}\frac{1}{n}\ln\left|\lambda^n\prod_{m=1}^{n}\mathrm{F}'(x_m)\right| \\ &= \lim_{n\to\infty}\frac{1}{n}\ln\left|\prod_{m=1}^{n}\mathrm{F}'(x_m)\right| + \ln|\lambda|\,. \end{aligned} \tag{7}$$

Note that the first term in (7) is precisely the LE of the local chaotic map F, that is, $\lim_{n\to\infty}\frac{1}{n}\ln\left|\prod_{m=1}^{n}\mathrm{F}'(x_m)\right| = \mathrm{LE_F}$. So, to calculate the LE of the 2D CML, one has to compute the second term $\ln|\lambda|$. Here, λ represents any eigenvalue of $\mathbf{K}$ and $\mathbf{K}$ is a block circulant matrix with most of the blocks being zero (i.e., $\mathbf{A}_3 = \cdots = \mathbf{A}_{R-1} = \mathbf{0}$). Its characteristic polynomial is [46]

$$\prod_{r=0}^{R-1}\left|\mathbf{A}_1 + \mathbf{A}_2\omega_r + \mathbf{A}_R\omega_r^{R-1} - \lambda\mathbf{I}\right|, \tag{8}$$

where $\mathbf{I}$ is the identity matrix and

$$\omega_r = \exp\left(i\frac{2\pi r}{R}\right) = \cos\frac{2\pi r}{R} + i\sin\frac{2\pi r}{R},$$

for $r = 0, 1, \cdots, R-1$. Expanding the innner part of (8), one gets

$$\begin{aligned}\mathbf{A} &= \mathbf{A}_1 + \mathbf{A}_2\omega_r + \mathbf{A}_R\omega_r^{R-1} - \lambda\mathbf{I} \\ &= \begin{bmatrix} j & \varepsilon/4 & 0 & \cdots & \varepsilon/4 \\ \varepsilon/4 & j & \varepsilon/4 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon/4 & 0 & 0 & \cdots & j \end{bmatrix}_{L\times L},\end{aligned}$$

where $j = 1 - \varepsilon + \frac{\varepsilon}{4}\omega_k + \frac{\varepsilon}{4}\omega_k^{R-1} - \lambda$. It is clear that $\mathbf{A}$ is still a circulant matrix. Applying the same technique as used above [46], one obtains

$$\lambda = 1 - \varepsilon + \frac{\varepsilon}{4}\omega_r + \frac{\varepsilon}{4}\omega_r^{R-1} + \frac{\varepsilon}{4}\mu_l + \frac{\varepsilon}{4}\mu_l^{L-1}, \tag{9}$$

where $\mu_l = \exp\left(i\frac{2\pi l}{L}\right) = \cos\frac{2\pi l}{L} + i\sin\frac{2\pi l}{L}$, with $l = 0, 1, \cdots, L-1$.

Substitute (9) into (7) gives all the $(R\times L)$ LEs of 2D CML, i.e.,

$$\text{LEs} = \mathrm{LE_F} + \ln\left|1 - \varepsilon + \frac{\varepsilon}{2}\left(\cos\frac{2\pi r}{R} + \cos\frac{2\pi l}{L}\right)\right|, \tag{10}$$

for $r = 0, 1, \cdots, R-1$ and $l = 0, 1, \cdots, L-1$. From this analytical formula, it is easy to get the following results.

Theorem 1. *The maximum Lyapunov exponent of the 2D CML model is determined by the local chaotic map* F.

Proof. According to (10), the maximum LE of 2D CML is given as $\mathrm{LE} = \mathrm{LE_F}$, when $r = 0$ and $l = 0$. □

Corollary 1. *The maximum Lyapunov exponent is independent of the size of the 2D CML model.*

Corollary 2. *In 2D CML, increasing* $\mathrm{LE_F}$ *will increase the maximum Lyapunov exponent of the whole 2D lattice.*

B. Pseudo-Randomness Extraction From the 2D CML

According to the above analyses, with appropriate selection of the local chaotic map F, the maximum Lyapunov exponent of the 2D CML will be positive and the orbits of the whole system will only run in fully developed chaotic state.

Given the initial conditionals of the 2D CML and the system parameters of F, if any, the distribution of the orbits is governed by a PRD[1]. Without loss of generality, assume that this PRD is not uniform. This is because only few local chaotic maps are known to have a uniform density, for example the tent map [47]. And, even if a uniform local map is used, the overall density is likely to be uneven after coupling the local maps together in the structure of 2D CML. Similarly to sampling a random variable from a true random distribution, the way of taking a sample from this PRD is to iterate the 2D CML.

The following discussion shows how to squeeze uniform bits from the non-uniform PRD that governs the 2D CML's orbits. This discussion starts with a general result that holds for both pseudo-random distribution and true random distribution.

Theorem 2. *For a random (or pseudo-random) distribution in* $[0, 1]$*, assume that the density function has bounded first-order derivative. For any sample* $x = 0.w_1w_2\cdots w_z$ *(*$w_i \in \{0, 1\}$ *and* $i \in [1, z]$*) from this distribution, one has*

$$\lim_{z\to\infty} P(w_z = 0) = \lim_{z\to\infty} P(w_z = 1).$$

Proof. See the appendix. □

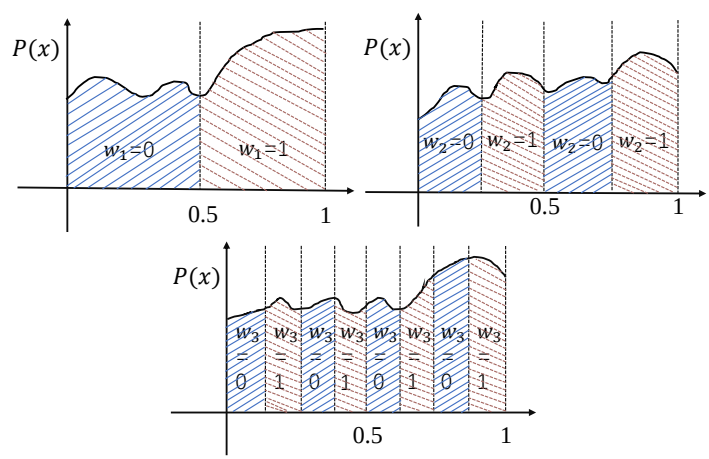


Fig. 2. Illustration of Theorem 2.

Remark: This theorem can be intuitively understood from the examples depicted in Fig. 2. For any (pseudo) random sample $x \in [0, 1]$, if there is only 1-bit ($z = 1$) used to represent this number, then the difference of the probabilities of $P(w_1 = 0)$ and $P(w_1 = 1)$ equals the difference of the areas associated with the two slices in blue and red (as shown in the top-left of Fig. 2). As z goes larger (for example, $z = 2$ and $z = 3$ in Fig. 2), the slices become smaller and their associated areas become similar. Thus, from the appendix, it is easy to see that $|P(w_z = 0) - P(w_z = 1)| = O(2^{-z})$.

[1]This kind of chaotic system is referred to as a 2D CML instance.

Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808

Remark: Although the theorem is proved by assuming that the density function has a bounded first-order derivative, which implies continuity, it can be easily extended to a density with countably many points of discontinuity, which is always the case of finite precision representations of real chaotic orbits.

Theorem 2 states the fact that, under fixed-point arithmetic, the farther away from the radix point, the more uniform the bits tend to be. For a given PRD determined by a 2D CML instance, a naive way of using Theorem 2 to extract (pseudo) random bits is to take the rightmost bit of a chaotic orbit under finite precision representation. If the precision is z bit, a single uniform bit with bias bounded by $O(2^{-z})$ can be obtained from one orbit. However, as discussed in Sec. I, the efficiency of this method is too low to meet practical requirements. Nevertheless, as will be seen in the following, this drawback can be avoided by making use of more CML instances and Property 1 given below.

For two different 2D CML instances, assume that their associated initial conditions and system parameters are selected independently. Consequently, the associated PRDs induced by these two instances are independent of each other, and so do the pseudo-random samples drawn from the two PRDs. For any sample $x = 0.x_1x_2\cdots x_z$ drawn from the first PRD and $y = 0.y_1y_2\cdots y_z$ drawn from the second PRD, by letting $w = x_i + y_j \pmod 2$, one can conclude that w is more uniform than both x_i and y_j ($i, j \in [1, z]$). Specifically, the bias of w is bounded by $O(2^{-(i+j)})$, as shown below.

Property 1. *Let x_i and y_j be independent binary random bits with bounded bias $O(2^{-i})$ and $O(2^{-j})$, respectively. Then, $w = x_i + y_j \pmod 2$ satisfies*

$$|P(w = 0) - P(w = 1)| = O(2^{-(i+j)}).$$

Proof. See the appendix. □

Extending this property further, the following corollary can be easily obtained. From this corollary, it is easy to see that, with more independent PRDs, one can obtain bits with arbitrarily small bias.

Corollary 3. *Let $w_1, w_2, \cdots, w_c$ be c mutually independent binary random variables, and each with bounded bias $O(2^{-k_i})$ ($i \in [1, c]$), and set $w = \sum_{i=1}^{c} w_i \pmod 2$. Then, one has*

$$|P(w = 0) - P(w = 1)| = O(2^{-\left(\sum_{i=1}^{c} k_i\right)}).$$

Property 1 holds only when the samples x and y are drawn from independent (pseudo) random distributions, which requires the two CML instances be selected independently. By making use of the result about LE in Sec. III-A, this assumption can be further relaxed. If the initial conditions or the system parameters of the CML instances are correlated but they both have large positive LE, then the correlation between their orbits will be dispersed fast after a few iterations. This transition period can be determined by applying the independence test from Sec. II-D to the orbits. In this way, the choice of the two 2D CML instances can be made arbitrary if both CMLs have large positive LEs and the orbits in transition are discarded.

Based on the discussion above, one can improve the efficiency of pseudo-random bits generation for z times by using two 2D CML instances, while keeping the bias be $O(2^{-z})$. The end-to-end extraction Algorithm 1 summarizes the details.

Algorithm 1: Extraction Algorithm **E**

Input: Two sets of initial conditions and system parameters 2D CML.
Output: Pseudo-random bits.
Function ModAdd (x, y) **:**
  $x = 0.x_1x_2\cdots x_z$
  $y = 0.y_1y_2\cdots y_z$
  $w = w_1w_2\cdots w_z$
  $w_i = x_i + y_{z+1-i} \pmod 2$
  return w
Function Main**:**
  Step 1 Run the two CML instances and collect their corresponding orbits $\{x^i\}_{i=0}^{K}$ and $\{y^i\}_{i=0}^{K}$, and update the states
  Step 2 Perform the independence test on the collected orbits
  if test passed **then**
    while necessary **do**
      Run the two instances to get (x^0, y^0)
      ModAdd(x^0, y^0)
    end
  else
    go back to **Step 1**
  end
  return 0

IV. Experimental Analysis

In this section, numeric tests are carried out on different 2D CML systems to assess the applicability and performance of the theoretic results developed in Sec. III. Some popular chaotic systems are used as the local map F of (1):

- The Logistic map

$$x_{n+1} = \mu x_n(1 - x_n),$$

where $\mu \in (0, 4]$ and $x_n \in (0, 1)$.

- The Tent map

$$x_{n+1} = \begin{cases} \mu x_n, & \text{if } x_n \in (0, 0.5), \\ \mu(1 - x_n), & \text{if } x_n \in [0.5, 1), \end{cases}$$

where $\mu \in (0, 2]$ is the system parameter.

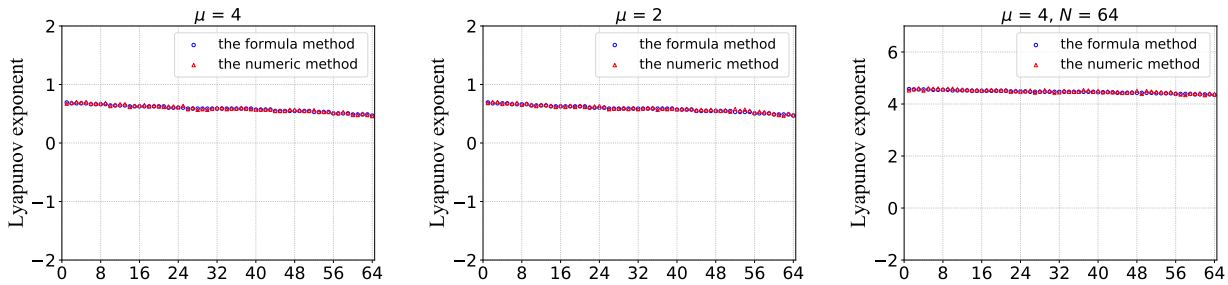
- The piecewise Logistic map (PLM)

$$x_{n+1} = \begin{cases} \mu N^2(x_n - \frac{i-1}{N})(\frac{i}{N} - x_n)\ , \\ \quad \text{if } \frac{i-1}{N} < x_n < \frac{i}{N}, \\ 1 - \mu N^2(x_n - \frac{i-1}{N})(\frac{i+1}{N} - x_n), \\ \quad \text{if } \frac{i}{N} < x_n < \frac{i+1}{N}, \end{cases}$$

where $\mu \in (0, 4]$ is the system parameter and N is the number of segments.

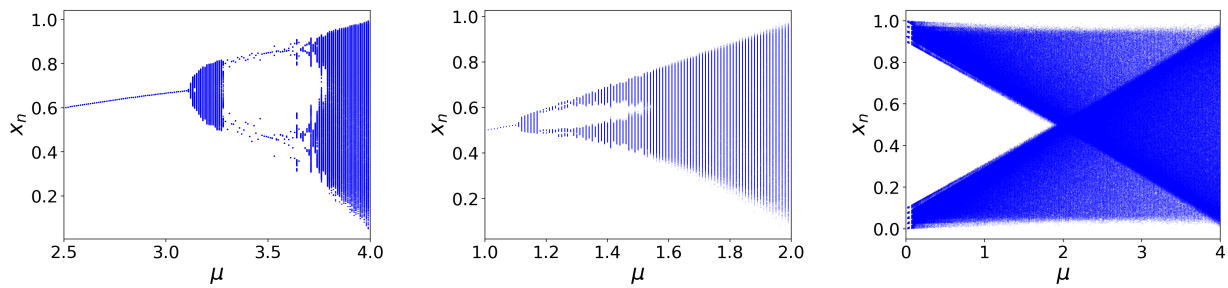
Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808



(a) LEs of L-CML (b) LEs of T-CML (c) LEs of PLM-CML

Fig. 3. LEs obtained from the numeric method in [41] and theoretic result in (10).



(a) (b) (c)

Fig. 4. Bifurcation diagram corresponding to the first node of L-CML, T-CML and PLM-CML, respectively.

It is worth mentioning that, other than the basic ones listed here, chaotic maps with more complex behavior, such as the ones suggested in [15], [16], can also be employed as the local map of the 2D CML system. For simplicity, refer the resultant chaotic systems to as L-CML, T-CML and PLM-CML, respectively.

A. Lyapunov Exponents and Bifurcation Diagrams

In this section, it is verified that the theoretic LE formula (10) of the 2D CML is correct. Then, the associated bifurcation diagram is used to demonstrate that the orbits of the CML instance indeed run in full chaotic state with appropriate parameters.

The reliable Wolf's scheme [41] is used to calculate the LEs of the 2D CML, and the result is plotted in Fig. 3 with $\mu = 4, 2, 4$ for L-CML, T-CML and PLM-CML, respectively. Here, the other parameters in (1) are set to $\varepsilon = 0.1$, $R = L = 8$, and the number of segments of PLM $N = 64$. It is clear that the theoretic scheme given by (10) perfectly aligns with the numeric method.

To demonstrate of chaos, the bifurcation diagram is used to plot output orbits of the system with respect to the change of the parameters. As shown in Theorem 1, the maximum LE of a 2D CML is solely determined by the local map $\mathrm{LE_F}$. And, the orbits extracted from the first node (i.e., $r = l = 0$ in (1)) achieves the maximum LE. By varying the parameter μ, one can exact the orbits of the first node of L-CML, T-CML, and PLM-CML, respectively, and their bifurcation diagrams are shown in Fig. 4. It is clear that the orbits of CML run in the chaotic state with an appropriate choice of the parameter of the local map. Once again, it should be pointed out that, with more complex chaotic systems, the $\mathrm{LE_F}$ can be made larger [15], [16], and the resulting chaotic performance will be better. This may bring up more benefits like a higher extraction rate of the pseudo-random numbers, but such possible improvement is left for future study. Moreover, other metrics, such as Kolmogorov entropy, fractal power spectra and correlation dimension, can be used to measure the performance of chaos, which likewise not the focus of this study.

B. Randomness Tests

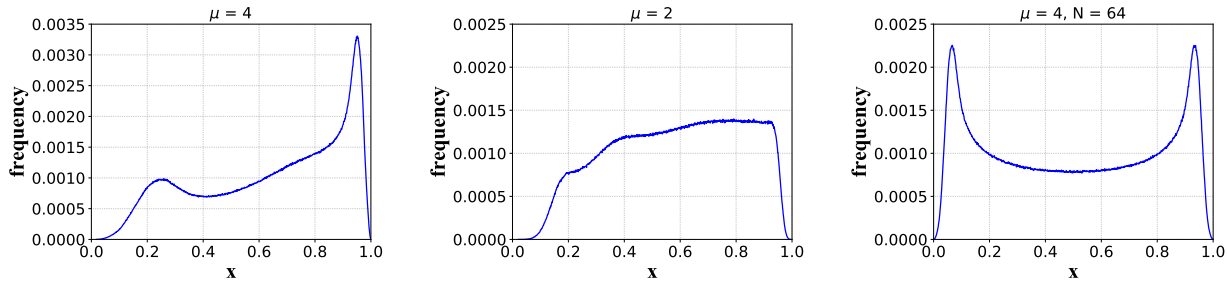
To assess the results presented in Sec. III-B, the pseudo-random distributions introduced by L-CML, T-CML and PLM-CML are first plotted. Under the same parameter settings, the resulting distributions are shown in Fig. 5.

Note that the distribution of the Tent map is known to be uniform when $\mu = 2$ [47], but the distribution of T-CML, like that of L-CML and PLM-CML, is not uniform. This finding supports the arguments in Sec. I, suggesting the need for a theoretically sound approach for squeezing uniform random numbers from chaotic orbits.

Next, statistical tests are performed on binary outputs of the extraction algorithm $\mathbf{E}$. The parameter settings are the same as above, while the precision z is set to 64 and the parameter K for the independence test is set to 10^3. With 3 local maps, there are $3 \times 3 = 9$ pairs of 2D CML instances as the input of $\mathbf{E}$. For each pair, one of the initial conditions of the CML instance is set to random and the other is obtained by perturbing the first one with difference up to 10^{-3}. Note that the initial conditions are strongly correlated, so this is the worst-case study for the extraction algorithm $\mathbf{E}$. But, as argued in Sec. III-B, positive LE, on the other hand, can disperse

Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808



(a) (b) (c)

Fig. 5. Orbits distribution of L-CML, T-CML and PLM-CML, respectively.

the correlation within initial conditions. The resultant orbits (and thus PRD) are pseudo-independent of each other after a transition period. The binary outputs are collected and tested against NIST SP800-22 [37] and TestU01 [38].

NIST SP800-22: This is a statistical randomness test suite for binary sequences. It is a collection of 15 tests, each of them outputs one (or, in a few cases, more than one) P-value. The suite is designed to be applied over binary sequences with length 10^6. The significance level α, which determines the region of acceptance of the assumption that the sequence being tested is random, is suggested to be set to 0.01. A sequence passes a test if P-value is not less than α.

Each test is designed such that the probability for an ideal generator to fail a test is α. So, $1000 > 1/\alpha$ sequences with length 10^6 produced by the extraction algorithm $\mathbf{E}$ are considered and tested. According to [37, Chap. 4], empirical test results with P-value $\geq \alpha$ should be examined further according to[2]: 1) the pass rate of the sub-tests; 2) the distribution of empirical P-values. The pass rate is approximated with a normal random variable, and under the current setting, a pass rate < 0.98056 is interpreted as evidence of non-randomness. The 1000 P-values from each sub-test will be further tested against a chi-square goodness-of-fit test with 10 bins, and this again produces another P-value$_T$ (i.e., a P-value of P-values). If P-value$_T \geq 0.0001$, the sequences are considered to be uniform.

TestU01: Generally speaking, it is a more systematic test suite for random numbers. TestU01 works for both random number over the interval $(0, 1)$ and binary sequences, and it contains six predefined batteries of tests. In this work, the batteries *Rabbit*, *Alphabit* and *BlockAlphabit* designed for binaries are selected. The batteries *Rabbit* and *Alphabit* have 38 and 17 statistical sub-tests, and *BlockAlphabit* applies the sub-tests in *Alphabit* repeatedly after reordering the bits by a block length specified by the software package. This test suite is run on binary sequences with length 2^{20}, 2^{25} and 2^{30}, and a P-value lies outside the range $[0.001, 0.999]$ is interpreted as fail; otherwise, it passes.

Taking two correlated L-CML instances as input, $\mathbf{E}$ produces $10^3 \times 10^6$ bits and these bits are tested against NIST 800-22. Table I shows the results of each sub-test of NIST 800-22, where the two-level testing procedure discussed above is applied. Note that the P-value displayed here is just a typical test output. It is obvious that the output of the extraction algorithm $\mathbf{E}$ passes all the sub-tests contained in the NIST 800-22 test suite. Similarly, taking two correlated T-CML instances as input, the TestU01 test suite is run on the outputs and the results are summarized in Table II. Clearly, the output of the extraction algorithm $\mathbf{E}$ also passes the TestU01 test. In addition, all additional 8 pairs of 2D CML instance combination of $\mathbf{E}$ pass both NIST 800-22 and TestU01 tests, but the results are omitted here owing to space constraints. The experimental results confirm that the developed theory provides a sound foundation for generating pseudo-random bits from digital chaos and the extraction algorithm $\mathbf{E}$ can be used to produce uniform bits for secure applications.

[2]This is called the two-level random number testing procedure [48].

TABLE I
NIST 800-22 TEST RESULTS ON $\mathbf{E}$ WITH TWO CORRELATED L-CML INSTANCES AS INPUT.

Sub-test	P-value†	Pass Rate	P-value$_T$
	≥ 0.01	≥ 0.98056	≥ 0.0001
Frequency	0.8902	0.9890	0.8677
Block Freq.	0.7511	0.9890	0.9179
Cumulative*	0.3775	0.9860	0.1835
Runs	0.1881	0.9900	0.0242
Longest Run	0.9101	0.9930	0.6931
Rank	0.5842	0.9900	0.1381
Spectral (FFT)	0.6464	0.9890	0.5341
Nonoverlap*	0.4994	0.9895	0.4702
Universal	0.8006	0.9820	0.0669
App. Entropy	0.4894	0.9830	0.4245
Random E.*	0.3198	0.9893	0.5463
Random E. V.*	0.3176	0.9915	0.4760
Serial*	0.2898	0.9880	0.3429
Linear Comp.	0.4376	0.9870	0.1581
Success Counts	15/15	15/15	15/15

* These tests generates more than one P-values; † Only one of the test results has been considered in the table.

C. Efficiency Analysis

To further assess the performance of the proposed extraction algorithm $\mathbf{E}$, it is evaluated through comparing with other chaotic PRNGs [16], [18], [24], [39]. It is noted that the design in [39] is the only previously known method that provides

Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808

TABLE II
TESTU01 TEST RESULTS ON **E** WITH TWO CORRELATED T-CML INSTANCES AS INPUT.

Length	*Rabbit*	*Alphabit*	*BlockAlphabit*
2^{30}	38/38	17/17	17/17
2^{25}	38/38	17/17	17/17
2^{20}	38/38	17/17	17/17

theoretically guaranteed uniform randomness from chaotic systems. By enhancing 1D chaotic systems, the work in [16] is a famous heuristic PRNG design due to its simplicity and thorough experimental evaluation. The methods in [18] and [24] are more recent heuristic proposals based on enhancing simple chaotic systems and using CML, respectively. For the proposed method, L-CML is used. For PRNGs in [16], [18], [24], [39], the same settings of the original works are used. All the algorithms are then implemented on a Laptop with the Core i7-10710U CPU and 16G RAM.

Table III lists the running time (averaged from $1,000$ tests) for generating 1 MByte binary stream from all these methods. With a running time of 40 *ms*, the proposed method is more efficient than the only other theoretical sound RPNG [39], and is also more efficient than the heuristic designs [18] (with running time 2996 *ms*) and [24] (with running time 58 *ms*), but inferior to the method in [16] (with running time 16 *ms*). However, looking further at the third and fourth column of Table III, it is clear that the proposed method provides theoretical randomness guarantee while the method in [16] does not.

TABLE III
RUNNING TIME COMPARISON.

Methods	Running Time (1MByte)	Theoretical Analysis	Experimental Analysis
Ours	40 *ms*	Yes	Yes
[16]	16 *ms*	No	Yes
[18]	2996 *ms*	No	Yes
[24]	58 *ms*	No	Yes
[39]	47 *ms*	Yes	Yes

To investigate the reason of the above experimental results, the amount of the basic operations for producing 8 pseudo-random bits is used as the metric to evaluate the complexity of the considered PRNGs. Referring the arithmetic details of the works in [16], [18], [24], [39], the basic operations are counted and the result is listed in Table IV. According to this table, the average number of the basic operations for the proposed method is 21, which is smaller than that of [18], [24], [39] but bigger than 13.33 of [16]. To summarize, the proposed method outperforms most of chaotic PRNGs [16], [18], [24], [39] in terms of efficiency. And the only method [16] that is more efficient than the proposed method does not provide guaranteed uniform randomness. From this sense, in real applications, one may choose to use the method in [16] when higher efficiency is needed, while opt to use the proposed method when higher security is desirable.

TABLE IV
NUMBER OF BASIC OPERATIONS FOR GENERATING 8 BITS.

No. of Operations	Ours	[16]	[18]	[24]	[39]
No. of Exclusive OR	8	0	0	0	0
No. of Interception	0	2/3	0	1	0
No. of Modulo	0	8	8	0	0
No. of Compare	0	1/3	4	0	32
No. of Inversion	8	0	0	0	0
No. of Addition/Subtraction	9/4	8/3	248	14	8
No. of Multiplication/Division	11/4	5/3	1368	26	0
No. of Real $\rightarrow$ Char	0	0	0	1	0
Total	21	13.33	1628	42	40

V. CONCLUSION

Coupling chaotic maps is a popular method for generating more complicated dynamic behavior, using for example 1D and 2D coupled map lattices in secure communications. This work presents the first theoretic study of the LEs of the 2D CML model and finds that the maximum LE is solely determined by the local map used for coupling. Moreover, by bridging true randomness and pseudo-randomness, it lays the theoretic foundation for deriving uniform pseudo-randomness from digitized chaotic orbits. Making use of this result, a random number extraction algorithm **E** is designed, which produces pseudo-random bits with bias bounded by $O(2^{-z})$, where z is the bit number of the precision. Extensive experiments are carried out and the results align perfectly with the theoretic formulas. Moreover, it is validated that the proposed algorithm possesses good efficiency. The theory may provide fresh insights into how chaos can be used to create pseudo-randomness. Future work will include more performance gains through parallel implementation and theoretical analysis of pseudo-random distributions produced by chaos.

APPENDIX
PROOF OF THEOREM 2 AND PROPERTY 2

Proof. Let $P(x)$ be the density function of $x \in [0,1]$, which bounds the first derivative $P'(x)$. Considering that x is represented with z bits, one has $P(w_z = 0) + P(w_z = 1) = 1$ and

$$P(w_z = 0) = \sum_{b=1}^{2^{z-1}} \int_{\frac{2(b-1)}{2^z}}^{\frac{2b-1}{2^z}} P(x)dx, \tag{11}$$

$$P(w_z = 1) = \sum_{b=1}^{2^{z-1}} \int_{\frac{2b-1}{2^z}}^{\frac{2b}{2^z}} P(x)dx. \tag{12}$$

Applying the mean-value theorem to (11) and (12), one gets

$$\begin{aligned} P(w_z = 0) &= \sum_{b=1}^{2^{z-1}} \int_{\frac{2(b-1)}{2^z}}^{\frac{2b-1}{2^z}} P(x)dx \\ &= \int_0^{\frac{1}{2^z}} P(x)dx + \int_{\frac{2}{2^z}}^{\frac{3}{2^z}} P(x)dx + \cdots \\ &\quad + \int_{\frac{2^z-2}{2^z}}^{\frac{2^z-1}{2^z}} P(x)dx \\ &= \frac{1}{2^z}\left[P(x_1) + P(x_2) + \cdots + P(x_{2^{z-1}})\right], \end{aligned}$$

Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808

where $\frac{2(i-1)}{2^z} \le x_i \le \frac{2i-1}{2^z}$ for $i \in [1, 2^{z-1}]$, and similarly,

$$\begin{aligned} P(w_z = 1) &= \sum_{b=1}^{2^{z-1}} \int_{\frac{2b-1}{2^z}}^{\frac{2b}{2^z}} P(x)dx \\ &= \frac{1}{2^z}\left[P(x'_1) + P(x'_2) + \cdots + P(x'_{2^{z-1}})\right], \end{aligned}$$

where $\frac{2i-1}{2^z} \le x'_i \le \frac{2i}{2^z}$ for $i \in [1, 2^{z-1}]$. Finally, one has

$$\begin{aligned} &\lim_{z\to\infty} |P(w_z = 1) - P(w_z = 0)| \\ &\le \lim_{z\to\infty} \frac{1}{2^z}\left(\sum_{i=1}^{2^{z-1}} |P(x_i) - P(x'_i)|\right) \\ &\le \lim_{z\to\infty} \frac{1}{2^z} \cdot \left(\sum_{i=1}^{2^{z-1}} |P'(\bar{x}_i)| \cdot \frac{2}{2^z}\right) \qquad (13) \\ &\le \lim_{z\to\infty} \frac{1}{2^z} \cdot \left(2^{z-1} \max |P'(\bar{x}_i)|\right) \cdot \frac{2}{2^z} \\ &\le \lim_{z\to\infty} \frac{1}{2^z} \cdot (\max |P'(\bar{x}_i)|) \\ &= 0, \end{aligned}$$

where $\bar{x}_i \in (x_i, x'_i)$ and (13) is derived based on the mean value theorem. □

Proof. By assumption, one has

$$\begin{aligned} |P(x_i = 0) - P(x_i = 1)| &= O(2^{-i}), \\ |P(y_j = 0) - P(y_j = 1)| &= O(2^{-j}). \end{aligned}$$

Since $w = x_i + y_j \pmod 2$ and x_i and y_j are independent of each other, one gets

$$\begin{aligned} P(w = 0) &= P(x_i = 0)P(y_j = 0) + P(x_i = 1)P(y_j = 1), \\ P(w = 1) &= P(x_i = 0)P(y_j = 1) + P(x_i = 1)P(y_j = 0). \end{aligned}$$

The bias of w can then be calculated as

$$\begin{aligned} &|P(w = 0) - P(w = 1)| \\ &= |P(x_i = 0)P(y_j = 0) + P(x_i = 1)P(y_j = 1) \\ &\quad - P(x_i = 0)P(y_j = 1) - P(x_i = 1)P(y_j = 0)| \\ &= |[P(x_i = 0) - P(x_i = 1)] \cdot [P(y_j = 0) - P(y_j = 1)]| \\ &= O(2^{-(i+j)}). \end{aligned}$$

Hence, the property is true. □

References

[1] G. Jakimoski and L. Kocarev, "Chaos and cryptography: Block encryption ciphers based on chaotic maps," *IEEE Trans. Circuits Syst. I-Fundam. Theor. Appl.*, vol. 48, no. 2, pp. 163–169, Feb. 2001.

[2] T. Yang, C. W. Wu, and L. O. Chua, "Cryptography based on chaotic systems," *IEEE Trans. Circuits Syst. I-Fundam. Theor. Appl.*, vol. 44, no. 5, pp. 469–472, May 1997.

[3] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Tech. J.*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

[4] L. Kocarev, "Chaos-based cryptography: A brief overview," *IEEE Circuits Syst. Mag.*, vol. 1, no. 3, pp. 6–21, 2001.

[5] P. Bergamo, P. D'Arco, A. De Santis, and L. Kocarev, "Security of public-key cryptosystems based on chebyshev polynomials," *IEEE Trans. Circuits Syst. I-Regul. Pap.*, vol. 52, no. 7, pp. 1382–1393, Jul. 2005.

[6] S. El Assad, M. Farajallah, and C. Vladeanu, "Chaos-based block ciphers: An overview," in *2014 10th International Conference on Communications (COMM)*. IEEE, 2014, pp. 1–4.

[7] L.-Y. Zhang, Y. Liu, F. Pareschi, Y.-S. Zhang, K.-W. Wong, R. Rovatti, and G. Setti, "On the security of a class of diffusion mechanisms for image encryption," *IEEE Trans. Cybern.*, vol. 48, no. 4, pp. 1163–1175, Mar. 2017.

[8] G. Alvarez, F. Montoya, M. Romera, and G. Pastor, "Breaking two secure communication systems based on chaotic masking," *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 51, no. 10, pp. 505–506, Oct. 2004.

[9] Z.-Y. Hua and Y.-C. Zhou, "One-dimensional nonlinear model for producing chaos," *IEEE Trans. Circuits Syst. I-Regul. Pap.*, vol. 65, no. 1, pp. 235–246, Jul. 2017.

[10] C. López-Caraballo, I. Salfate, J. Lazzús, P. Rojas, M. Rivera, and L. Palma-Chilla, "Mackey-glass noisy chaotic time series prediction by a swarm-optimized neural network," in *Proceedings of Journal of Physics: Conference Series*, vol. 720, no. 1, 2016, p. 012002.

[11] X.-C. Cao, L. Du, X.-X. Wei, D. Meng, and X.-J. Guo, "High capacity reversible data hiding in encrypted images by patch-level sparse representation," *IEEE Trans. Cybern.*, vol. 46, no. 5, pp. 1132–1143, Apr. 2015.

[12] A. Miranian and M. Abdollahzade, "Developing a local least-squares support vector machines-based neuro-fuzzy model for nonlinear and chaotic time series prediction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 2, pp. 207–218, Dec. 2012.

[13] Y. Wang, Z.-L. Liu, J.-B. Ma, and H.-Y. He, "A pseudorandom number generator based on piecewise logistic map," *Nonlinear Dyn.*, vol. 83, no. 4, pp. 2373–2391, Nov. 2015.

[14] Y.-C. Zhou, Z.-Y. Hua, C.-M. Pun, and C. P. Chen, "Cascade chaotic system with applications," *IEEE Trans. Cybern.*, vol. 45, no. 9, pp. 2001–2012, Oct. 2014.

[15] Z.-Y. Hua, S. Yi, Y.-C. Zhou, C.-Q. Li, and Y. Wu, "Designing hyperchaotic cat maps with any desired number of positive lyapunov exponents," *IEEE Trans. Cybern.*, vol. 48, no. 2, pp. 463–473, Feb. 2017.

[16] Z.-Y. Hua and Y.-C. Zhou, "Dynamic parameter-control chaotic system," *IEEE Trans. Cybern.*, vol. 46, no. 12, pp. 3330–3341, Dec. 2015.

[17] M. Alawida, A. Samsudin, J. S. Teh *et al.*, "Deterministic chaotic finite-state automata," *Nonlinear Dyn.*, vol. 98, no. 3, pp. 2403–2421, Oct. 2019.

[18] M. Alawida, A. Samsudin, and J. S. Teh, "Enhanced digital chaotic maps based on bit reversal with applications in random bit generators," *Inf. Sci.*, vol. 512, pp. 1155–1169, Feb. 2020.

[19] M. Alawida, A. Samsudin, N. Alajarmeh, J. S. Teh, M. Ahmad *et al.*, "A novel hash function based on a chaotic sponge and dna sequence," *IEEE Access*, vol. 9, pp. 17 882–17 897, Jan. 2021.

[20] M. Alawida, A. Samsudin, J. S. Teh *et al.*, "Digital cosine chaotic map for cryptographic applications," *IEEE Access*, vol. 7, pp. 150 609–150 622, Oct. 2019.

[21] Z. Li, C.-G. Peng, L.-R. Li, and X.-Y. Zhu, "A novel plaintext-related image encryption scheme using hyper-chaotic system," *Nonlinear Dyn.*, vol. 94, no. 2, pp. 1319–1333, Jun. 2018.

[22] F. Dachselt and W. Schwarz, "Chaos and cryptography," *IEEE Trans. Circuits Syst. I-Regul. Pap.*, vol. 48, no. 12, pp. 1498–1509, Dec. 2001.

[23] P. Li, Z. Li, W. A. Halang, and G.-R. Chen, "A multiple pseudorandom-bit generator based on a spatiotemporal chaotic map," *Phys. Lett. A*, vol. 349, no. 6, pp. 467–473, Jan. 2006.

[24] X.-P. Lv, X.-F. Liao, and B. Yang, "A novel pseudo-random number generator from coupled map lattice with time-varying delay," *Nonlinear Dyn.*, vol. 94, no. 1, pp. 325–341, Jun. 2018.

[25] Q.-X. Wang, S.-M. Yu, C.-Q. Li, J. Lü, X.-L. Fang, C. Guyeux, and J. M. Bahi, "Theoretical design and FPGA-based implementation of higher-dimensional digital chaotic systems," *IEEE Trans. Circuits Syst. I-Regul. Pap.*, vol. 63, no. 3, pp. 401–412, Mar. 2016.

[26] K. Kaneko, "Pattern dynamics in spatiotemporal chaos: Pattern selection, diffusion of defect and pattern competition intermettency," *Physica D*, vol. 34, no. 1-2, pp. 1–41, 1989.

[27] Y. Wang, K.-W. Wong, and D. Xiao, "Parallel hash function construction based on coupled map lattices," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 16, no. 7, pp. 2810–2821, Jul. 2011.

[28] S. Coombes and A. H. Osbaldestin, "Period-adding bifurcations and chaos in a periodically stimulated excitable neural relaxation oscillator," *Phys. Rev. E*, vol. 62, no. 3, p. 4057, Sept. 2000.

[29] S. Kumar, M. Kumar, R. Budhiraja, M. Das, and S. Singh, "A cryptographic model for better information security," *J. Inf. Secur. Appl.*, vol. 43, pp. 123–138, Dec. 2018.

[30] K.-W. Wong, B. S. H. Kwok, and W. S. Law, "A fast image encryption scheme based on chaotic standard map," *Phys. Rev. A*, vol. 372, no. 15, pp. 2645–2652, Apr. 2008.

Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808

[31] X.-Q. Fu, B.-C. Liu, Y.-Y. Xie, W. Li, and Y. Liu, "Image encryption-then-transmission using DNA encryption algorithm and the double chaos," *IEEE Photonics J.*, vol. 10, no. 3, pp. 1–15, Apr. 2018.
[32] I. Hussain, J. Ahmed, and A. Hussain, "An image encryption technique based on coupled map lattice and one-time S-boxes based on complex chaotic system," *J. Intell. Fuzzy Syst.*, vol. 29, no. 4, pp. 1493–1500, 2015.
[33] Y. Wang, X.-F. Liao, D. Xiao, and K.-W. Wong, "One-way hash function construction based on 2D coupled map lattices," *Inf. Sci.*, vol. 178, no. 5, pp. 1391–1406, Mar. 2008.
[34] Y.-Q. Zhang and X.-Y. Wang, "A symmetric image encryption algorithm based on mixed linear-nonlinear coupled map lattice," *Inf. Sci.*, vol. 273, pp. 329–351, Jul. 2014.
[35] M. E. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Phys. Rev. E*, vol. 69, no. 2, p. 026113, Feb. 2004.
[36] S. Biswas and A. Das, "Patterns, bifurcations, multistability and hysteresis in an inhomogeneous coupled map lattice," *Int. J. Bifurcation and Chaos.*, vol. 26, no. 03, p. 1630008, Mar. 2016.
[37] L. E. Bassham III, A. L. Rukhin *et al.*, *A statistical test suite for random and pseudorandom number generators for cryptographic applications*. National Institute of Standards & Technology (NIST) Special Publication 800-22, Rev. 1a, 2010.
[38] P. L'Ecuyer and R. Simard, "Testu01: A C library for empirical testing of random number generators," *ACM Transactions on Mathematical Software (TOMS)*, vol. 33, no. 4, pp. 1–40, Aug. 2007.
[39] S.-J. Li, X.-Q. Mou, and Y.-L. Cai, "Pseudo-random bit generator based on couple chaotic systems and its applications in stream-cipher cryptography," in *Proceedings of International Conference on Cryptology in India*. Springer, Nov. 2001, pp. 316–329.
[40] K. Kaneko, "Overview of coupled map lattices," *Chaos*, vol. 2, no. 3, pp. 279–282, Jul. 1992.
[41] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, "Determining Lyapunov exponents from a time series," *Physica D*, vol. 16, no. 3, pp. 285–317, Jul. 1985.
[42] N.-S. Liu, "Pseudo-randomness and complexity of binary sequences generated by the chaotic system," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 16, no. 2, pp. 761–768, Feb. 2011.
[43] J. von Neumann, "Various techniques used in connection with random digits," *Appl. Math Ser.*, vol. 12, no. 36-38, p. 5, 1951.
[44] L. Devroye, "Sample-based non-uniform random variate generation," in *Proceedings of the 18th Conference on Winter Simulation*, 1986, pp. 260–265.
[45] M.-Z. Ding and W.-M. Yang, "Stability of synchronous chaos and on-off intermittency in coupled map lattices," *Phys. Rev. E*, vol. 56, no. 4, p. 4009, Oct. 1997.
[46] H. Wang and X. Guo, "Characteristic polynomials and spectra of some block circulant graphs," *Polycycl. Aromat. Compd.*, vol. 33, no. 2, pp. 83–96, Mar. 2013.
[47] J. Heidel, "The existence of periodic orbits of the tent map," *Phys. Lett. A*, vol. 143, no. 4-5, pp. 195–201, Jan. 1990.
[48] F. Pareschi, R. Rovatti, and G. Setti, "On statistical tests for randomness included in the NIST SP800-22 test suite and based on the binomial distribution," *IEEE Trans. Inf. Forensic Secur.*, vol. 7, no. 2, pp. 491–505, Apr. 2012.

Yong Wang is a Professor in Chongqing University of Posts and Telecommunications, China. He graduated with a PhD degree in computer science from Chongqing University, China, in 2007. His current research interests involve cryptography, differential privacy and information management, and he has published more than 50 refereed journal and conference articles in these fields.

Zhuo Liu is a Lecturer with the School of Mathematics and Big Data, Guizhou Education University, Guiyang, China. She is currently learning for her PhD degree in Chongqing University of Posts and Telecommunications, Chongqing, China. Her current research interests focus on the chaos-based cryptography fields of pseudo random number generation (PRNG), the image encryption and the dynamic behavior analysis in the higher-dimensional chaotic system, and she has published more than 5 refereed journal and conference articles in those fields.

Leo Yu Zhang (M'17) is currently a Lecturer with the School of Information Technology, Deakin University, VIC, Australia. He received the bachelor's and master's degrees in computational mathematics from Xiangtan University, Xiangtan, China, in 2009 and 2012, respectively, and the Ph.D. degree from the City University of Hong Kong, Hong Kong, in 2016. Prior to joining Deakin, he held various research positions with the City University of Hong Kong, the University of Macau, Macau, China, the University of Ferrara, Ferrara, Italy, and the University of Bologna, Bologna, Italy. His current research interests include applied cryptography and AI-related security, and he has published more than 60 refereed journal and conference articles in these fields.

Fabio Pareschi (Senior Member, IEEE) received the Dr. Eng. degree (Hons.) in electronic engineering from the University of Ferrara, Italy, in 2001, and the Ph.D. degree in information technology from the University of Bologna, Italy, in 2007, under the European Doctorate Project (EDITH). He is currently an Associate Professor with the Department of Electronic and Telecommunication, Politecnico di Torino. He is also a Faculty Member with ARCES, University of Bologna. His research activity focuses on analog and mixed-mode electronic circuit design, statistical signal processing, compressed sensing, dc-dc converters, random number generation and testing, and electromagnetic compatibility. Dr. Pareschi was a recipient of the 2019 IEEE BioCAS Transactions Best Paper Award. He also received the Best Paper Award at ECCTD 2005 and the Best Student Paper Award at EMC Zurich 2005 and IEEE EMCCompo 2019. He served as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–PART II from 2010 to 2013. He is currently Associate Editor for the IEEE OPEN JOURNAL OF CIRCUITS AND SYSTEMS.

Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCYB.2021.3129808

Gianluca Setti (Fellow, IEEE) received the Ph.D. degree in electronic engineering and computer science from the University of Bologna in 1997. From 1997 to 2017, he was with the School of Engineering, University of Ferrara, Italy. Since 2017 is a Professor of Electronics for Signals and Data Processing at the Politecnico di Torino, Italy. He is also a permanent Faculty Member of ARCES, University of Bologna. His research interests include nonlinear circuits, recurrent neural networks, electromagnetic compatibility, compressive sensing and statistical signal processing, biomedical circuits and systems, power electronics, design and implementation of IoT nodes, circuits and systems for machine learning, and ML and AI algorithms for anomaly detection. He was a recipient of the 2013 IEEE CAS Society Meritorious Service Award and a co-recipient of the 2004 IEEE CAS Society Darlington Award, the 2013 IEEE CAS Society Guillemin-Cauer Award, the 2019 IEEE BioCAS Transactions Best Paper Award, the Best Paper Award at ECCTD2005, and the Best Student Paper Award at EMCZurich2005, IEEE ISCAS2011 and IEEE EMCCompo2019. He held several editorial positions and served, in particular, as the Editor-in-Chief for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–PART II from 2006 to 2007 and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–PART I from 2008 to 2009. He was the Technical Program Co-Chair of ISCAS2007, ISCAS2008, ICECS2012, and BioCAS2013 and the General Co-Chair of NOLTA2006 and ISCAS2018. He was a Distinguished Lecturer of the IEEE CAS Society from 2004 to 2005 and from 2014 to 2015 and a member of its Board of Governors from 2005 to 2008. He served as the 2010 President for CASS. He held several other volunteer positions for the IEEE. Among these, from 2013 to 2014, he was the First Non North-American Vice President of the IEEE for Publication Services and Products, and, since 2019 he is the First non-US Editor-in-Chief of the Proceedings of the IEEE, the Flagship publication of the IEEE.

Guanrong (Ron) Chen (M'89, SM'92, F'97, LF'19) received the MSc degree in Computer Science from Sun Yat-sen University, Guangzhou, China in 1981 and the PhD degree in Applied Mathematics from Texas A&M University, USA in 1987. Since year 2000, he has been a Chair Professor and the founding director of the "Centre for Chaos and Complex Networks" at City University of Hong Kong.

Professor Chen was elected Fellow of the IEEE in 1997, awarded the 2011 Euler Gold Medal from Russia, and conferred Honorary Doctor Degrees by the Saint Petersburg State University, Russia in 2011 and by the University of Normandy, France in 2014. He is a Member of the Academy of Europe (since 2014) and a Fellow of The World Academy of Sciences (since 2015).

Professor Chen's research interests are in the fields of complex networks, nonlinear dynamics and control systems. He has been a Highly Cited Researcher in Engineering for many years according to Clarivate (Web of Science).

Copyright (C) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org