DYNAMICAL AND STATISTICAL ANALYSIS OF A NEW LOZI FUNCTION FOR RANDOM NUMBERS GENERATION

Andrea Espinel Rojas, Ina Taralova, René Lozi

DYNAMICAL AND STATISTICAL ANALYSIS OF A NEW LOZI FUNCTION FOR RANDOM NUMBERS GENERATION

Andrea Espinel, Ina Taralova
IRCCyN, UMR CNRS 6597
Ecole Centrale de Nantes
France
andrea.espinel-rojas, ina.taralova@irccyn.ec-nantes.fr

René Lozi
Laboratoire J.A. Dieudonné, UMR CNRS 6621
Université de Nice Sophia-Antipolis
France
lozi@unice.fr

Abstract
This paper presents the first results of the statistical and dynamical analysis of a new function showing random properties firstly proposed by Lozi. The phase plane analysis via the critical lines tool allowed to delimit analytically the holes in the chaotic attractor and to follow their evolution. In addition, the results of the statistical NIST tests for pseudo-randomness showed to be successful and significantly improved after an under-sampling of the output signal.

Key words
Lozi map, chaotic map, Random Number Generator, NIST test

1 Introduction

The incessantly increasing demand for secure data storage and transmission (e-banking, e-payments, personal data encoding...) motivates the research for newer and more secure data encryption techniques. The latter are classically performed via Pseudo Random Number Generators (PRNG) which, besides being highly reliable, should be able to generate as many different encoding sequences (hidden in the encryption keys) as possible. The encryption keys lie in the system parameters, since the structure is always supposed to be known by the pirates. For this reason, during the last decade, there has been a plethora of papers devoted to the nonlinear maps used for chaotic encryption.

Indeed, the well known intrinsic sensitivity to small parameter changes and initial conditions exhibited by the chaotic maps makes them perfect candidates for encryption. Thus, for each - even infinitesimal - parameter change, a different chaotic sequence will be generated, so in theory an infinite number of encoding sequences can be obtained - and therefore, an infinite number of keys (if we make abstraction of the quantization). Nevertheless, designing a chaotic PRNG remains a very tough problem, because chaoticity is only a necessary, but not a sufficient feature. Indeed, the encrypting sequence has to exhibit also a set of statistical properties [3], [4] and therefore not all chaotic maps are suitable for encryption purposes. However, most of the authors simply neglected the statistical properties, which have to be satisfied by the chaotic map, if used as PRNG. This is typically the case when the basin of attraction is not dense or exhibits holes, so the state variables are not equidistributed.

The most widely and universally used test to validate PRNG is the National Institute of Standards and Technology Test, known as NIST tests.

2 System Definition

The system under consideration has been proposed first by Lozi in [1] who emphasized its random features. It is defined on the p-dimensional torus $T^p = [-1,1[^{\,p}$ by the map $M_p : T^p \Rightarrow T^p$

$$M_p : \begin{cases} x^1_{n+1} = 1 - 2\left|x^1_n\right| + k^1 \times x^2_n \\ x^2_{n+1} = 1 - 2\left|x^2_n\right| + k^2 \times x^3_n \\ \vdots \\ x^p_{n+1} = 1 - 2\left|x^p_n\right| + k^p \times x^1_n \end{cases} \quad (1)$$

where the parameters $k^i = (-1)^{i+1}$ or $k^i = 1$, the latter case being considered hereafter. The flow x is contained on the torus:

$$\begin{array}{l} if \quad x^j_{n+1} = 1 - 2\left|x^j_n\right| + k^j \times x^{j+1}_n < -1 \\ \quad add \quad 2 \\ if \quad x^j_{n+1} = 1 - 2\left|x^j_n\right| + k^j \times x^{j+1}_n \geq 1 \\ \quad substract \quad 2 \end{array} \quad (2)$$

$|x_n|$ denotes the absolute value of x_n, therefore the map (1) is a noninvertible map (i.e. the backward iterates are not unique, or do not exist).

Hereafter we deal with the dynamical analysis of the second order system M_2 on the torus T^2, here simply denoted as M.

Therefore, considering that $|x|$ can take two different values, there are four regions in T^2 with locally linear behaviour (two for x^1, and two for x^2). Thus the map can also be considered as a piece-wise linear one.

2.1 Analysis of the 2D-System

Singularities of type 1. Fixed and periodic points.

The fixed points are to be studied independently in the four regions of T^2.They are defined by:

$$M(x) = x \qquad (3)$$

Keeping in mind that $x = (x^1, x^2)$ and $x^i \in [-1,1[$, for $x^1 > 0$, $x^2 > 0$, the fixed point is located at $(x^1, x^2) = (0.5,0.5)$. It is unstable with eigenvalues $(\lambda_1, \lambda_2) = (-3,-1)$. But it is numerically stable because of the structure of the floating point numbers.

For $x^1 < 0$, $x^2 < 0$, each point of the line $x^2 = -1 - x^1$ is an unstable fixed point with positive eigenvalues $(\lambda_1, \lambda_2) = (3,1)$.

For $x^1 \geq 0$, $x^2 < 0$, there is only one fixed point at $(x^1, x^2) = (0, -1)$, which is unstable. The eigenvalues are $(\lambda_1, \lambda_2) = (\sqrt{5}, -\sqrt{5})$.

For $x^1 < 0$, $x^2 \geq 0$, the single fixed point $(x^1, x^2) = (-1, 0)$ has the same eigenvalues as before, and is also unstable.

Due to the piece-wise linear nature of (Eq1– Eq.2), closed formulas can be found for every periodic solution, see fig.1 with all period-2 cycles:

$$x_n = x_{n+2}$$

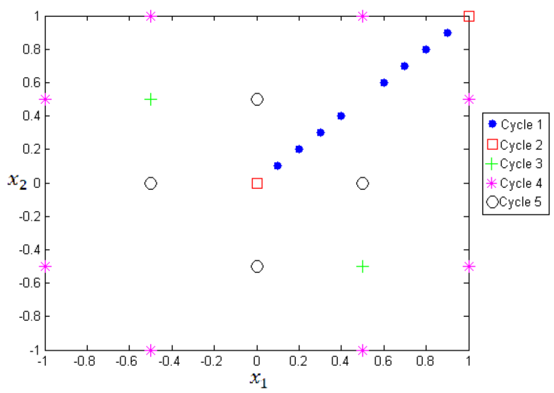

Figure 1. Period 2 solutions of the map M_2 (1) on the torus $T^2 = [-1,1[^2$

Singularities of type 2. Critical lines.

The critical lines CL [2] are singularities of dimension 1 and represent an important tool for the analysis of noninvertible maps. By definition, the critical lines separate regions of the phase space with different number of preimages (backward iterates). In the case of piece-wise linear maps, they are the first iterates of the lines of discontinuity CL_{-1} of the system.

For the two dimensional system M_2 there are four groups of critical lines CL with preimages CL_{-1} given by:

Critical Lines A

For $CL^A_{-1} : x^1 = 0$, we have:

$$CL_1^{A1} : x^2 = -2\,x^1 - 1 \quad if\ x^2 > 0$$
$$CL_1^{A2} : x^2 = 2\,x^1 - 1 \quad if\ x^2 < 0$$

Critical Lines B

For $CL^B_{-1} : x^1 = -1$

$$CL_1^{B1} : x^2 = 2\,x^1 \quad if\ x^2 < 0,\ x^1 \in [0,0.5]$$
$$CL_1^{B2} : x^2 = -2\,x^1 - 2 \quad if\ x^2 > 0,\ x^1 \in [-1,-0.5]$$
$$CL_1^{B3} : x^2 = 2\,x^1 - 2 \quad if\ x^2 < 0,\ x^1 \in [0.5,1[$$
$$CL_1^{B4} : x^2 = -2\,x^1 \quad if\ x^2 > 0,\ x^1 \in [-0.5,0]$$

Critical Lines C

For $CL^C_{-1} : x^2 = 0$

$$CL_1^{C1} : x^2 = -\frac{1}{2}\left(x^1 + 1\right) \quad if\ x^1 > 0$$
$$CL_1^{C2} : x^2 = \frac{1}{2}\left(x^1 + 1\right) \quad if\ x^1 < 0$$

Critical Lines D

For $CL^D_{-1} : x^2 = -1$

$$CL_1^{D1} : x^2 = \frac{x^1}{2} \quad if\ x^1 < 0,\ x^2 \in [0,0.5]$$
$$CL_1^{D2} : x^2 = -\frac{x^1}{2} - 1 \quad if\ x^1 > 0,\ x^2 \in [-1,-0.5]$$
$$CL_1^{D3} : x^2 = -\frac{x^1}{2} \quad if\ x^1 > 0,\ x^2 \in [-0.5,0]$$
$$CL_1^{D4} : x^2 = \frac{x^1}{2} + 1 \quad if\ x^1 < 0,\ x^2 \in [0.5,1[$$

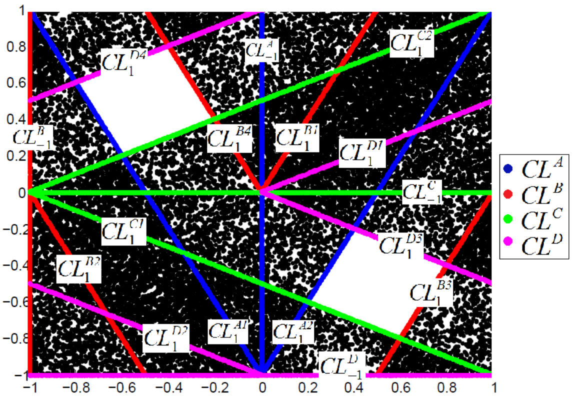

Figure 2. Critical lines of the map M_2 (1) on the torus $T^2 = [-1,1[^{\ 2}$

Fig.2 shows the invariant measure associated to the chaotic mapping for the second order system (the transient of the first 10^6 iterations has been cut off). It can be observed that this measure is not constant everywhere and vanishes on two diamond-like holes. The latter are completely delimited by segments of the critical lines CL_1^{A1}, CL_1^{B4}, CL_1^{C2}, CL_1^{D4}, and CL_1^{A2}, CL_1^{B3}, CL_1^{C1}, CL_1^{D3} so there evolution can be analytically followed under parameter or system order variation. Moreover, as the holes are symmetrical, the signal could be considered as symmetrically distributed as well (i.e. there are as many points in the half plane x<0 than in the half plane x>0). In the case where k^1=1 and k^2=-1, the invariant measure does not present the same pattern, see fig. 3, and needs a more sophisticated study.

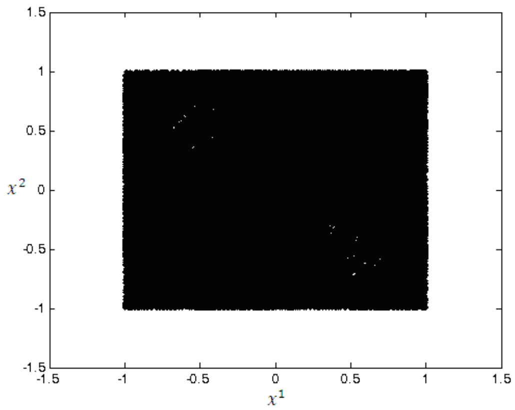

Figure 3. Invariant measure associated to the map M_2 (1) on the torus $T^2 = [-1,1[^{\ 2}$ where k^1=1 and k^2=-1

3 NIST Tests for the 4D-System

The measure of the holes decreases with the increase of the system dimension, and becomes neglectable for the four dimensional Lozi system M_4 (1) which shall be considered hereafter. As an example Fig 4 (resp.5) shows the projection of the invariant measure of 3D (resp.4D) - systems onto the (x^1, x^2) coordinates for values of x^3 and x^4 chosen inside the diamond-like holes, $0.49 \le x^3 \le 0.51$ (resp. $0.49 \le x^3, x^4 \le 0.51$).

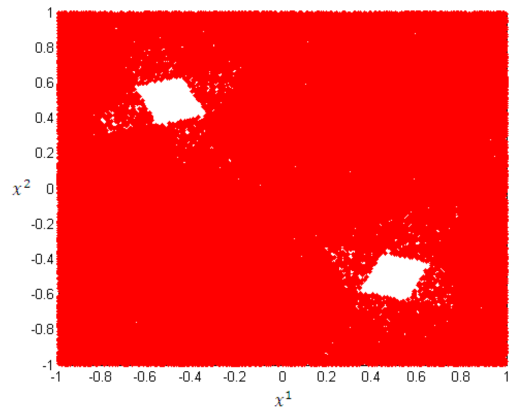

Figure 4. Invariant measure associated to the map M_3 (1) on the torus $T^3 = [-1, 1[^{\ 3}$

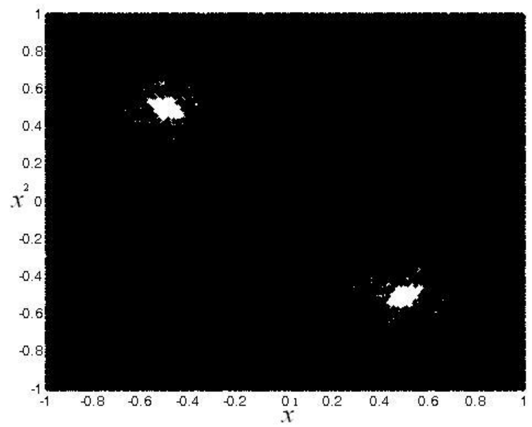

Figure 5. Invariant measure associated to the map M_4 (1) on the torus $T^4 = [-1, 1[^{\ 4}$

The output of the system has been arbitrary chosen as:

$$y_n = x_n^1$$

The NIST tests proposed by the National Institute of Standards and Technology [4] require a binary signal, whereas the generated chaotic output contains real values (floating point numbers), therefore an appropriate binarization has to be performed prior to applying the tests. Several types of floating-point binarization have been tested, such as the standard IEEE754 [5] that allows the conversion in single and double formats (32 or 64 bits). However, as the map (1) generates only numbers between $[-1,1[$, a significant number of bits remain always constant. Thus, periodic patterns appear necessarily in the sequence.

Finally, the selected approach in this paper has been to choose a threshold value, and classify all the numbers above this threshold as ones, and all the others as zeros.

The optimal threshold is a tricky problem, for the system (1) it has been naturally chosen as:

$$\begin{array}{ll} if\ y_n \geq 0 & b_n = 1 \\ else & b_n = 0 \end{array} \quad (4)$$

For the 4th order system M_4, the holes become much smaller, and it can be assumed that the output signal is equidistributed.

Length of the original sequence: 10^8 bits.

Initial condition (randomly chosen)

$$x_0 = [0.6324, -0.0975, 0.2785, -0.5469]$$

The statistical test evaluates the randomness of the sequence: the null hypothesis (H0) assumes that it is random, and the alternative hypothesis (Ha) assumes that it is not random.
For a successful test, the sequence must be accepted as being a random. The P-value (table 1) is a complex quantifier used to measure if the zeroes and the ones in the sequence can be considered as uniformly distributed.
A test is successful if the p-value is superior to the significance level (for this case 0.01). In addition, the minimum pass rate for each statistical test with the exception of the random excursion (variant) test is approximately = 57 for a sample size = 60 binary sequences (approximately = 96 for a sample size = 100). The minimum pass rate for the random excursion (variant) test is approximately = 46 for a sample size = 49 binary sequences (approximately = 64 for a sample size = 68). The failing tests are denoted by an asterisk (*).

For the NIST test, each bit stream is considered as a different sequence, so in order to evaluate the results, different lengths bit stream have been tested and compared.

For the first test, the original sequence has been divided into ten sub-sequences (or bit strings) having a length of 10^7 points.

Test 1 (shown in Table one)

Length of bit string: 1.666.666
Quantity of bit strings: 60

P-VALUE		PROPORTION		STATISTICAL TEST
0.500934		60/60		Frequency
0.020085		59/60		BlockFrequency
0.468595		60/60		CumulativeSums
0.862344		60/60		CumulativeSums
0.862344		59/60		Runs
0.082177		57/60		LongestRun
0.437274		59/60		Rank
0.148094		60/60		FFT
0.000000	*	47/60	*	NonOverlappingTemplate
0.000000	*	46/60	*	OverlappingTemplate
0.671779		59/60		Universal
0.000000	*	1/60	*	ApproximateEntropy
0.509162		46/47		RandomExcursions
0.098036		46/47		RandomExcursionsVariant
0.000000	*	54/60	*	Serial
0.706149		59/60		LinearComplexity

Table 1.

Conclusion: four of the tests fail.

For the second test, the original sequence has been divided into 100 bit strings, each of length 10^6.

Test 2 (shown in Table two)
Length of bit string: 1M
Quantity of bit strings: 100

P-VALUE		PROPORTION		STATISTICAL TEST
0.514124		99/100		Frequency
0.000474		100/100		BlockFrequency
0.719747		99/100		CumulativeSums
0.334538		99/100		CumulativeSums
0.000296		99/100		Runs
0.017912		96/100		LongestRun
0.419021		100/100		Rank
0.058984		98/100		FFT
0.000000	*	91/100	*	NonOverlappingTemplate
0.000000	*	85/100	*	OverlappingTemplate
0.419021		99/100		Universal
0.000000	*	20/100	*	ApproximateEntropy
0.392456		62/63		RandomExcursions
0.484646		60/63		RandomExcursionsVariant
0.013569		95/100	*	Serial
0.437274		99/100		LinearComplexity

Table 2.

Conclusion: again, four of the tests fail.

To improve the results, we applied an under-sampling which has been shown to improve the statistical properties of the signal [3]. For a sequence *S*, we take one bit out of ten, periodically.

$$S_{(10k)},\ k \in \mathbb{N}$$

Same initial condition:

$$x_0 = [0.6324, -0.0975, 0.2785, -0.5469]$$

Test 1
Length of bit string: 1.666.666
Quantity of bit strings: 60

P-VALUE	PROPORTION	STATISTICAL TEST
0.407091	60/60	Frequency
0.074177	60/60	BlockFrequency
0.232760	60/60	CumulativeSums
0.568055	60/60	CumulativeSums
0.602458	60/60	Runs
0.178278	59/60	LongestRun
0.148094	59/60	Rank
0.232760	59/60	FFT
0.834308	59/60	NonOverlappingTemplate
0.468595	59/60	OverlappingTemplate
0.500934	59/60	Universal
0.275709	60/60	ApproximateEntropy
0.611108	47/49	RandomExcursions
0.027405	49/49	RandomExcursionsVariant
0.772760	59/60	Serial
0.602458	59/60	Serial
0.253551	59/60	LinearComplexity

Table 3.

Test 2
Length of bit string: 1M
Quantity of bit strings: 100

P-VALUE	PROPORTION	STATISTICAL TEST
0.224821	100/100	Frequency
0.678686	99/100	BlockFrequency
0.334538	100/100	CumulativeSums
0.035174	100/100	CumulativeSums
0.383827	99/100	Runs
0.955835	97/100	LongestRun
0.739918	99/100	Rank
0.494392	99/100	FFT
0.437274	98/100	NonOverlappingTemplate
0.678686	100/100	OverlappingTemplate
0.455937	99/100	Universal
0.678686	99/100	ApproximateEntropy
0.392456	62/63	RandomExcursions
0.756476	61/63	RandomExcursionsVariant
0.779188	99/100	Serial
0.739918	98/100	LinearComplexity

Table 4.

Now, all tests are statistically successful; moreover it should be emphasized that the results do not depend on the bit strings quantity

4 Conclusion

Dynamical and statistical analysis demonstrated the efficiency of a new map firstly proposed by Lozi. The NIST tests carried out have been improved using a constant under-sampling. Future work with chaotic under-sampling has to be envisaged. The main difference between this model and other chaotic pseudo-random number generators is that this map not only provides one single stream of pseudo-random number but several uncorrelated parallel streams of numbers. This property is very useful in the case of simulation of multi-agent complex problem.

References

[1] R. Lozi, “Random properties of ring-coupled tent maps on the torus”, submitted to Discrete and continuous Dynamical Systems Series-B

[2] C. Mira et al, “Chaotic dynamics in two-dimensional noninvertible maps”, World Scientific Series on Nonlinear Science, Series A - Vol. 20, 1996

[3] S. Hénaff, I. Taralova, R.Lozi, “Dynamical Analysis of a new statistically highly performant deterministic function for chaotic signals generation”, International Conf. on Physics and Control (PhysCon), Catania, Sicily, September 2009.

[4] A. Rukhin, et al, “A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications”, NIST (2001), http://csrc.nist.gov/rng/

[5] W. Kahan, “IEEE Standard 754 for Binary FloatingPoint Arithmetic,” Lecture Notes on the Status of IEEE 754, Elect. Eng. & Computer Science University of California, Berkeley CA 94720-1776, May 1996.