



Density-based recurrence measures from microstates

Felipe Eduardo Lopes da Cruz ^{1,2,3,*}, Thiago de Lima Prado,³ Sergio Roberto Lopes,³
Norbert Marwan,^{2,4,5} and Jürgen Kurths ^{1,2}


¹*Humboldt Universität zu Berlin, Berlin 10099, Germany*

²*Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, Potsdam 14473, Germany*

³*Departamento de Física, Universidade Federal do Paraná, Curitiba 81531-980, Brazil*

⁴*Institute for Geosciences, University of Potsdam, Potsdam 14476, Germany*

⁵*Institute for Physics and Astronomy, University of Potsdam, Potsdam 14476, Germany*

 (Received 23 October 2024; revised 4 March 2025; accepted 25 March 2025; published 17 April 2025)

Recurrence analysis is a powerful tool for nonlinear time series analysis deeply rooted in the theory of dynamical systems, finding applications across many areas of science. It works by mapping recurrences of a time series or phase space trajectory into a logical matrix. Recurrence quantification analyses (RQAs) are computed from its internal structures, such as recurrence density and the distribution of diagonal and vertical lines. Here, we link the density-based recurrence measures such as determinism and laminarity to the concept of microstates. We present a way to obtain the histogram of both diagonal and vertical lines from recurrence microstates, which are small square submatrices of the recurrence matrix. This approach opens up a line of research by reframing traditional RQAs in terms of microstates. Therefore, we establish a bridge between concepts of traditional lines-based RQA and recurrence microstates, and illustrate this for various paradigmatic systems.

DOI: [10.1103/PhysRevE.111.044212](https://doi.org/10.1103/PhysRevE.111.044212)

I. INTRODUCTION

Correctly understanding data provided by experiments or simulations has become an increasingly evident need in science. Perhaps the most important example in our days is information obtained by machine learning approaches, which depend decisively on sufficient sampling and completeness of data sets. In this sense, retrieving the maximum information from the data is essential for building a complete set of reliable information from the systems. Such sets, in general, are composed of space transformations of the data [1–4], as well as a collection of their representative quantifiers [5,6].

A basic property of dynamical systems is recurrence [7], and a very successful approach for studying recurrence properties is the use of recurrence plots (RPs). RPs are visual representations of a square logical matrix with elements generated by

$$R_{i,j} = \Theta(\varepsilon - \|\mathbf{x}_i - \mathbf{x}_j\|), \quad (1)$$

with $\Theta(\cdot)$ being the Heaviside function, equal to 1 when its argument is non-negative and 0 otherwise. $\mathbf{x}_{i(j)}$ is the $i(j)$ th element of a time series \mathbf{x} (or state space trajectory), and ε is a vicinity threshold. This method is fundamentally based on the results by Poincaré [7], who strongly contributed to our understanding of recurrence properties of dynamical systems.

Since its development in 1987 by Eckmann *et al.* [8], who originally proposed this technique as a tool to visualize even high-dimensional trajectories, several theoretical extensions and applications of RPs have been proposed [9,10]. In particular, the recurrence quantification analysis (RQA) [6,9,11–13], was introduced shortly after the original introduction of the RP, which consists of quantifying its internal structures, such as density of points [12], and distribution of diagonal [12,13] and vertical [14] lines. RQA then quickly became the standard technique for recurrence analysis.

The quantities extracted from these structures are called recurrence measures or recurrence quantifiers. Among them are the density of recurrence points, the recurrence rate (RR), the determinism (DET—the fraction of points that form diagonal lines), and the laminarity (LAM—analogue to the determinism, but for vertical lines). These quantities help us detect transitions in the dynamical regime of a system under study, e.g., DET can identify transitions between regular and chaotic dynamics (and vice versa) through local extrema [15], whereas LAM can reveal transitions between different chaotic regimes or intermittency, having larger values during laminar states [14]. Moreover, RQA can also be used for characterization of dynamics, which is especially useful when paired with learning methods.

Traditionally, the calculation of DET and LAM comes from the statistics of the diagonal and vertical lines within the RP. As a consequence, we need two separate distributions. Proposing alternatives for the calculation and approximation of these quantifiers is currently an important problem [10,16,17] that needs to be addressed due to the high computational cost when large time series are analyzed, and also for algorithms that must be based in real-time recurrence quantification [18].

*Contact author: felipe.eduardo@ufpr.br

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

In this work, we introduce an alternative approach for calculating DET and LAM based on recurrence microstates. Recurrence microstates are small submatrices of a RP, representing local recurrence patterns [18–22] which, within the framework of recurrence network analysis, [23] can be considered as graph motifs [24,25]. This approach only needs a single distribution of microstates, hence simplifying the calculation process for both DET and LAM, by extracting information without the need to distinguish between different types of line structures. Moreover, the values of DET and LAM derived from this innovative method are equivalent to those obtained through the traditional line-based approach, keeping the same interpretation and application of these measures. This approach not only simplifies the computation but also offers a different perspective on the quantification of recurrence structures, such as the *possibility of obtaining simultaneously different RQA measures from a unique probability distribution, as well as improving theoretical understanding of the small-scale structures in RPs.*

Here, we derive the expressions for DET and LAM based on microstate distributions, showing a comparison with the traditional line-based approach. We demonstrate that the proposed method retains the accuracy of the existing measures with sampling only a fraction of points in the RP.

This work is structured as follows: Section II derives the general expressions for recurrence microstates. Section III relates these probabilities to the histograms of diagonal and vertical lines. Following this, there is a section in which we perform a numerical analysis of the precision of our method for paradigmatic models with different dynamics, showing their errors and general validity. In Sec. V we conclude the paper with final remarks and discuss the possibilities created by this approach.

II. PROBABILITIES OF RECURRENCE MICROSTATES

A recurrence microstate [19,26] is defined as a matrix of size k embedded in a RP matrix of size N [8], usually, but not necessarily, for $k \ll N$ and defined as

$$\mathbf{M}^{(k)} \doteq \begin{pmatrix} b_{1,1} & \cdots & b_{1,k} \\ \vdots & \ddots & \vdots \\ b_{k,1} & \cdots & b_{k,k} \end{pmatrix}, \quad (2)$$

where b assumes the binary values 1 or 0 for a pair of recurrent or nonrecurrent elements of the data. In this way, a RP matrix may be composed of microstates of different sizes k ranging from 1 (single point) to N (full RP), denoted by $\mathbf{M}^{(k)}$ and exemplified by Eq. (3),

$$\mathbf{R} \doteq \begin{pmatrix} (R_{1,1}) & \cdots & R_{1,N} \\ \vdots & & \vdots \\ \vdots & \begin{pmatrix} R_{i,j} & R_{i,j+1} \\ R_{i+1,j} & R_{i+1,j+1} \end{pmatrix} & \vdots \\ R_{N,1} & \cdots & R_{N,N} \end{pmatrix}. \quad (3)$$

Let us call β an element of a string labeling all possible microstates of size k , e.g., for $k = 1$, $\beta \in [1, 2]$, for $k = 2$, $\beta \in [1, 2, \dots, 15, 16]$ or, in general, for any $k \in [1, N]$, it is true that $\beta \in [1, 2, \dots, 2^{k^2}]$. In this way, $\mathbf{M}_\beta^{(k)}$ is one of all

possible microstates of size k and embedded in the RP, e.g.,

$$\begin{aligned} \mathbf{M}_1^{(1)} &\doteq (0), & \mathbf{M}_2^{(1)} &\doteq (1), \\ \mathbf{M}_1^{(2)} &\doteq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \dots, \mathbf{M}_9^{(2)} &\doteq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \dots \\ \mathbf{M}_{16}^{(2)} &\doteq \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \end{aligned} \quad (4)$$

Intuitively, the total number of 1's in the RP can be counted by adding each possible value $R_{i,j}$ within the RP. This is used to calculate the RR, since normalizing this number by the total size of the RP results in a density-like quantity. Therefore, as the RP is made of binary values, the total number of 0's is just the size of the RP subtracted by the number of 1's, and a single 0 is just $1 - R_{i,j}$. It is also possible to search for *patterns* instead of single points. This is seen in expressions for the histograms of lines and can be performed by multiplying individual comparisons. We can apply the same logic to search for microstates.

As an example, let us consider the case of $k = 2$ microstates. With the previous arguments, the probabilities of finding the completely nonrecurrent $\mathbf{M}_1^{(2)}$ ($\beta = 1$), namely, the microstate pattern [0 0 0 0], and the completely recurrent $\mathbf{M}_{16}^{(2)}$ ($\beta = 16$), with pattern [1 1 1 1], are given, respectively, by

$$\begin{aligned} p_1^{(2)} &= \sum_{i,j=1}^{N-1} \frac{(1 - R_{i,j})(1 - R_{i,j+1})(1 - R_{i+1,j})(1 - R_{i+1,j+1})}{(N-1)^2}, \\ p_{16}^{(2)} &= \sum_{i,j=1}^{N-1} \frac{R_{i,j}R_{i,j+1}R_{i+1,j}R_{i+1,j+1}}{(N-1)^2}, \end{aligned} \quad (5)$$

by associating 0's to $(1 - R)$ and 1's to (R) (each term with its individual indices) and multiplying them to represent a Boolean “logical equivalence” operation. The term $(N - 1)^2$ is a normalization factor, otherwise the sum would return an integer number.

When sampling a $k \times k$ region from the RP, starting from (i, j) , we end up with a data-dependent submatrix, given the dependence of \mathbf{R} on the data \mathbf{x} shown in Eq. (1). Therefore, we should compare them to preexistent patterns, as in Eq. (2), to give them a label β . These patterns are the recurrence microstates. Evidently, in practical applications, the sampled submatrix is just interpreted as a binary string and converted to an integer, otherwise the sheer number of comparisons for each sample could potentially render the process unfeasible.

Algebraically, we can compute the probability of finding any particular microstate in a RP with a scalar expression as in

$$\begin{aligned} p_\beta^{(k)} &= \frac{1}{(N - k + 1)^2} \sum_{i,j=1}^{N-k+1} \prod_{\mu,v=0}^{k-1} (-1)^{\bar{b}_{1+\mu,1+v}} \\ &\quad \times (R_{i+\mu,j+v} - \bar{b}_{1+\mu,1+v}), \end{aligned} \quad (6)$$

where $R_{i,j}$ are the elements of the RP and $\bar{b}_{i,j}$ are the *negated elements* of the β th $k \times k$ microstate, i.e., if $b_{i,j} = 1$ (0), then $\bar{b}_{i,j} = 0$ (1). We can also write, explicitly, $\bar{b}_{i,j} = 1 - b_{i,j}$, keeping in mind that although β is an integer index labeling any possible microstate of size k , both $R_{i,j}$ and $b_{i,j}$ are binary

TABLE I. Symmetries for the $k = 2$ case. These are always valid as they are merely a consequence of the symmetric nature of a recurrence plot [not considering fixed amount of neighbors (FAN) criterion of choice for ε]. When using FAN, the RP ceases to be symmetric. These are not all 2×2 microstates, they are just the subset of microstates of this size that are related by transposition. The remaining microstates of this size are symmetric. We are also not seeing all possible recurrence correlations, just the subset related by the translation of indices from Eq. (11).

Symmetries		
Recurrence Microstates	Microstate Probabilities	Recurrence Correlations
$\mathbf{M}_3^{(2)} = [\mathbf{M}_5^{(2)}]^\top$	$p \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = p \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\langle R_{i+1,j} \rangle^{(2)} = \langle R_{i,j+1} \rangle^{(2)}$
$\mathbf{M}_4^{(2)} = [\mathbf{M}_6^{(2)}]^\top$	$p \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = p \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$	$\langle R_{i+1,j} R_{i+1,j+1} \rangle^{(2)} = \langle R_{i,j+1} R_{i+1,j+1} \rangle^{(2)}$
$\mathbf{M}_{11}^{(2)} = [\mathbf{M}_{13}^{(2)}]^\top$	$p \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = p \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$	$\langle R_{i,j} R_{i+1,j} \rangle^{(2)} = \langle R_{i,j} R_{i,j+1} \rangle^{(2)}$
$\mathbf{M}_{12}^{(2)} = [\mathbf{M}_{14}^{(2)}]^\top$	$p \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = p \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$	$\langle R_{i,j} R_{i+1,j} R_{i+1,j+1} \rangle^{(2)} = \langle R_{i,j} R_{i,j+1} R_{i+1,j+1} \rangle^{(2)}$

variables, since they are just elements of the RP and the microstates, respectively.

So, the straightforward way to compute probabilities of specific microstates is to multiply combinations of $R_{i+\mu,j+\nu}$ and $(1 - R_{i+\mu,j+\nu})$ terms with the respective indices to represent the desired microstate pattern, where $R_{i+\mu,j+\nu}$ should be used on the microstate element where we want to find 1 and $(1 - R_{i+\mu,j+\nu})$ where we want to find 0. From now on, we compact the equations for the probabilities, assuming the notation

$$\frac{1}{(N-k+1)^2} \sum_{i,j=1}^{N-k+1} [\cdot] \rightarrow \langle \cdot \rangle^{(k)} \quad (7)$$

such that Eqs. (5) can be written as

$$\begin{aligned} p_1^{(2)} &= \langle (1 - R_{i,j})(1 - R_{i,j+1})(1 - R_{i+1,j})(1 - R_{i+1,j+1}) \rangle^{(2)}, \\ p_{16}^{(2)} &= \langle R_{i,j} R_{i,j+1} R_{i+1,j} R_{i+1,j+1} \rangle^{(2)}. \end{aligned} \quad (8)$$

We now proceed to assign a meaning for each product term in Eqs. (8). Note that by expanding the products, we have terms from degree zero to degree k^2 in $R_{i,j}$. For example, the first degree terms are accurate approximations for RR, since

$$\frac{1}{(N-k+1)^2} \sum_{i,j=1}^{N-k+1} R_{i+\mu,j+\nu} \rightarrow \langle R_{i+\mu,j+\nu} \rangle^{(k)} \approx RR \quad (9)$$

for any combination of indices $(i + \mu, j + \nu)$ and μ, ν ranging from 0 to $k - 1$ for $k \ll N$, which is the standard practice. The previous expression is exact for $k = 1$. Expanding the expressions derived from Eq. (6), we get functions of the same form as correlation functions, since the correlation \mathcal{C} between n points in a signal x can be calculated by averaging the product of the values of the signal at different times, denoted

$$\mathcal{C}(t_1, \dots, t_n) = \langle x(t_1) \cdots x(t_n) \rangle. \quad (10)$$

With this, we can safely call these quantities in the new notation ‘‘recurrence correlations,’’ given that they are, in fact, very similar (see Table II) to correlations between recurrences at different locations in the RP. For a fixed recurrence threshold and a metric as distance function in Eq. (1), the RP is a

symmetric matrix. This means that if we look at its transposed counterpart we can see, with help from Fig. 2, that the density lost below (above) the main diagonal is recovered above (below) it. Hence, the operation $row \leftrightarrow column$ makes the densities invariant. In fact, Table I shows that we can get the basic symmetries of the recurrence microstates from these functions. This clarifies that these functions (the products of different $R_{i,j}$) can be used in the same context as the recurrence microstates. Since the RP is symmetric, the equally valued recurrence correlations are found by their invariance under translation of indices

$$\langle R_{i+\mu,j+\nu} \rangle^{(k)} = \langle R_{i+\nu,j+\mu} \rangle^{(k)}, \quad \mu \leftrightarrow \nu, \quad (11)$$

with (μ, ν) defined as in Fig. 1 and, for the products of more than one term, like

$$\langle R_{i+\mu_1,j+\nu_1} R_{i+\mu_2,j+\nu_2} \rangle^{(k)} = \langle R_{i+\nu_1,j+\mu_1} R_{i+\nu_2,j+\mu_2} \rangle^{(k)}, \quad (12)$$

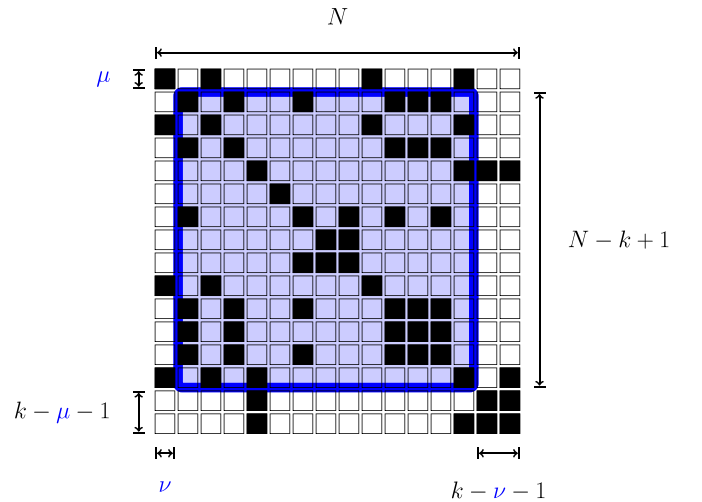


FIG. 1. A square $(N - k + 1) \times (N - k + 1)$ region (in blue) from a recurrence plot representing the set of indices $(i + \mu, j + \nu)$ allowed in the sum of Eq. (11). The shifts (μ, ν) present in this same equation are defined as the number of rows and columns to ignore prior to selecting the square blue region.

the invariance is with respect to *simultaneous* translations of the type

$$\mu_1 \leftrightarrow \nu_1, \quad \mu_2 \leftrightarrow \nu_2, \quad (13)$$

being readily generalized for the product of any number of recurrences. We define a vector whose components are these recurrence correlations, following a standard lexicographical ordering. Let us call this vector $\mathbf{c}^{(k)}$. For the 2×2 case, it can be written as

$$\mathbf{c}^{(2)} = \begin{pmatrix} 1 \\ \langle R_{i,j} \rangle^{(2)} \\ \langle R_{i,j+1} \rangle^{(2)} \\ \langle R_{i+1,j} \rangle^{(2)} \\ \langle R_{i+1,j+1} \rangle^{(2)} \\ \langle R_{i,j} R_{i,j+1} \rangle^{(2)} \\ \langle R_{i,j} R_{i+1,j} \rangle^{(2)} \\ \langle R_{i,j} R_{i+1,j+1} \rangle^{(2)} \\ \langle R_{i,j+1} R_{i+1,j} \rangle^{(2)} \\ \langle R_{i,j+1} R_{i+1,j+1} \rangle^{(2)} \\ \langle R_{i,j} R_{i,j+1} R_{i+1,j} \rangle^{(2)} \\ \langle R_{i,j} R_{i,j+1} R_{i+1,j+1} \rangle^{(2)} \\ \langle R_{i,j+1} R_{i+1,j} R_{i+1,j+1} \rangle^{(2)} \\ \langle R_{i,j} R_{i+1,j} R_{i+1,j+1} \rangle^{(2)} \\ \langle R_{i,j+1} R_{i+1,j} R_{i+1,j+1} \rangle^{(2)} \end{pmatrix}. \quad (14)$$

These are the fundamental terms that constitute the probabilities of all recurrence microstates, including the constant component 1, which represents the linear combination given by the sum of all probabilities. All components of the vectors $\mathbf{c}^{(k)}$ are distinct linear combinations of the components of $\mathbf{p}^{(k)}$, and, conversely, the reciprocal is also true. Therefore, we can define a family of isomorphisms $\Lambda^{(k)}$ represented by a family of full-rank $2^{k^2} \times 2^{k^2}$ matrices relating the vectors $\mathbf{c}^{(k)}$ and $\mathbf{p}^{(k)}$. Let us denote their inverses $[\Lambda^{(k)}]^{-1}$ as $\Lambda^{-(k)}$, for simplicity. This can be translated into the following equations:

$$\mathbf{p}^{(k)} = \Lambda^{(k)} \mathbf{c}^{(k)}, \quad (15)$$

and its inverse

$$\mathbf{c}^{(k)} = \Lambda^{-(k)} \mathbf{p}^{(k)}. \quad (16)$$

These correlations can be visually explained with the help of Fig. 2. They are densities of certain regions of the RP, in the case of those of $\mathcal{O}(R)$ (e.g., of the form $\langle R_{i+\mu,j+\nu} \rangle^{(k)}$), that are delimited by the microstates size k and the lags μ and ν . For the higher-order ones (e.g., of the form $\langle R_{i+\mu_1,j+\nu_1} R_{i+\mu_2,j+\nu_2} \rangle^{(k)}$), they are densities of the Hadamard (element-wise) product of their respective regions, which are also delimited by k and their respective lags (μ_1, ν_1) and (μ_2, ν_2) . Figure 2 also visually explains the higher-order ones (bottom). In all cases, the regions are always $(N - k + 1) \times (N - k + 1)$. From this, we see that, in general, we can expect $\langle R \rangle^{(k)} \geq \dots \geq \langle R \dots R \rangle^{(k)}$.

The transformation matrices, $\Lambda^{(k)}$ and $\Lambda^{-(k)}$, that take us from correlations to microstates and vice versa, can easily be pregenerated, since they do not depend on any specific system

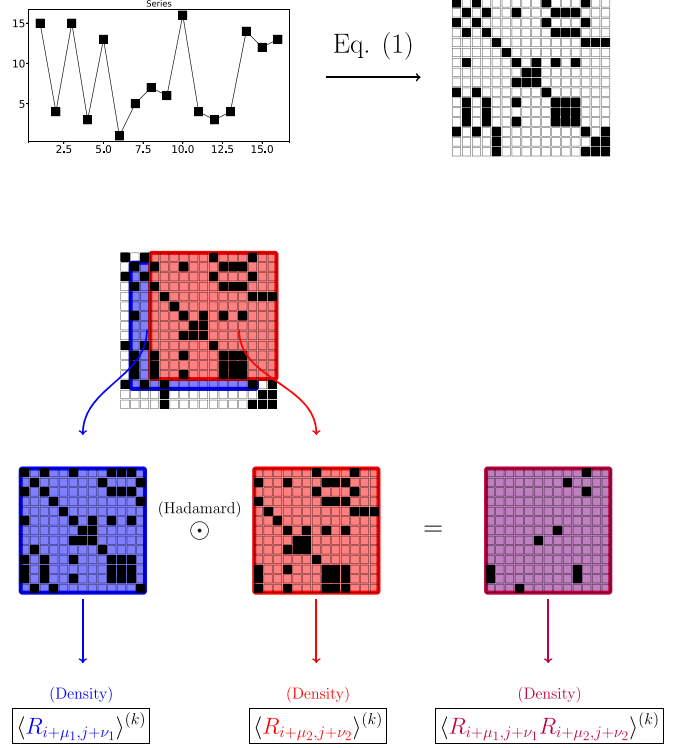


FIG. 2. Schematic diagram illustrating one possible interpretation of the recurrence correlations. The top part shows the use of Eq. (1) to generate a RP from a time series. The bottom part presents an interpretation of the correlations $\langle R_{i+\mu_1,j+\nu_1} \rangle^{(k)}$ and $\langle R_{i+\mu_2,j+\nu_2} \rangle^{(k)}$ as the local densities of points of the blue and red regions. The interpretation of $\langle R_{i+\mu_1,j+\nu_1} R_{i+\mu_2,j+\nu_2} \rangle^{(k)}$ is that it is the density of the Hadamard product of those two regions. They are defined by the parameters k, μ_1 , and ν_1 , and k, μ_2 , and ν_2 , respectively, as demonstrated in Fig. 1.

and are only a property of the sets of microstates. To understand the meaning behind these matrices we first must remember that we are dealing with vectors, more specifically, a probability vector containing the probabilities of microstates $\mathbf{p}^{(k)} = (p_1^{(k)}, \dots, p_{2^{k^2}}^{(k)})^\top$ and a vector of recurrence correlations $\mathbf{c}^{(k)} = (1, \langle R_{i,j} \rangle^{(k)}, \dots, \langle R_{i,j} \dots R_{i+k-1,j+k-1} \rangle^{(k)})^\top$, generated from the expansion of the polynomial terms in Eq. (6). The matrices $\Lambda^{(k)}$ and $\Lambda^{-(k)}$ then represent how one set of variables should be combined to generate the other set. The special case for $k = 3$ is depicted in Fig. 3. These relationships are independent of the system being analyzed, but depend solely on how we order the elements of $\mathbf{p}^{(k)}$ and $\mathbf{c}^{(k)}$.

The inverse transformation matrix $\Lambda^{-(3)}$ depicted as a special case in Fig. 3 reveals a pattern that strongly resembles the Sierpiński gasket [27] (the resolution of the pattern increases with k). It is possible to perform a permutation to this matrix such that it is directly mapped to a projection of the actual Sierpiński gasket into a square. The aforementioned permutation can be easily found through a \mathcal{QR} factorization [28], in which we find an orthogonal matrix \mathcal{Q} and an upper triangular \mathcal{R} such that

$$\Lambda^{-(3)} = \mathcal{Q}^{(3)} \mathcal{R}^{(3)}. \quad (17)$$

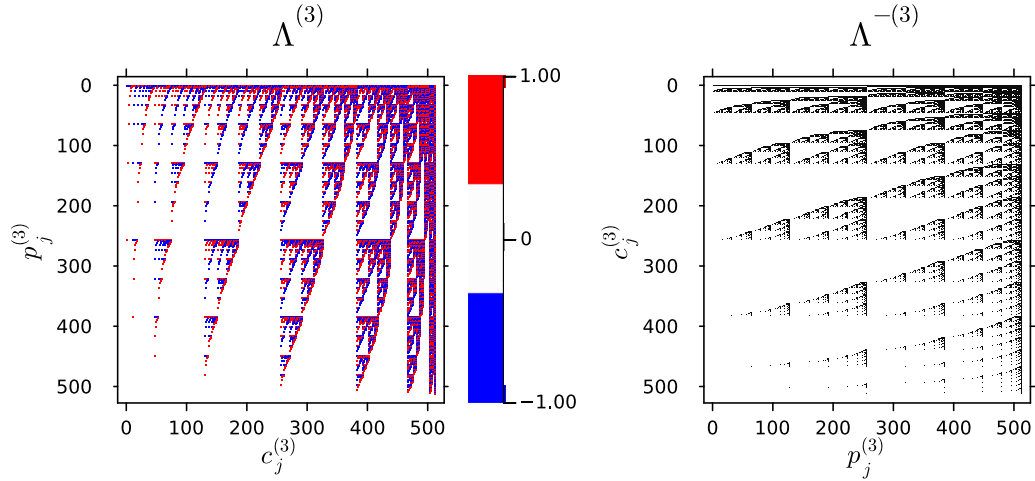


FIG. 3. On the left, the matrix $\Lambda^{(3)}$ with the following color scheme: blue for -1 , white for 0 , and red for 1 . This is the matrix that allows us to obtain the recurrence microstates from the recurrence correlations. On the right, its inverse with the following color scheme: white for 0 and black for 1 . This is the matrix that allows us to obtain the recurrence correlations from the recurrence microstates. The axes labels represent the elements of which vectors are associated with each row or column of the matrices.

The \mathcal{QR} factorization for Λ^{-3} proves the appearance of this famous fractal (Fig. 4). This fractal structure arises in logical conjunctions between variables organized in a lexicographical manner, exactly as we have in the current problem, where the elements of either of the vectors $\mathbf{p}^{(3)}$ and $\mathbf{c}^{(3)}$ have a definitive ordering. We can, however, find a way of generating directly the fractal, thus avoiding having to perform this factorization. This is briefly explored in the next subsection.

Optimal ordering for the vector of recurrence correlations

Here, we briefly find the ordering of the vectors $\mathbf{c}^{(k)}$ such that we do not need to perform a \mathcal{QR} decomposition, and just

consider Λ^{-k} as the Sierpiński gasket projections $\mathcal{R}^{(k)}$ from the very beginning.

From the definition of this type of matrix factorization, $\mathcal{Q}^{(k)}$ are orthogonal (or unitary, depending of the matrix being decomposed). In our case, they are just permutation matrices. This implies that their inverses are just their transposes, i.e., they are still just permutation matrices. Therefore, the following expression is valid:

$$\mathcal{Q}^{-k} \mathbf{c}^{(k)} = \mathcal{R}^{(k)} \mathbf{p}^{(k)}, \quad (18)$$

providing us with a way to automatically retrieve the Sierpiński triangle by shuffling the elements of $\mathbf{c}^{(k)}$ using a

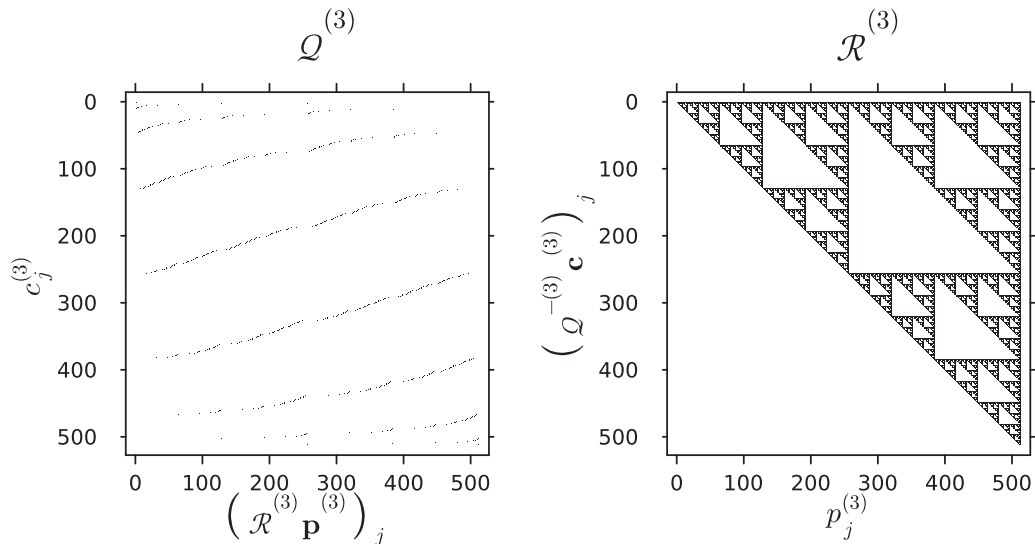


FIG. 4. \mathcal{QR} decomposition of the matrix Λ^{-3} . White cells represent 0 and black cells represent 1 . We see that the orthogonal component \mathcal{Q} is just a permutation matrix, showcasing which rows of \mathcal{R} we should swap to return to Λ^{-3} . It is also visible that the pattern on the component \mathcal{R} is the Sierpiński triangle. The notation for the inverse is maintained for $[\mathcal{Q}^{(3)}]^{-1} = \mathcal{Q}^{-3}$ for clarity. As all $\mathcal{Q}^{(k)}$ are orthogonal, their inverses are just their transposes. The axes labels represent the elements of which vectors are associated with each row or column of the matrices. This is not surprising since it is already known that this pattern arises during bit-wise operations—more specifically, with the logical conjunction. As this matrix basically contains the information on which densities we must combine to form a microstate, the appearance of the Sierpiński triangle is expected.

TABLE II. Relationships between the ordering of the recurrence correlations and the ordering of recurrence microstates for the optimal generation of the fractal matrix $\Lambda^{(k)}$ without the need for any type of matrix decomposition such that $\Lambda^{(k)} = \mathcal{R}^{(k)}$. The ξ are unknowns that can be *independently* either 0 or 1, and the numbers 1 are located in the same positions as the indices on the averages—from (i, j) corresponding to “top left” all the way down to $(i + 1, j + 1)$ corresponding to “bottom right.”

Optimal ordering				
Recurrence Correlation	Matrix Configuration	Recurrence Microstate	Binary Equivalent	Final Index
1	$\begin{pmatrix} \xi & \xi \\ \xi & \xi \end{pmatrix}$	$\mathbf{M}_1^{(2)} \doteq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	0000	1
$\langle R_{i+1,j+1} \rangle^{(2)}$	$\begin{pmatrix} \xi & \xi \\ \xi & 1 \end{pmatrix}$	$\mathbf{M}_2^{(2)} \doteq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$	0001	2
\vdots	\vdots	\vdots	\vdots	\vdots
$\langle R_{i,j} R_{i,j+1} R_{i+1,j} \rangle^{(2)}$	$\begin{pmatrix} 1 & 1 \\ 1 & \xi \end{pmatrix}$	$\mathbf{M}_{15}^{(2)} \doteq \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$	1110	15
$\langle R_{i,j} R_{i,j+1} R_{i+1,j} R_{i+1,j+1} \rangle^{(2)}$	$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$\mathbf{M}_{16}^{(2)} \doteq \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	1111	16

certain family of permutations. This new ordering of $\mathbf{c}^{(k)}$ is indeed very basic, we just need to apply the same ordering scheme as we do for microstates: a binary ordering.

We must remember that each pair of indices represent a position within a microstate, while the element 1 (first element of every $\mathbf{c}^{(k)}$) is not related to any location, since it has no index. The positions identified by the indices can be interpreted as a binary string and, by pairing each correlation to their equivalent microstates, we do not need to perform the \mathcal{QR} decomposition and $\Lambda^{-(k)}$ is already equal to $\mathcal{R}^{(k)}$ (the fractal) right from the start. Let us explain by using the $k = 2$ case as an example (16 microstates and correlations) for clarification. The pairing should follow the arrangement shown in Table II.

This ensures the fractal relationship for every microstate size k without any further factorization. It is also possible to justify why we can alternate between the microstates description of the RP and the recurrence correlations description with Table II. It is clear from there that there is a one-to-one correspondence between the elements of both sets, since both of them can be mapped into binary sequences.

A final consequence worth mentioning is that, with the results in Table II, we readily know which microstates have to be combined to generate any recurrence correlation. This can be achieved in the following manner: First, consider the average on the second row, $\langle R_{i+1,j+1} \rangle^{(2)}$. Now, fix the associated position (in this case, the bottom-right corner) and sum the probabilities of *every* microstate with a recurrence in that position. For this specific average we would sum the probabilities of eight microstates, since we have three empty positions and $2^3 = 8$. That is why we need all microstates for the element 1, since its binary equivalent has four empty locations, leading to $2^4 = 16$ microstates. For the recurrence correlation with all indices, there are no empty locations (last row of Table II), resulting in only $2^0 = 1$ microstate. In summary, if a recurrence correlation forms a specific binary pattern as shown in Table II, we just have to sum the probability of every microstate that displays the same pattern to calculate its numerical value.

From this point on we will replace $\Lambda^{-(k)} \rightarrow \mathcal{R}^{(k)}$.

III. RECURRENCE MICROSTATES AND LINE STRUCTURES

In the RP, the number of diagonal lines of length ℓ is given by [6]

$$P_\ell(\ell) = \sum_{i,j=1}^{N-\ell-1} (1 - R_{i,j})(1 - R_{i+\ell+1,j+\ell+1}) \prod_{\ell'=1}^{\ell} R_{i+\ell',j+\ell'}, \quad (19)$$

and its equivalent in terms of verticals (or horizontals) is

$$P_\nu(\nu) \approx \sum_{i,j=1}^{N-\nu-1} (1 - R_{i,j})(1 - R_{i,j+\nu+1}) \prod_{\nu'=1}^{\nu} R_{i,j+\nu'}. \quad (20)$$

The approximate sign (\approx) is used for vertical lines because the summation runs as $\sum_{i=1}^N \sum_{j=1}^{N-\nu-1} [\cdot]$. However, we neglect (with also negligible loss) the last $\nu + 1$ rows by summing as $\sum_{i,j=1}^{N-\nu-1} [\cdot]$. This also has the structure of a set of certain recurrence correlations, except it is lacking a normalization factor $(N - \ell - 1)^2$ or $(N - \nu - 1)^2$. By including this, we always end up with four recurrence correlations, regardless of the values of ℓ and ν . Of course, these sets of four correlations are different for diagonals and verticals.

For the trivial case of $\ell = \nu = 1$, we have for the diagonal case

$$\begin{aligned} P_\ell(1) &= \sum_{i,j=1}^{N-2} (1 - R_{i,j})(1 - R_{i+2,j+2}) R_{i+1,j+1} \\ &= \sum_{i,j=1}^{N-2} (1 - R_{i,j} - R_{i+2,j+2} + R_{i,j} R_{i+2,j+2}) R_{i+1,j+1}, \end{aligned} \quad (21)$$

which, by distributing the right term into the parentheses, dividing both sides by the normalization factor $(N - 2)^2$,

and using the compact notation from Eq. (7), becomes (it is analogous for the verticals)

$$\begin{aligned} \frac{P_\ell(1)}{(N-2)^2} &= \langle R_{i+1,j+1} \rangle^{(3)} - \langle R_{i,j} R_{i+1,j+1} \rangle^{(3)} \\ &\quad - \langle R_{i+1,j+1} R_{i+2,j+2} \rangle^{(3)} \\ &\quad + \langle R_{i,j} R_{i+1,j+1} R_{i+2,j+2} \rangle^{(3)}, \\ \frac{P_\nu(1)}{(N-2)^2} &\approx \langle R_{i,j+1} \rangle^{(3)} - \langle R_{i,j} R_{i,j+1} \rangle^{(3)} - \langle R_{i,j+1} R_{i,j+2} \rangle^{(3)} \\ &\quad + \langle R_{i,j} R_{i,j+1} R_{i,j+2} \rangle^{(3)}. \end{aligned} \quad (22)$$

To model this, we can think of very sparse vectors $\mathbf{d}^{(\ell)}$ and $\mathbf{v}^{(\nu)}$ with the same number of elements as $\mathbf{p}^{(k)}$, but with mostly 0's, *except* on the indices of the correlations that appear by expanding the expression for the lines [Eqs. (19) and (20)]. This setup allows us to select the required correlations. In those indices, there will either be the elements 1 or -1 . The signs just represent that the corresponding elements are multiplied either by 1 or -1 [as in Eq. (22)]. Then, we can write

$$\begin{aligned} \frac{P_\ell(\ell)}{(N-\ell-1)^2} &= \mathbf{d}^{(\ell)} \cdot \mathcal{R}^{(\ell+2)} \mathbf{p}^{(\ell+2)}, \\ \frac{P_\nu(\nu)}{(N-\nu-1)^2} &\approx \mathbf{v}^{(\nu)} \cdot \mathcal{R}^{(\nu+2)} \mathbf{p}^{(\nu+2)}. \end{aligned} \quad (23)$$

With this, for lines of length ℓ or ν , we would need microstates of length $\ell+2$ or $\nu+2$ to describe them. We can circumvent this problem for the case of the RQA measures DET and LAM, given by

$$\begin{aligned} \text{DET} &= \frac{\sum_{\ell=\ell_{\min}}^N \ell P_\ell(\ell)}{\sum_{\ell=1}^N \ell P_\ell(\ell)}, \\ \text{LAM} &= \frac{\sum_{\nu=\nu_{\min}}^N \nu P_\nu(\nu)}{\sum_{\nu=1}^N \nu P_\nu(\nu)}. \end{aligned} \quad (24)$$

We shift the summation from the longer lines to the shorter ones, by using the fact that the denominator is just the total number of recurrences, given by $N^2 \text{RR}$ and separating the numerator into two sums, one till $\ell_{\min}-1$ or $\nu_{\min}-1$ and the other after. With this, we write these measures entirely in terms of lines shorter than the minimum line length as in

$$\begin{aligned} \text{DET} &= 1 - \frac{1}{N^2 \text{RR}} \sum_{\ell=1}^{\ell_{\min}-1} \ell P_\ell(\ell), \\ \text{LAM} &= 1 - \frac{1}{N^2 \text{RR}} \sum_{\nu=1}^{\nu_{\min}-1} \nu P_\nu(\nu). \end{aligned} \quad (25)$$

As a consequence, we successfully require only microstates up to size $\ell_{\min}+1$ and $\nu_{\min}+1$. For the usual value of these parameters, which is 2, we just have to sample 3×3 microstates, which is completely feasible. For this case, we

have

$$\begin{aligned} \text{DET} &\approx 1 - \frac{\mathbf{d}^{(1)} \cdot \mathcal{R}^{(3)} \mathbf{p}^{(3)}}{\text{RR}}, \\ \text{LAM} &\approx 1 - \frac{\mathbf{v}^{(1)} \cdot \mathcal{R}^{(3)} \mathbf{p}^{(3)}}{\text{RR}}, \end{aligned} \quad (26)$$

assuming $(N-2)^2/N^2 \approx 1$. The recurrence rate RR can also be obtained from the same microstate distribution $\mathbf{p}^{(3)}$ via the inner (scalar) product

$$\text{RR} \approx \mathbf{rr}^{(3)} \cdot \mathbf{p}^{(3)}, \quad (27)$$

with $\mathbf{rr}^{(3)}$ being the vector of the individual recurrence rates for each of the 3×3 microstates, so that this inner product is just an average of local densities over all sampled microstates [29]. The microstates in Eq. (4) provide useful explicit examples for illustration. Their individual recurrence rates—and hence the respective elements of the vectors $\mathbf{rr}^{(1)}$ and $\mathbf{rr}^{(2)}$ —are $rr_1^{(1)} = 0$, $rr_2^{(1)} = 1$, $rr_1^{(2)} = 0$, $rr_9^{(2)} = 0.25$, and $rr_{16}^{(2)} = 1$.

By performing the sampling process of recurrence microstates as small as 3×3 just once, we can simultaneously calculate the most important recurrence quantifiers, RR, DET, and LAM. The only requirement is that we must know eight rows of the matrix $\mathcal{R}^{(3)}$ beforehand, four of them for DET and the remaining four for LAM. This matrix, moreover, is independent of anything but the size of the microstates. Thus, for fixed-sized 3×3 microstates, the matrix is also fixed, allowing us to calculate it beforehand, which prevents us from having to calculate the matrices in Fig. 3 over and over again.

IV. NUMERICAL ANALYSIS

In this section we perform a series of comparisons between our approximation method and the recurrence measures calculated using lines. The following paradigmatic systems are used for this: (1) uniform white noise (UWN), (2) logistic map (with $r = 4$), (3) x component of the Lorenz system with parameters $(\sigma, b, r) = (10, 8/3, 28)$ [30] (with total time being 100 units, providing a time step of ~ 0.1 , comparable to real-world data resolution), and (4) a harmonic oscillator with five oscillations. All time series (x) are initially mapped to the unit interval ($x \in [0, 1]$) for simplicity, without loss of generality. All series have $2^{10} = 1024$ data points, also without loss of generality.

For each system, we calculate the recurrence measures DET and LAM using the conventional method and the microstates approximation provided above. The range of the threshold parameter ε for this analysis is from 1% to 100% of the series' maximum distances. Following Ref. [19], we select only a fraction of possible microstates, such that only 5% of the total number of microstates are used for the calculation. We test a similar approach for the standard (line-based) calculation by using only a fraction of the diagonal (vertical) lines through a bootstrapping procedure described in Refs. [31,32]. This procedure allows us to better evaluate the improvements by using the microstates approach by ensuring our results are within the correct statistical bounds. Thus, we consider ten different fractions of the diagonal (vertical) lines: from 10% to 90% of the average number of lines in the bootstrapped

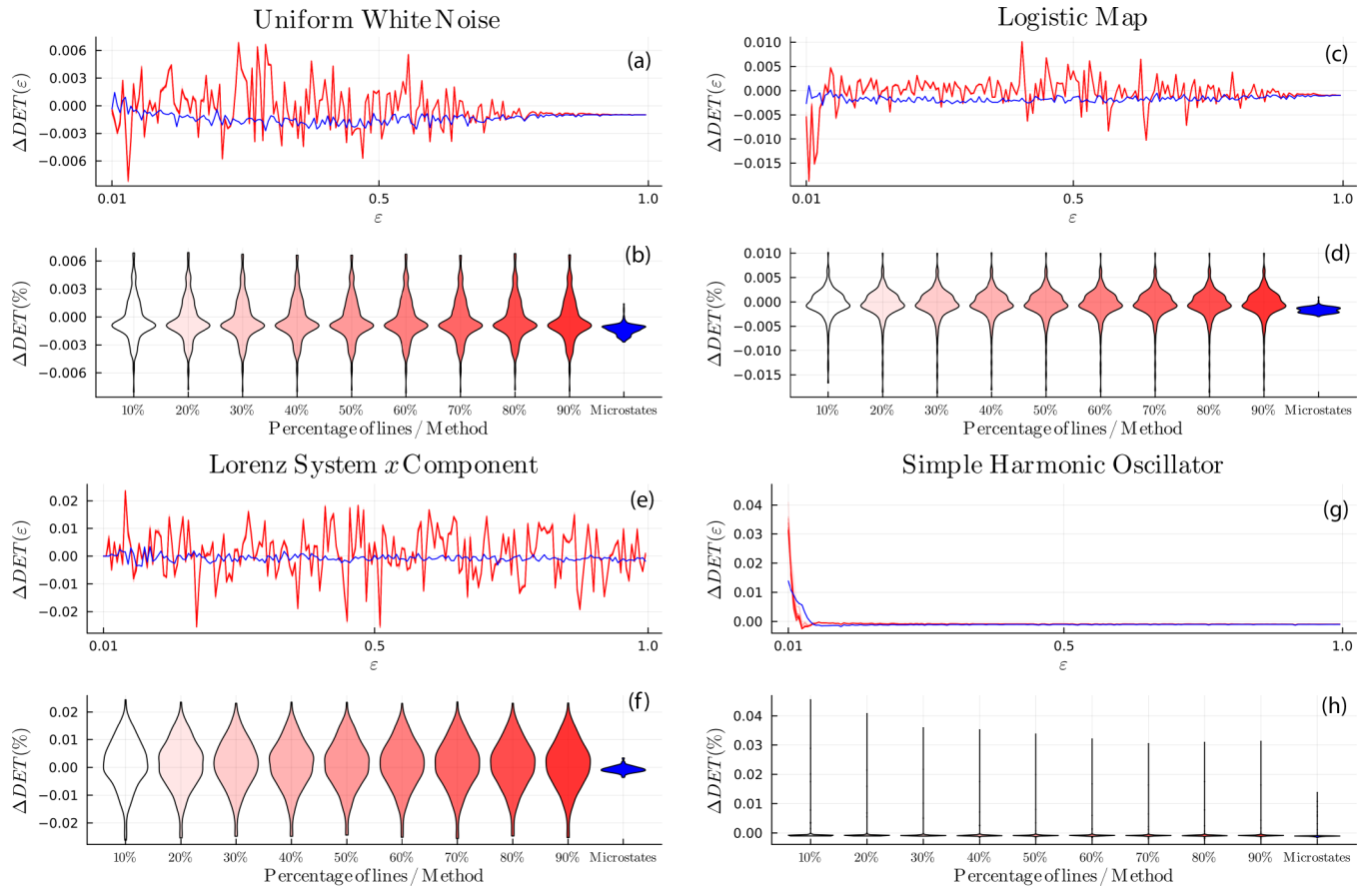


FIG. 5. Differences between the ground truth determinism calculated from different fractions of lines (shades of red) and from recurrence microstates (blue). The fractions are 10%–90% of the average number of lines in the windows they were calculated from (through a bootstrapping procedure) in steps of 10% as a function of the threshold ϵ [curves on top panels—(a), (c), (e), and (g)], as well as their violin plots as a function of the percentage of lines [bottom panels—(b), (d), (f), and (h)]. The window size is half the entire series, being 512 data points, while all the series have 1024 points, with window step of 32 points, generating a total of 17 windows. The systems were chosen such that they have different dynamics to showcase the dynamics-independent nature of the method: (1) the uniform white noise (UWN), (2) logistic map with $r = 4.0$ (in a chaotic region), (3) the x component of the Lorenz system (also in a chaotic region), and (4) for a simple harmonic oscillator. All curves were calculated for $1\% \leq \epsilon \leq 100\%$ of the series diameter.

windows, and 100%, which is considered as the “ground truth” result to which we compare the approximations. We further consider the differences Δ between ground truth DET (LAM) and the microstates approximation and line-based approximation (Figs. 5 and 6).

The bottom panels (b, d, f, h) of Fig. 5, in red, show the *violin plots* (a way of showcasing the probability density functions for different variables in the same figure) of the cumulative errors between the DET calculated from the bootstrapped windows and a certain percentage of lines (from 10% to 90% of the average number of lines) and the ground truth quantifier. They appear to have absolute values consistently below 0.02 for all systems tested. In blue, there are the errors using recurrence microstates. In all cases, it is consistently lower than the alternatives. Figure 6 shows the results for LAM (green). They also have absolute values mostly below 0.02 for lines and significantly less for microstates. In both figures, the top panels (a, c, e, g) represent the errors as functions of the recurrence threshold ϵ for all percentages of lines and microstates. This is strong evidence that microstates, as a tool, are a viable option for these calculations.

The bootstrapping procedure for the lines is the reason for the shapes of the violin plots in Figs. 5 and 6, since we are using the average number of lines of the windows for the calculations. If instead we simply sampled the percentages, we would need less lines to match the microstates error. This also explains why they are all similar (but not equal). However, in order to know the exact percentage of lines, we would still need to scan the full RP, since we do not know how many lines are there without counting all of them first. We do know, however, the number of microstates already from the start, so there is no need to scan the full RP in our approach.

We find that the errors for every system—regardless of its dynamics—and for each fraction of sampled lines are very small, generally the order $\mathcal{O}(10^{-2})$ (Figs. 5 and 6). However, the microstates approximation tends to overestimate the measures slightly, given that the errors are skewed toward negative values $\Delta < 0$. The reason why errors are so small is understandable, since this approximation consisted mostly of algebraic manipulations, with the few assumptions made being $(N - 2)^2/N^2 \approx 1$ (which was not even needed, since this was just for simplifying the expression), and that $RR \neq 0$.

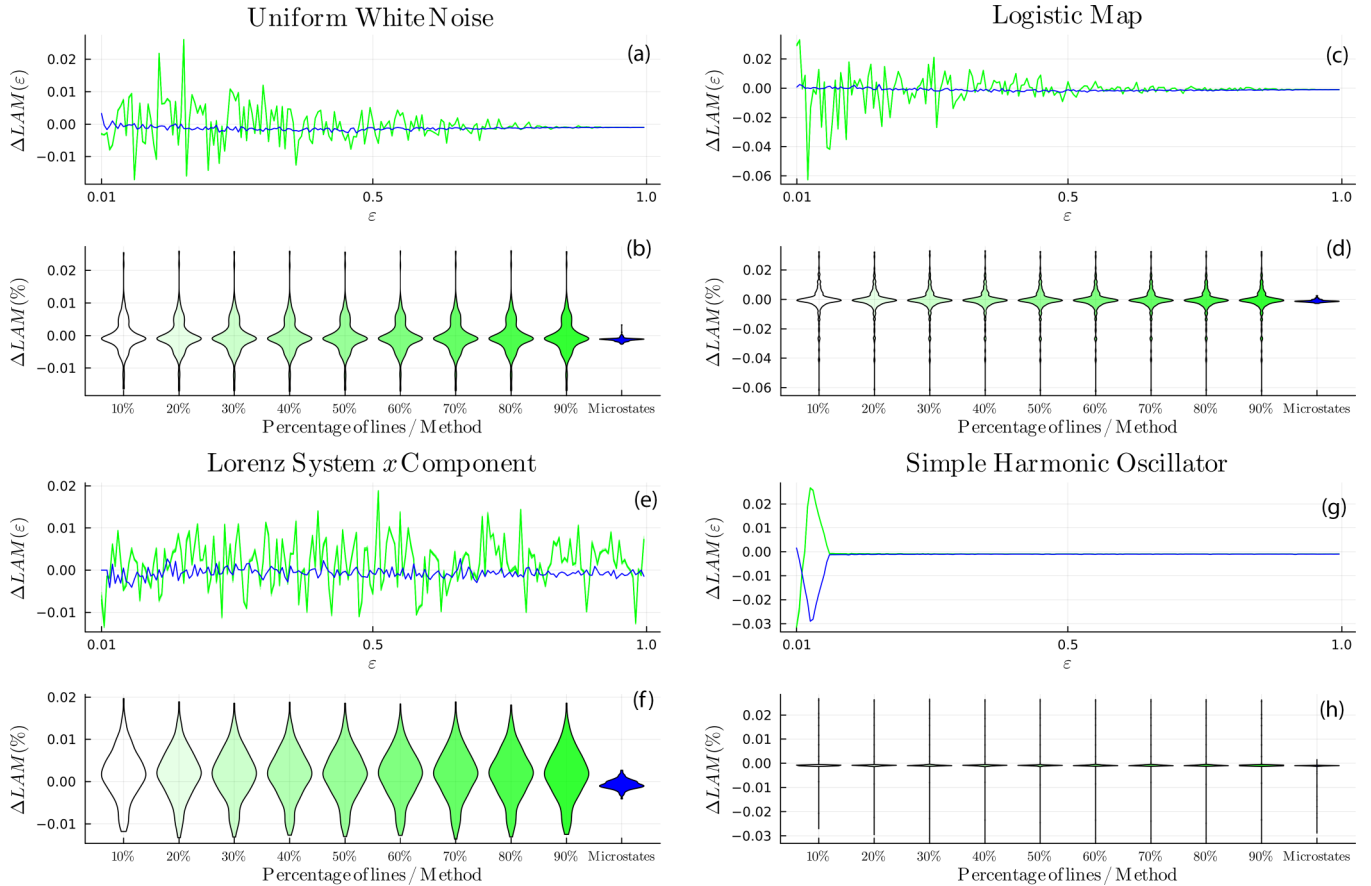


FIG. 6. Same as Fig. 5 but for laminarity (LAM) (shades of green).

In practice, we might encounter numerical problems for very small thresholds if we are not careful enough. This happens because we expect the RP not to have any diagonals when $\varepsilon \ll 1$, so we must avoid sampling points within the LOI, since it would induce a larger number of diagonals, which will affect the final DET. However, this does not happen with LAM, which maintains small errors even for $\varepsilon \rightarrow 0$, since the LOI is only N verticals with $\nu = 1$.

As mentioned in previous works [19], only about 5% of the total number of microstates were used, which, for a series with approximately 1000 data points, is already sufficient for the sampled distribution to be representative of the correct one. This occurs because, in a recurrence plot of this size, the number of 3×3 microstates that can be sampled is of the order $\mathcal{O}(10^6)$. However, there are only 512 distinguishable microstates, so sampling 5% [$\mathcal{O}(10^5)$] still provides a sufficient sample size. With this in mind, if we consider equally optimized algorithms for generating a RP and sampling specific parts of it (such as microstates), this approximation could require significantly fewer steps compared to sampling traditional line distributions, as only a single distribution would be needed.

One important advantage of the microstates approach against sampling random lines is that, for the lines, we would still have to navigate a large portion of the RP to actively search for them, which is not a problem for our method—because we need complete lines, not fraction of lines. With microstates, we do not need to worry about typical problems

like where we are in the RP (starting point), to which direction we are going, or if the neighbor elements are equal to the element at our current location. The only task is to count the amount of a few small randomly sampled squares and apply a predefined matrix multiplication to the resulting probability vector.

Concerning numerical complexity, for raw calculations over all lines and microstates, the overall cost for lines is about $\mathcal{O}(N^2)$ and for $k \times k$ microstates is $\mathcal{O}(k^2 N^2)$. However, both approaches can benefit from optimization due to the existence of symmetries, and this is accentuated on the microstates, given that the invariance under transposition of the RP is not their only source of symmetries. Moreover, efforts to optimize algorithms for efficient sampling and processing of recurrence microstates are already underway.

V. CONCLUSION

In this work we provide an alternative way of calculating two important RQA measures, DET and LAM, based on the sampling of just one single distribution of 3×3 microstates, given the minimum line length of two points. The findings discussed here integrate the concept of recurrence microstates into the one of computation of RQA measures. They also contribute to the mathematical understanding of how smaller local structures in the RP are related to each other.

Although it is an approximation, the errors encountered are very small, being most of the time negligible, for the

recurrence threshold $\varepsilon \geq 1\%$ of the series diameter (maximum pairwise distance). The results are useful in several different scenarios, mainly when large RP are analyzed, given that we do not have to scan all the RP repeatedly. Instead of generating separate distributions for diagonal and vertical lines, we only sample small random parts of it to obtain a single distribution.

We show that the recurrence microstate analysis (RMA) not only provides new quantifiers, such as microstate entropy, but also proves to be an alternative method for calculating traditional quantifiers based on lines (diagonal and vertical). In this way, we present a generalization of RP line sampling algorithms based on recurrence microstates.

This method, albeit promising, depends on the efficient computation of recurrence microstates. Moreover, its strength will be increasingly more expressive as the number of quantifiers we can calculate with a single microstates distribution becomes larger, as well as more computational optimizations and conceptual generalizations are proposed. *Therefore, this work serves both as a theoretical advance and to provide us with a solid reason for pursuing computational optimization in this area.*

Although several research problems still need to be addressed in recurrence microstates analysis—such as improving computational efficiency for large microstates, developing new microstate-based measures, investigating scaling behaviors, and deriving traditional quantifiers from microstates—our findings represent an important advancement. Our approach addresses the current challenge of having to generate the full RP for calculating RQA measures, as microstate-based analyses allow us to skip this step. We hope this work will open new avenues for efficient computations and contribute to the ongoing development of new techniques within the scope of recurrence microstates analysis.

ACKNOWLEDGMENTS

We greatly thank Andreas Amann for asking important questions that lead to some of the results of this work, which ended up improving the overall quality of the manuscript. This work is supported by the Brazilian research agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Grants No. 308441/2021-4, No. 307907/2019-8, No. 407072/2022-5, No. 305189/2022-0, No. 408254/2022-0, and No. 300064/2023-3; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Grants No. 88881.895032/2023-01 and No. 88887.935598/2024-00; and DAAD PPP Grant No. 57705568.

APPENDIX: FORMAL PROOF OF EQ. (6)

This is a direct formal proof of the expression for the probability of an arbitrary microstate, given by Eq. (6).

Definition A1 (Logical Operations).

(i) *Logical equivalence.* ($X \equiv Y$) is a function returning 1 if $X = Y$ and 0 otherwise. For binary inputs $X, Y \in \{0, 1\}$,

$$(X \equiv Y) = (-1)^{\bar{Y}}(X - \bar{Y}), \quad \text{where } \bar{Y} := 1 - Y.$$

(ii) *Logical AND.* Defined for $X, Y \in \{0, 1\}$ as

$$X \wedge Y := XY,$$

where $X \wedge Y = 1$ if $X = Y = 1$ and 0 otherwise.

Definition A2 (Recurrence Matrices and Microstates).

Let $\Omega \subset \{0, 1\}^{N \times N}$ denote the set of all possible recurrence matrices, with $\mathbf{R} \in \Omega$ being a recurrence matrix. A $k \times k$ **submatrix** of \mathbf{R} , starting at (i, j) , is defined as

$$\mathbf{R}_{i,j}^{(k)} := (R_{i+\mu, j+\nu})_{0 \leq \mu, \nu \leq k-1},$$

where $1 \leq i, j \leq N - k + 1$ and $1 \leq k \leq N$. The **recurrence microstates** of size $k \times k$ are elements of the set $\Omega^{(k)} := \{0, 1\}^{k \times k}$, containing all 2^{k^2} possible logical matrices. The β th recurrence microstate is denoted as

$$\mathbf{M}_{\beta}^{(k)} := (b_{1+\mu, 1+\nu})_{0 \leq \mu, \nu \leq k-1}.$$

Definition A3 (Measure Space of Submatrices). For a given recurrence matrix \mathbf{R} , define

(i) $\Omega^{(k)}(\mathbf{R})$ as the set of all $(N - k + 1)^2$ submatrices of \mathbf{R} , i.e.,

$$\Omega^{(k)}(\mathbf{R}) := \{\mathbf{R}_{i,j}^{(k)} \mid 1 \leq i, j \leq N - k + 1\};$$

(ii) $\Omega_{\beta}^{(k)}(\mathbf{R})$ as the subset of $\Omega^{(k)}(\mathbf{R})$ that matches the pattern of the microstate $\mathbf{M}_{\beta}^{(k)}$, given by

$$\Omega_{\beta}^{(k)}(\mathbf{R}) := \{\mathbf{R}_{i,j}^{(k)} \mid (\mathbf{R}_{i,j}^{(k)} \equiv \mathbf{M}_{\beta}^{(k)}) = 1\}.$$

The tuple $(\Omega^{(k)}(\mathbf{R}), \mathcal{F}, \omega)$ forms a measure space, where $\mathcal{F} = \mathcal{P}(\Omega^{(k)}(\mathbf{R}))$ is the power set, and the counting measure ω satisfies

$$\omega(A) := |A|, \quad \forall A \in \mathcal{F},$$

with $\omega(\Omega^{(k)}(\mathbf{R})) = (N - k + 1)^2$.

Proposition A1. The probability of finding the β th $k \times k$ recurrence microstate within a $N \times N$ recurrence matrix is given by Eq. (6).

Proof. We want a normalized $\omega(\Omega_{\beta}^{(k)}(\mathbf{R}))$. Since from Definition A3 it requires comparisons between matrices, each of their elements must match individually, i.e.,

$$\begin{aligned} (\mathbf{R}_{i,j}^{(k)} \equiv \mathbf{M}_{\beta}^{(k)}) &= (R_{i,j} \equiv b_{1,1}) \wedge \cdots \wedge (R_{i+k-1, j+k-1} \equiv b_{k,k}) \\ &= \prod_{\mu, \nu=0}^{k-1} (R_{i+\mu, j+\nu} \equiv b_{1+\mu, 1+\nu}), \end{aligned}$$

where we used Definition A1 on (\wedge) . Since $\{R\}$ and $\{b\}$ are binary values, then applying Definition A1 on (\equiv) yields

$$(\mathbf{R}_{i,j}^{(k)} \equiv \mathbf{M}_{\beta}^{(k)}) = \prod_{\mu, \nu=0}^{k-1} (-1)^{\bar{b}_{1+\mu, 1+\nu}} (R_{i+\mu, j+\nu} - \bar{b}_{1+\mu, 1+\nu}). \quad (\text{A1})$$

Since $\Omega_{\beta}^{(k)}(\mathbf{R}) \subseteq \Omega^{(k)}(\mathbf{R}) \in \mathcal{F}$, we can apply the measure ω as [respecting the constraints on (i, j) in Definition A2]

$$\omega(\Omega_{\beta}^{(k)}(\mathbf{R})) = \sum_{i, j=1}^{N-k+1} (\mathbf{R}_{i,j}^{(k)} \equiv \mathbf{M}_{\beta}^{(k)}). \quad (\text{A2})$$

The probability of a particular microstate is defined as the previous quantity normalized by the measure of all found

$k \times k$ microstates, i.e.,

$$p_{\beta}^{(k)} := \frac{\omega(\Omega_{\beta}^{(k)}(\mathbf{R}))}{\omega(\bigcup_{\beta} \Omega_{\beta}^{(k)}(\mathbf{R}))}. \quad (\text{A3})$$

However, $\Omega^{(k)}(\mathbf{R})$ is partitioned disjointly into microstates, so $\omega(\bigcup_{\beta} \Omega_{\beta}^{(k)}(\mathbf{R})) = (N - k + 1)^2$, since each position $\mathbf{R}_{i,j}^{(k)}$ can only hold a single pattern $\mathbf{M}_{\beta}^{(k)}$. Therefore, from the prior argument and Eqs. (A1)–(A3), we prove Eq. (6). ■

-
- [1] C. Torrence and G. Compo, A practical guide to wavelet analysis, *Bull. Am. Meteorol. Soc.* **79**, 61 (1998).
- [2] T. Schreiber and A. Schmitz, Surrogate time series, *Physica D* **142**, 346 (2000).
- [3] C. Bandt, Ordinal time series analysis, *Ecol. Modell.* **182**, 229 (2005).
- [4] Y. Hirata and J. Amigó, A review of symbolic dynamics and symbolic reconstruction of dynamical systems, *Chaos Interdiscip. J. Nonlinear Sci.* **33**, 052101 (2023).
- [5] C. Bandt and B. Pompe, Permutation entropy: a natural complexity measure for time series, *Phys. Rev. Lett.* **88**, 174102 (2002).
- [6] N. Marwan, M. Romano, M. Thiel, and J. Kurths, Recurrence plots for the analysis of complex systems, *Phys. Rep.* **438**, 237 (2007).
- [7] J. H. Poincaré, Sur le problème des trois corps et les équations de la dynamique, *Acta Math.* **13**, 1 (1890).
- [8] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, Recurrence plots of dynamical systems, *Europhys. Lett.* **4**, 973 (1987).
- [9] N. Marwan, A historical review of recurrence plots, *Eur. Phys. J. Spec. Top.* **164**, 3 (2008).
- [10] N. Marwan and K. Kraemer, Trends in recurrence analysis of dynamical systems, *Eur. Phys. J.: Spec. Top.* **232**, 5 (2023).
- [11] N. Marwan, *Recurrence Quantification Analysis* (Springer International Publishing, Cham, 2016).
- [12] J. P. Zbilut and C. L. Webber, Embeddings and delays as derived from quantification of recurrence plots, *Phys. Lett. A* **171**, 199 (1992).
- [13] C. Webber and J. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies, *J. Appl. Physiol.* **76**, 965 (1994).
- [14] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, Recurrence-plot-based measures of complexity and their application to heart-rate-variability data, *Phys. Rev. E* **66**, 026702 (2002).
- [15] L. Trulla, A. Giuliani, J. Zbilut, and C. Webber, Recurrence quantification analysis of the logistic equation with transients, *Phys. Lett. A* **223**, 255 (1996).
- [16] D. Schultz, S. Spiegel, N. Marwan, and S. Albayrak, Approximation of diagonal line based measures in recurrence quantification analysis, *Phys. Lett. A* **379**, 997 (2015).
- [17] T. Rawald, M. Sips, and N. Marwan, PyRQA – conducting recurrence quantification analysis on very long time series efficiently, *Comput. Geosci.* **104**, 101 (2017).
- [18] T. Prado, B. Boaretto, G. Corso, G. Santos Lima, J. Kurths, and S. Lopes, A direct method to detect deterministic and stochastic properties of data, *New J. Phys.* **24**, 033027 (2022).
- [19] G. Corso, T. Prado, G. Lima, J. Kurths, and S. Lopes, Quantifying entropy using recurrence matrix microstates, *Chaos Interdiscip. J. Nonlinear Sci.* **28**, 083108 (2018).
- [20] S. Lopes, T. Prado, G. Corso, S. G. Lima, and J. Kurths, Parameter-free quantification of stochastic and chaotic signals, *Chaos, Solitons & Fractals* **133**, 109616 (2020).
- [21] T. Prado, E. Macau, and S. Lopes, Detection of data corruption in stationary time series using recurrence microstates probabilities, *Eur. Phys. J.: Spec. Top.* **230**, 2737 (2021).
- [22] G. You and Y. Ke, ENRM: An alternative tool for studying dynamical systems, *Chaos, Solitons & Fractals* **174**, 113889 (2023).
- [23] R. Donner, Y. Zou, J. Donges, N. Marwan, and J. Kurths, Recurrence networks—a novel paradigm for nonlinear time series analysis, *New J. Phys.* **12**, 033025 (2010).
- [24] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Network Motifs: Simple building blocks of complex networks, *Science* **298**, 824 (2002).
- [25] Y. Hirata and M. Shiro, Recurrence plots bridge deterministic systems and stochastic systems topologically and measure-theoretically, *Chaos* **33**, 8 (2023).
- [26] T. Prado, G. Corso, G. Santos Lima, R. Budzinski, B. Boaretto, F. Ferrari, E. Macau, and S. Lopes, Maximum entropy principle in recurrence plot analysis on stochastic and chaotic systems, *Chaos* **30**, 043123 (2020).
- [27] K. Falconer, *Fractal Geometry* (John Wiley & Sons, New York, 2003).
- [28] L. Trefethen and D. Bau, *Numerical Linear Algebra* (Society for Industrial and Applied Mathematics, Philadelphia, 1997).
- [29] T. Lima Prado, V. Machado, G. Corso, G. Santos Lima, and S. Lopes, How to compute suitable vicinity parameter and sampling time of recurrence analysis, *Nonlinear Dyn.* **112**, 1141 (2023).
- [30] E. Lorenz, Deterministic nonperiodic flow, *J. Atmos. Sci.* **20**, 130 (1963).
- [31] N. Marwan, S. Schinkel, and J. Kurths, Significance for a recurrence based transition analysis, *Proceedings of International Symposium on Nonlinear Theory and its Applications (NOLTA'08)* (Academia, 2008), pp. 412–415.
- [32] S. Schinkel, N. Marwan, O. Dimigen, and J. Kurths, Confidence bounds of recurrence-based complexity measures, *Phys. Lett. A* **373**, 2245 (2009).