DAMAGE SPREADING AND INFORMATION DISTANCE IN CELLULAR AUTOMATA

PROPAGACIÓN DE DAÑOS Y DISTANCIA INFORMACIONAL EN ATÓMATAS CELULARES

K. García-Medina[a], D. Estevez-Moya[a] y Ernesto Estevez-Rams [a,b†]

a) Facultad de Física, Universidad de la Habana, San Lázaro y L. CP 10400. La Habana. Cuba.
b) Instituto de Ciencias y Tecnología de Materiales, University of Havana (IMRE); estevez@fisica.uh.cu†
† corresponding author

Recibido 18/1/2022; Aceptado 12/11/2022

Using the concept of information distance derived from Kolmogorov randomness, we study damage spreading for elementary cellular automata acting on a one-dimensional lattice. In contrast to previous definitions of the Lyapunov exponent based on Hamming distance, the new magnitude allows a better clustering of chaotic rules. The combined use of the Lyapunov exponent, Hamming, and information distance-based, results in a more robust characterization of cellular automata behavior. An extension of the type analysis shown can be directly made to other one-dimensional time and space discrete dynamical systems.

La propagación de daños en autómatas celulares es estudiada utilizando distancia informacional, una magnitid derivada del uso de la complejidad algorítmica. Los autómatas celulares estudiados son los llamados elementales actuando sobre un arreglo unidimensional de sitios. Se define un exponente de Lyapunov derivado de la distancia algorítmica y se compara con definiciones más tradicionales del mismo. Una caracetrización más robusta de los autómatas es lograda a partir del uso combinado de exponentes de Lyapunov, distancia de Hamming y distancia informacional. El método expuesto, queda claro, es extensible a otros sitemas dinámicos que sean discretos en el espacio y en el tiempo.

PACS: Information theory (teoría de la información), 89.70.-a; Entropy in information theory (entropía en teoría de la información), 89.70.Cf; Dynamical systems (sistemas dinámicos), 05.45.-a.

I. INTRODUCTION

Lyapunov exponents usually characterize sensitivity to initial conditions in dynamical systems, but in a discrete space system, it is not directly applicable [1]. By taking two (infinitesimal) close initial conditions, the maximum Lyapunov exponent can be loosely defined as the rate of exponential divergence of the distance between two trajectories in the limit of infinite time [2]. In a discrete space system, there is no infinitesimal close initial condition. Nevertheless, how two initially close configurations diverge when submitted to the same evolution dynamics is still an important measure, and ways to define and extend Lyapunov exponent analysis to discrete systems have been proposed [1,3,4].

In a discrete spatial system, initial conditions are said to be closed when the distance between two configurations is that given by the smaller possible change between them, complying with the discrete spatial nature. The need to define a distance measure is therefore needed. The simplest one, usually taken, is the Hamming distance between the configurations, or the number of cells where the two configurations differ. In such a case, the minimum distance is effectively one or the equivalent normalized value. As both systems evolve, damage spreading is quantified by the same Hamming measure [1]. Several behaviors can be observed:

1. The initial perturbation may disappear, and the two configurations, after some finite number of steps, are the same and, therefore, evolve in the same manner. In the limit, the Hamming distance is zero.
2. The initial perturbation, beyond a certain finite number of time steps, does not grow anymore, although it can have a trajectory of its own: the Hamming distance is constant after a finite number of time steps.
3. The initial perturbation grows, and so does the divergence between the initial configurations: The Hamming distance is monotonically increasing with time.

If the interaction range in the system is finite, then, due to the discrete spatial nature, damage spreading can only occur linearly in the time limit.

The hamming distance can be deceiving. Consider, for example, a configuration given by alternating ones and zeroes, such that $s_i = i \ mod \ 2$, another configuration shifted only by one site (or an odd number of sites), resulting in $s_i = (i+1) \ mod \ 2$, will have a maximum Hamming distance while, it is clear, that they are conveying the same information.

Instead of using Hamming distance, other authors have used Kolmogorov complexity as the tool of choice [5]. Kolmogorov complexity of a system is given by the length of the shortest algorithm, able to describe the system running on a Universal Turing Machine [6, 7]. Using a Universal Turing Machine in the definitions makes the defined measure absolute up to a constant value. Kolmogorov complexity measures the randomness instead of complexity as it is currently understood (in what follows, we will

use the term Kolmogorov randomness). Distance metrics using Kolmogorov randomness have been defined [8]. Emmert-Streib has used one of such definitions to derive a hierarchical clustering of elementary cellular automata (ECA) [9], which are dynamical systems with nearest neighborhood range interaction acting over a boolean one-dimensional lattice [10]. This approach considers small configurations of cellular automata (CA) with 50 cells and 10^3 time steps. The system spatiotemporal evolution is analyzed as a unique string, by concatenating sequences derived from the temporal evolution [9].

Unlike a Hamming distance, information distances measure the closeness of two configurations by the computational effort of going from one to the other: it is an information distance [11]. The Kolmogorov-based distance will be small for the shift example described above with maximum Hamming distance.

In dynamical systems, such as cellular automata, damage spreading does not need to be a production and random walk process but, instead, can involve, besides the production of new symbols, the structuring of the perturbation as times evolve. This structuring behavior will not be captured by the Hamming-based distance and may affect the information-based distance. For two states, discrete time and space systems, Bagnoli et al. [3] extended the definition of Lyapunov exponents λ by using boolean derivatives, first introduced by Vichniac [12]. They applied the definition to ECA systems. They introduced a parameter μ that proved relevant as a control parameter, related to the fraction p of sites equal to one on the three principal diagonals of the Jacobian matrix [3]. The Hamming distance between two configurations, starting from a single site perturbation, is given by the Boolean Jacobian matrix of the global evolution function F [13]. The Jacobian of F is a tridiagonal matrix of ones and zeros. For a given critical value $p = p_c$, three groups can be identified according to the values of (μ, λ) : a first group with $(\mu, \lambda) = (0, -\infty)$; a second group $\mu > p_c$ and $\lambda > 0$ and; a third group with $\mu \approx p_c$ anad $\lambda > 0$. Furthermore, for ECA with $0 < \mu < p_c$ a random perturbation or small noise in the time evolution produces a collapse of the defined Lyapunov exponent to 0 or $-\infty$. In such a way, they could characterize ECA rules through Hamming distance, yet their approach carries the limitations of using Hamming distance as the system metric.

In this article, we will use an information distance based on Kolmogorov randomness to define a Lyapunov exponent. The information distance will be estimated by Lempel-Ziv factorization [14]. Using the μ parameter defined by Bagnoli et al. [3], ECA will be characterized. It will be shown that information distance allows a better partition of ECA rules and captures features of damage spreading that could not be followed by previous measures.

II. CELLULAR AUTOMATA

One dimensional cellular automata (CA) are defined over a lattice of length N (N can be infinite) with sites at time t, labeled as s_i^t, $i = 0, 1, \ldots, N$. s_i^t can take values, for a boolean CA, from an alphabet $\Sigma = \{0, 1\}$ of cardinality two. The whole lattice configuration at time t will be denoted by s^t. A local updating rule f, of range r, is defined such that at time $t + 1$, the value of the site s_i is given by

$$s_i^{t+1} = f[s_{i-r}^t, \ldots, s_i^t, \ldots, s_{i+r}^t] = f[{}_i^r\boldsymbol{s}^t] \in \{0, 1\} \tag{1}$$

where ${}_i^r\boldsymbol{s}^t$ stands for the subsequence $s_{i-r}^t, \ldots, s_i^t, \ldots, s_{i+r}^t$.

In what follows circular boundary conditions are enforced: $s_i = s_{(i \bmod N)}$ where (i mód N) means that i is reduced module N.

An overlapping partition of the global configuration s^t in blocks ${}_i^r\boldsymbol{s}^t$ of length $2r + 1$ allows applying the local rule f to every block, updating the whole configuration to s^{t+1} and therefore, inducing a global map F. If no ambiguity arises, the r superscript in ${}_i^r\boldsymbol{s}^t$ will be dropped: ${}_i\boldsymbol{s}^t$.

When $r = 1$, the cellular automata are called elementary (ECA), and to each ECA rule f, Wolfram proposed a notation number R that can be computed as [10]:

$$R = f(0,0,0)2^0 + f(0,0,1)2^1 + \ldots + f(1,1,1)2^7. \tag{2}$$

ECA rules can be partitioned into 88 equivalence classes as a result of mirror and reversion symmetries [10], and the analysis can then be reduced to a representative member of each type.

Wolfram devised a classification scheme for cellular automata that is still the most used one [1]. Starting from an arbitrary random initial configuration, CA is classified as:

- W1 : System evolves to a homogeneous state;
- W2 : System evolves to a (time) periodic behavior;
- W3 : System evolves to aperiodic chaotic patterns;
- W4 : System evolves to configurations with complex patterns and long-lived, correlated, and localized structures.

Reference to Wolfram classification of rules will be made, which are taken from Wolfram Alpha knowledge engine [15]. It has been proven that Wolfram classification is problematic for CA in classes W3 and W4, as, for a given rule, determining its class turns to be undecidable [16]. Additionally, some rules show high sensitivity to initial conditions, and different initial configurations can lead to different behaviors [10], making their classification even more ambiguous.

III. BOOLEAN DERIVATIVES OF CA, PERTURBATIONS AND LYAPUNOV EXPONENT

Consider a n-variable boolean function f, the partial derivative with respect to the k^{th} variable x_k, is another n-variable boolean function defined as [12, 13]

$$D_k f(\boldsymbol{x}) \equiv \frac{\partial f(\boldsymbol{x})}{\partial x_k} = f(\boldsymbol{x}) \oplus f(\boldsymbol{x} \oplus \boldsymbol{e}_k) \tag{3}$$

$\boldsymbol{x} = (x_i, \ldots, x_{i+n})$ is the n-variable argument of the function, defined in a n-dimensional vector space over the Galois field $\mathbb{F} = \{0, 1\}$, and $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ are the standard basis vectors,

$\boldsymbol{e}_1 = (1, 0, \ldots, 0), \boldsymbol{e}_2 = (0, 1, \ldots, 0), \ldots, \boldsymbol{e}_n = (0, 0, \ldots, 1)$.

$\oplus$ is the addition module 2 operation or boolean XOR

$$\varsigma_1 \oplus \varsigma_2 = \begin{cases} 1 & \varsigma_1 \neq \varsigma_2 \\ 0 & \varsigma_1 = \varsigma_2 \end{cases}$$

The addition $\oplus$ between two vectors is defined as the XOR operation between its components,

$$\boldsymbol{x} \oplus \boldsymbol{y} = (x_1 \oplus y_1, x_2 \oplus y_2, \ldots, x_n \oplus y_n).$$

The boolean derivative (3) will be zero if, given the vector $\boldsymbol{x}$, the value of the function f does not depend on the value of x_k; it will be one otherwise.

The boolean function will correspond to our local rule for the ECA. In such case we can define the Jacobian matrix $\boldsymbol{J}^t(F)$ at time t of the global Boolean function F, with entries

$$J_{ik}^t = \frac{\partial f[_i\boldsymbol{s}^t]}{\partial s_k} \tag{4}$$

If $|i-k| > 1$, then $J_{ik} = 0$, which renders the Jacobian matrix for ECA rules as a tridiagonal matrix with entries of 1s or 0s. This last fact allows us to deal effectively with the matrix even if we consider the CA lattice infinite. If the lattice has N sites, the Jacobian matrix will be $N \times N$.

Consider an initial configuration $\boldsymbol{s}^0$ and the initial perturbed configuration $\boldsymbol{q}^0 = \boldsymbol{s}^0 \oplus \boldsymbol{\delta}^0$, where $\boldsymbol{\delta}^0$ is a perturbation vector (which is usually taken to be one of the $\boldsymbol{e}_i$, a minimum perturbation). The perturbed vector at time t will be $\boldsymbol{\delta}^{t+1} = \boldsymbol{s}^{t+1} \oplus \boldsymbol{q}^{t+1}$, which can be shown [13] to be given, in a linear approximation, by

$$\boldsymbol{\delta}^{t+1} \approx \bigoplus_{i=1,\ldots,N} \delta_i^t \wedge D_i F(\boldsymbol{s}^t) \tag{5}$$

$\wedge$ is the boolean AND operator

$$\varsigma_1 \wedge \varsigma_2 = \begin{cases} 1 & \varsigma_1 = \varsigma_2 = 1 \\ 0 & \text{otherwise} \end{cases}$$

For the ECA, equation (5) reduces to

$$\delta_i^{t+1} \approx \delta_{i-1}^t \wedge \frac{\partial F}{\partial x_{i-1}} \oplus \delta_i^t \wedge \frac{\partial F}{\partial x_i} \oplus \delta_{i+1}^t \wedge \frac{\partial F}{\partial x_{i+1}} \tag{6}$$

Bagnoli et al. [3] derived a Lyapunov exponent from equation (5) (or eq.6) by the following procedure. Define $\boldsymbol{\Gamma}^0 = \boldsymbol{\delta}^0$, then

$$\boldsymbol{\Gamma}^{t+1} = \boldsymbol{J}^t \cdot \boldsymbol{\Gamma}^t \tag{7}$$

The maximum Lyapunov exponent is defined as

$$\lambda_\Gamma(T) = \frac{1}{T} \sum_{t=1}^{T} \log \vartheta_t \tag{8}$$

where the perturbation propagation rate is defined by

$$\vartheta_t = \frac{|\Gamma^{t+1}|}{|\Gamma^t|} \tag{9}$$

where $|\Gamma| = \sum \Gamma_i$. In the infinite limit $\lambda_\Gamma \equiv \lambda_\Gamma(\infty)$.

For ECA, consider $\mu(t)$ as the fraction of sites that are different from zero on the three principal diagonals of $\boldsymbol{J}(F)$

$$\mu(t) = \frac{1}{3N} \sum_{i=1}^{N} (J_{i,i-1} + J_{i,i} + J_{i,i+1}) \tag{10}$$

which allows to define a parameter $\mu(T)$ as the geometric mean for large enough time T

$$\mu(T) = \left(\prod_{t=1}^{T} \mu(t) \right)^{1/T}. \tag{11}$$

$\mu(T)$ has been used to characterize the Jacobian matrix and therefore the corresponding ECA rule [3].

To understand the meaning of Γ, consider the limit of small initial perturbation, which corresponds to a single site difference in the initial configurations; when, during evolution, n defects appear, replicas for each produced perturbed site are considered. By doing so, perturbed sites are taken as individual non-interacting objects, and the annihilation of defects is impossible. Γ_i^t is the number of replicas carrying the defect $\boldsymbol{e}_i^t$ at time t [3].

It must be noticed that the Lyapunov exponent, as defined by equation (8), characterizes the damage production but not its structuring. They are insensible to whether perturbations are generated in a specific pattern or if they are generated and undergo some unpredictable random walk. To capture the possible structuring of perturbation production, we must resort to a different measure, and for that, Kolmogorov algorithmic randomness will be used.

IV. KOLMOGOROV BASED NORMALIZED INFORMATION DISTANCE, PERTURBATIONS AND LYAPUNOV EXPONENT

Given the shortest program s^* that, when running in a Universal Turing Machine (UTM), reproduces the string s, the Kolmogorov complexity or Kolmogorov randomness $K(s)$ of the string is the length of s^* [7],

$$K(s) = |s^*|. \tag{12}$$

Kolmogorov randomness measures how easier it is to convey the information in s using a shorter description s^*. At most, the Kolmogorov randomness will be equal to the string length when the string has no redundancy, which happens for a completely random sequence. The UTM condition is needed to make the definition sound. While Kolmogorov randomness is maximum for a random string, it will attain its smallest value for a constant sequence of only one symbol. The conditional Kolmogorov randomness $K(s|p)$ is the length of the shortest program that can compute s from the string p. If p conveys much information regarding s, then $K(s|p)$ will be small. The joint Kolmogorov randomness $K(s,p)$ is the size of the smallest program that computes both strings s and p. From a theoretical point of view, the most common type of program allowed for defining Kolmogorov randomness is prefix-free,

where no program is a proper prefix of another program [7]. It holds that

$$K(s,p) \cong K(s) + K(p|s*) = K(p) + K(s|p*) \tag{13}$$

where $\cong$ denotes that equality is valid up to a constant value independent of p and s.

Kolmogorov randomness is uncomputable, a fact related to the halting problem [7], but the entropy density defined by

$$h(s) = \lim_{|s|\to\infty} \frac{K(s)}{|s|} \tag{14}$$

can be estimated in several ways. The entropy density defined in (14) is equivalent to the entropy density defined within Shannon information theory, which opens the possibility to use estimates of Shannon entropy rates as estimates of $h(s)$.

It is straightforward to see that the information about s contained in p can be defined by

$$I(s:p) = K(s) - K(s|p^*). \tag{15}$$

Which, as a result of equation (13), is symmetrical, up to a constant, on its arguments: $I(s:p) \cong I(p:s)$.

Now we are in a condition to define the information distance between two sequences s and p as [8]:

$$d_{ID}(s,p) = \frac{max\{K(s|p^*), K(p|s^*)\}}{max\{K(s), K(p)\}} = 1 - \frac{I(s:p)}{K(p)} \tag{16}$$

d_{ID} complies, up to a constant value, with the triangle inequality, the symmetry axiom, and the identity axiom.

$d_{ID}(s,p)$ differs from a Hamming-type distance; it is not a direct damage field measure but, instead, is an information distance because it quantifies how correlated two sequences are from the algorithmic or information perspective. Two sequences that can be derived one from the other by a small-sized algorithm will have a small d_{ID}. In comparison, two strings that are not algorithmically correlated will have a d_{ID} near one. From the definition, the reader can see that if two sequences are highly random, but one is very similar site by site, they will have a small Hamming distance and a small d_{ID} value. However, two sequences, one being a shift of the other, can have a large Hamming distance and yet will deliver a small d_{ID} value, as the shifting program is simple and small compared to the string lengths.

We can cast equation (16) in terms of entropy density by dividing the numerator and the denominator by the string length and considering that $|s| = |p|$

$$d_h(s,p) = \frac{max\{h(s|p^*), h(p|s^*)\}}{max\{h(s), h(p)\}} \tag{17}$$

which, using (13), can be written as

$$d_h(s,p) = \frac{h(s|p) - min\{h(s), h(p)\}}{max\{h(s), h(p)\}} \tag{18}$$

The halting problem makes it impossible to compute Kolmogorov randomness, which is the main drawback in its practical use; instead, alternatives based on the compressibility of the mathematical description of the system are usually introduced [11]. The Lempel-Ziv factorization will be used to estimate the entropy density, as the supplementary material explains.

Consider an initial configuration s^t and the perturbed configuration $q^t = s^t \oplus \delta^t$; for any time step, the information distance between both sequences can be computed $d_h(s^t, q^t) \equiv d_h^t$. A similar expression to (9) can be defined,

$$\zeta_t = \frac{d_h^{t+1}}{d_h^t} \tag{19}$$

From there, a corresponding Lyapunov exponent follows

$$\lambda_h(T) = \frac{1}{T}\sum_{t=1}^{T} \log \zeta_t \tag{20}$$

Again, $\lambda_h(\infty) \equiv \lambda_h$.

V. SIMULATION CONDITIONS

1. For the calculation of λ_Γ, a lattice of 512 cell was used, 5000 time steps were taken.
2. For the calculation of λ_h a lattice of 10^4 cell was used, 10^4 time steps were taken.

The difference in the system sizes is justified in terms of computation efficiency. Lempel-Ziv estimations need longer sequences to guarantee a better estimation as the convergence of the estimates is slow [11].

Ten different random initial conditions were used in both cases, taken from http://www.random.org. Perturbation on the initial conditions was done by changing a single site at the centre of the initial lattice configuration.

VI. DAMAGE SPREADING CHARACTERIZATION BY LYAPUNOV EXPONENTS

Bagnoli et al. [3] definition of the Lyapunov parameter λ_Γ is not exactly a sum over the Hamming distance spreading. By making replicas of each error as it appears, the annihilation of defects is not taken into account; this results in the Hamming field may saturate while the Lyapunov factor continues to grow.

By definition, if $\lambda_\Gamma < 0$, the damage field is contractive, and after several steps, both the perturbed and the non-perturbed systems will behave identically. As defect interaction is not taken into account, the contractive nature of the damage field can only be a result of the interaction of the defect with the underlying configuration, identical in both perturbed and non-perturbed systems. Error annihilation overcomes error production. If $\lambda_\Gamma > 0$, error production increases with time, overcoming error annihilation by interaction with the underlying configuration. For $\lambda_\Gamma = 0$, after a sufficient number of steps, error production and annihilation equal.

Simulation results lead to the characterization of rules' behavior concerning their maximum Lyapunov exponent λ_Γ in five types:

- Rules have strong (single) Lyapunov value $\lambda_\Gamma = -\infty$; that is, the value of the exponent is independent of the initial random condition.
- Rules have strong (single) Lyapunov value $\lambda_\Gamma = 0$ also independent of the initial random condition.
- Rules can have two Lyapunov values $\lambda = -\infty$ or $\lambda_\Gamma = 0$ depending on the initial random condition.
- Rules have strong (single) Lyapunov value $\lambda_\Gamma > 0$ independent of the initial random condition.
- Depending on the initial random condition, rules have a mixture of two or more behaviors, one resulting in a Lyapunov value of $\lambda_\Gamma > 0$.

Maximum Lyapunov exponents for all rules are listed in Table 1, while rules with strong behaviors are listed in Table 2.

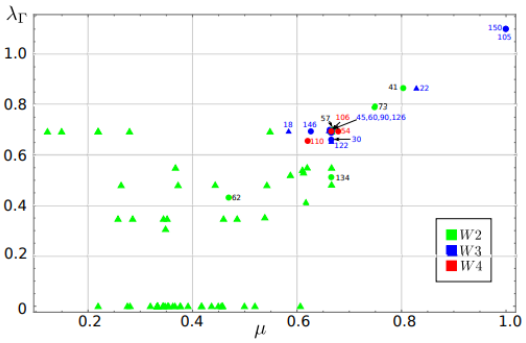

Figure 1. Non-negative maximum Lyapunov exponent (eq. (8)) as a function of the μ parameter for the all minimum equivalent ECA. Results were obtained from simulations carried over each cellular automaton with a 512-cell random initial condition evolving during 2500 time steps using a single site perturbation at the initial time. The average value over 10 realizations is shown. Triangles correspond to rules with also negative or zero Lyapunov exponents.

Fig. 1 shows the maximum Lyapunov exponent λ_Γ for each non-equivalent ECA for which $\lambda_\Gamma \geq 0$, as a function of the μ parameter. Up to $\mu < 0.42$, all rules belong to Wolfram class $W1$ and $W2$ and have at least one λ_Γ negative or zero value. A number of these rules can also show a positive maximum Lyapunov exponent, depending on the initial condition. Strong $\lambda_\Gamma = -\infty$ rules all have a $\mu = 0$ value, and they belong to $W1$ type in the Wolfram classification scheme. All rules with strong $\lambda_\Gamma = 0$ values, have $\mu < 0.44$ except rule 56 with $\mu = 1/2$. According to Bagnoli et al. [3], a second-order phase transition occurs in the random matrix approximation at $\mu_c = 0.441$. The same study found that for rules with $\mu < \mu_c$, the maximum Lyapunov exponent collapses to a negative or zero value if noise is added to their evolution. Our simulations show that for all rules with a μ value below the critical value μ_c, the probability of a positive maximum Lyapunov exponent is much lower than the probability of a non-positive exponent.

[1] $LCM(a, b)$ is the least common multiple of integer numbers a and b.

All rules with strong positive maximum Lyapunov exponent have μ values larger than μ_c.

Rules with constant Jacobian matrix have $\lambda_\Gamma = \log(3\mu)$.

$\lambda_\Gamma > 0$ rules can be in any of the Wolfram classifications, except $W1$, which is not surprising for $W3$ and $W4$ rules but can come as a surprise for $W2$ type rules. $W2$ rules for long enough times to settle into fixed configurations or periodic behavior. In either case, this fact and the discrete nature of time imply that the perturbation field cannot grow indefinitely. Suppose the orbit of the unperturbed and perturbed initial condition will end in a periodic cycle of periods T_0 and T_p, respectively. In that case, the corresponding damage field will also exhibit a periodic cycle of period $LCM(T_0 \times T_p)$[1]. As a result, for all $W2$ rules, an ever-expanding damage field should not happen, and the perturbation value d^t will eventually saturate.

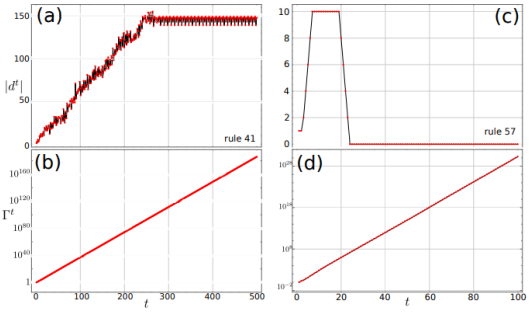

Figure 2. W2 rules with strong $\lambda_\Gamma > 0$ condition. Rule 41 (a) after an initial spreading of defects settles into a periodic behavior, yet (b) Γ^t increases exponentially. Rule 57 (c) collapses its damage field to zero after a few time steps, yet (d) Γ^t still shows an exponential increase.

Rules $W2$ that can show positive λ_Γ values show two trends; in one case, for certain rules and initial conditions, the ECA evolution ends in a periodic behavior with a period larger than one. Rules $1, 3, 5-7, 9, 25-29, 33, 35, 37, 38, 41, 57, 62, 73, 74, 94, 108, 134, 154, 156$ show such damage field cycles with periods larger than one. Fig. 2a and b shows rule 41 as an example of this class. Observe that in spite that Γ^t has exponential growth, the defect field, for sufficiently long times, has a periodic behavior. No periodic cycle can be identified in the second case, even for long running times. In this case, despite the exponential growth of the Γ^t values (Fig. 2d), the damage field ends into a constant value (Fig. 2c). In both cases, defects are massively produced by the ECA rule, as reflected by the exponential growth of Γ, but also massive defect annihilation occurs as a consequence of defect interaction, mediated by the whole of the cell configuration. In both cases, $\lambda_\Gamma > 0$ does not signal uncontrolled damage spreading in the context of the particular ECA rule evolution.

In those $W2$ rules with strong positive λ_Γ values (rules $41, 57, 62, 73, 134$), it is the rule which enforces both defect production and annihilation, independent of the lattice configuration. This dynamic results in either the freezing of the damage field into a local periodic pattern or its complete disappearance (see Fig. 3). For those rules where $\lambda_\Gamma > 0$ coexist with other behaviors, the production of defects is mediated by the cell environment.

It should be noticed that $W3$ rules 18, 22, 122, and 126 are usually perturbation expansive ($\lambda_\Gamma > 0$) but, for certain initial conditions, can show a negative maximum Lyapunov exponent. Finally, rules 105 and 150 exhibit the largest λ_Γ value of $\log 3$, corresponding to a constant Jacobian matrix with all derivatives equal to 1. The damage field of these two rules shows fractal behavior (see Fig. 3 for rule 150).

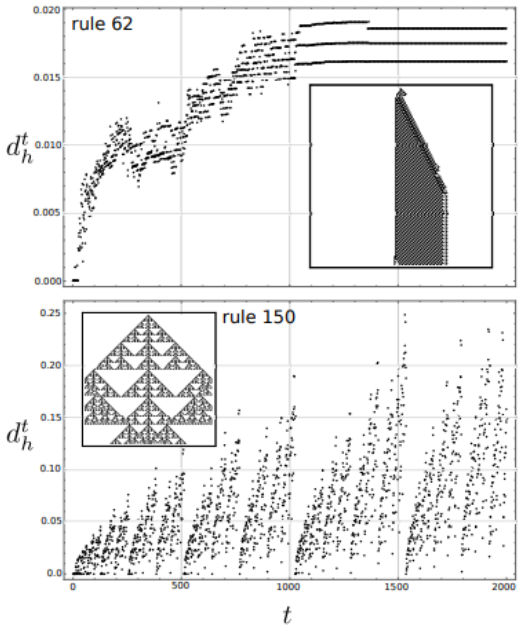

Figure 3. Information distance behaviour for two rules with strong $\lambda_\Gamma > 0$. Perturbed sequence has as initial sequence the same as the unperturbed sequence, with only the central site negated. Damage field were calculated by plotting in black those sites were the unperturbed and perturbed values are different, and are shown as inset to the figures, as a function of time. The Rule 62 (above) settles into a damage field (inset) with periodic spatial behavior. This results in a small d_h^t, of the order 10^{-2}, value for all times, in spite that the damage field initially spreads and finally saturates. Rule 150 (below) has a fractal damage field (inset). The d_h^t values are a order of magnitude larger than in rule 42.

When λ_h is plotted against λ_Γ, as shown in Fig. 4, there is no linear relation between both magnitudes; furthermore, there is no strict monotonic relation between them. Certainly, rules with the higher λ_Γ also correspond to the higher λ_h values, which also conform mostly to $W3$ and $W4$ type rules. The second interesting fact of Fig. 4 is that all $W2$ rules have smaller λ_h values than the more complex rules of type $W3$ and $W4$. Third, some rules with large $\lambda_\Gamma (> 0.6)$ have λ_h values near zero (rules 72, 104, 108, 200, 232). Even a strong $\lambda_\Gamma > 0$ rule, such as 134, shows a near zero λ_h value. Rule 57 with $\lambda_\Gamma > 0.6$ has a small λ_h value compared to the $W3$ and $W4$ rules. The only $W2$ rule with $\lambda_\Gamma > 0.6$ that have a larger λ_h value compared to the other rules of the same type is rules 73, which shows an asymptotic temporal periodic pattern of period larger than one and with a rich structured spatial organization, resulting in large damage field spread before freezing. This rule is far from producing a trivial spatiotemporal map.

The rules with the largest λ_h values are Wolfram numbers 60, 90, 105, and 150, with values well above other rules. The three rules have fractal damage spreading field.

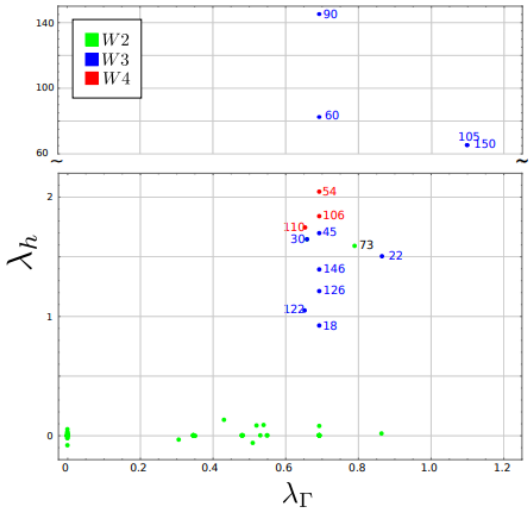

Figure 4. λ_h vs maximum Lyapunov exponent based on Hamming distance λ_Γ (eq. 8). λ_h was estimated using Lempel-Ziv factorization over cellular automaton lattice of 10^4 sites and averaged over ten different random initial conditions using a single site perturbation at the initial time. λ_Γ was estimated same as in figure 5. Only $\lambda_\Gamma > 0$ are shown.

Compare the behavior of the d_h^t function for rules 62 and 150. The first, which has a strong $\lambda_\Gamma > 0$ behavior, has a damage field that ends in periodic behavior (Fig. 3 above, inset). The d_h^t function reflects the long-term periodic behavior, characterized by a tree value cycle (Fig. 3 above). In any case, the d_h^t does not exceed the value of 0.020. For the 150 rule, as already explained, the damage field has a fractal character (Fig. 3 below, inset) which is captured by the d_h^t function (Fig. 3 below). The value of the d_h^t is an order of magnitude larger than the 62 rule. In both cases, Γ^t shows exponential growth, unable to capture the differences in the spatiotemporal evolution of both rules.

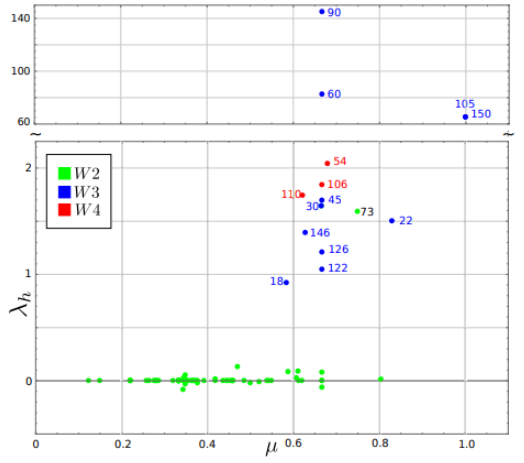

Figure 5. Lyapunov exponent based on information distance λ_h (eq. (20)) as a function of the μ parameter. Lyapunov exponents were estimated as in figure 4. A clear clustering of W3 and W4 type rules can be seen.

Fig. 5 plots the λ_h as a function of μ. Compared to fig 1, the use of λ_h allows a clear clustering of $W3$ and $W4$ rules from the non-complex one.

VII. CONCLUSION

In this work, we have shown that it is possible to introduce a Lyapunov exponent based on information distance as defined through Kolmogorov randomness. The measure capture characteristics of ECA evolution, which are not captured by previously defined maximum Lyapunov exponent based on Hamming distance. As information distance is related to the algorithmic burden of going from one sequence to a second one, information distance is not a Hamming-type distance but is sensible to perturbation structuring. In this last sense, it complements the use of λ_Γ. While positive λ_Γ values can falsely identify unlimited damage spreading, simulations show that λ_h allows a better clustering of ECA rules in correspondence to better characterizing chaotic behavior.

The Lyapunov exponent λ_h is not limited to cellular automata analysis but can be readily extended to other discrete time and space one-dimensional series.

VIII. APPENDIX: LEMPEL-ZIV FACTORIZATION AND ESTIMATION OF ENTROPY DENSITY

Consider a string s, of finite length N, made out of characters s_i taken from finite alphabet (e.g $\{0,1\}$). Let $s(i,j) = s_i \dots s_j$ be a substring of the string s; if $j < i$, then we take the substring as empty. A factorization of s, is a non-overlapping partition of the string s in substrings, such that, if we consider ab as the concatenation of string a and b, then we can write $s = s(1,l_1)s(l_1+1,l_2)\dots,s(l_{c-1},l_c = N)$. The Lempel-Ziv factorization **[?]** $F(s)$ is constraint by the following two conditions for each substring in the factorization:

1. $s(l_{k-1}+1,l_k)\pi \subset s(1,l_k)\pi^2$
2. $s(l_{k-1}+1,l_k) \not\subset s(1,l_k)\pi$ except, perhaps, for the last factor $s(l_{m-1}+1,N)$.

Where the "drop" operator π is defined as

$$s(i,j)\pi = s(i,j-1)$$

and, consequently,

$$s(i,j)\pi^k = s(i,j-k).$$

The partition $F(s)$ is unique for every string.

For example, the Lempel-Ziv factorization of the sequence $u = 10011101001011$ is $F(s) = 1.0.01.11.010.0101.1$, where a dot delimits each factor.

The LZ76 complexity $C_{LZ}(s)$ $(= |F(s)|)$ of the sequence s, is defined as the number of factors in its factorization. In the example above, $C_{LZ}(s)$=7.

Defining

$$c_{LZ}(s) = \frac{C_{LZ}(s)}{N/\log N}. \tag{21}$$

Ziv [17] proved that, if s is the output from an ergodic source, then

$$\lim_{N\to\infty} \sup c_{LZ}(s) = h(s). \tag{22}$$

Where $h(s)$ is the same entropy rate given by equation (14). This is the base of using c_{LZ} as an estimate of h_μ for $N \gg 1$.

Eq. (22) is valid in the infinite limit. In practical cases, eq. (21) is used as an estimate for the entropy density (further details can be found in [11]).

IX. ACKNOWLEDGEMENTS

EER and KG will like to thank AvH for their financial support. DEM and EER will like to thank MPI-PKS for financial support under visiting grants.

REFERENCES

[1] S. Wolfram, Phys. D **10**, 1 (1984).
[2] H. Kantz and T. Schreiber, Nonlinear time series analysis (Cambridge University Press, Cambridge, New York, 2003).
[3] F. Bagnoli, R. Rechtman and A. Enciso, Chaos **7**, 688 (1997).
[4] J. Urias, R. Rechtman and A. Enciso, Chaos **7**, 688 (1997).
[5] J. C. Dubacq, B. Durand and E. Formenti, Th. Comp. Science **259**, 271 (2001).
[6] A. N. Kolmogorov, Probl. Inf. Transf. (English Trans.) **1**, 1 (1965).
[7] M. Li and P. Vitányi, An Introduction to Kolmogorov Complexity and Its Applications, (Springer Verlag, 1993).
[8] M. Li, X. Chen., X. Li, B. Ma and P. M. B. Vitanyi, IEEE Trans. Inf. Th. **50**, 3250 (2004).
[9] F. Emmert-Streib, Phys. Rev. E **81**, 026103 (2010).
[10] S. Wolfram, A new kind of science (Wolfram media Inc., Champaign, Illinois, 2002).
[11] E. Estevez-Rams, R. L. Serrano, C. A. J. Nunes and B. Aragon-Fernández, Chaos **25**, 123106 (2015).
[12] G. Vichniac, Physica D **45**, 62 (1997).
[13] A. M. del Rey, G. R. Sánchez and A. de la Villa Cuenca, Appl. Math. Lett. **25**, 739 (2012).
[14] A. Lempel and J. Ziv, IEEE Trans. Inf. Th IT-22, 75 (1976).
[15] Wolfram Research Inc., Wolfram alpha knowledge engine (2019), (Available online at http://www.wolframalpha.com/).
[16] K. Culik and S. Yu, Complex Systems **2**, 177 (1988).
[17] J. Ziv, IEEE Transf. Inf. Th IT-24, 405 (1978).

Table 1. List of maximum Lyapunov exponent λ_Γ and μ values for all non-equivalent ECA rules. Blue corresponds to $W3$ type rules and red to $W4$ type rules[2].

No.	λ_Γ	μ	No.	λ_Γ	μ
0	-∞ = log(3μ)	0	56	0	1/3
1	-∞;0;1/2 log 3	0.368(2)	57	log 2 = log (3μ)	2/3
2	-∞;0	1/3	58	-∞; 0	0.417(6)
3	-∞;0;1/2 log 2	0.352(6)	60	log 2 = log (3μ)	2/3
4	-∞;0	0.344(4)	62	0.43(2)	0.47(2)
5	-∞;0;1/2 log 2	0.345(6)	72	-∞; log 2	0.15(2)
6	-∞;0;0.54(1)	0.611(3)	73	0.79(3)	0.749(4)
7	-∞;0;1/2 log 2	0.258(7)	74	-∞;0	0.520(6)
8	-∞	0	76	-∞;0	0.32(1)
9	-∞;0	0.607(2)	77	-∞;0	0.22(2)
10	-∞;0	1/3	78	-∞;0;0.480(3)	0.373(3)
11	-∞;0	0.3774(8)	90	log 2 = log (3μ)	2/3
12	-∞;0	1/3	94	-∞;0;1/2 log 2	0.486(5)
13	-∞;0.48(2)	0.26(1)	104	-∞; log 2	0.12(2)
14	-∞;0	0.34(1)	105	log 3 = log (3μ)	1
15	0 = log (3μ)	1/3	106	log 2 = log (3μ)	2/3
18	-∞; log 2	0.584(2)	108	0; log 2	0.549(8)
19	-∞;0;1/2 log 2	0.28(1)	110	0.654(3)	0.6212(2)
22	-∞;0.865(2)	0.830(1)	122	-∞;0.652(3)	2/3
23	-∞;0; log 2	0.22(2)	126	-∞; log 2	2/3
24	-∞;0	0.458(4)	128	-∞	0
25	-∞;0.52(1)	0.588(2)	130	-∞;0	1/3
26	-∞;0;0.411(3)	0.6176(5)	132	0	0.355(6)
27	-∞;0	0.458(4)	134	0.51(1)	2/3
28	0;0.48(1)	0.543(3)	136	-∞	0
29	0;1/2 log 2	0.460(6)	138	-00;0	0.367(7)
30	0.6596(6)	0.6665(1)	140	-∞;0	0.353(9)
32	-∞	0	142	-∞;0.306(3)	0.349(9)
33	0;0.55(7)	0.620(2)	146	log 2	0.6273(6)
34	-∞;0	1/3	150	log 3 = log (3μ)	1
35	-∞;0	0.437(2)	152	-∞;0	0.456(3)
36	-∞;0	0.377(9)	154	0;0.482(9)	2/3
37	0;0.352(4)	0.539(2)	156	0; log 2; 1/2 log 3	2/3
38	-∞;0;0.53(1)	0.613(2)	160	-∞	0
40	-∞	0	162	-∞;0	0.276(7)
41	0.864(1)	0.8036(8)	164	0	0.392(6)
42	-∞;0	0.363(3)	168	-∞	0
43	-∞;0	0.34(1)	170	0 = log (3μ)	1/3
44	-∞;0;0.481(1)	0.444(5)	172	-∞;0	0.418(4)
45	log 2 = log (3μ)	2/3	178	-∞;0; log 2	0.22(1)
46	-∞;0	0.450(8)	184	0	0.344(7)
50	-∞;0	0.281(8)	200	-∞;0; log 2	0.28(1)
51	0 = log (3μ)	1/3	204	0 = log (3μ)	1/3
54	log 2	0.680(1)	232	-∞;0; log 2	0.22(1)

Table 2. Strong type rules. Blue corresponds to W3 rules and W4 rules in red.

strong λ_Γ <0, μ = 0								
rule	0	8	32	40	128	136	160	168
strong λ_Γ = 0, W2								
rule	15	51	56	132	164	170	184	
strong λ_Γ >0								
rule	30	41	45	54	57	60	62	73
	90	105	106	110	134	146	150	

[2]The number between parentheses is the error in the calculation.