

Discriminating disorders and their complexity

Jean Sire Armand Eyebe Fouda,¹ Norbert Marwan,² Jürgen Kurths,² and Wolfram Koepf³

¹⁾*Department of Physics, University of Yaoundé I, P.O. Box 812 Yaoundé, Cameroon.*^{a)}

²⁾*Potsdam Institute for Climate Impact Research, Member of the Leibniz Association, P.O. Box 601203, Potsdam, D-14412, Germany*

³⁾*Institute of Mathematics, University of Kassel, Heinrich-Plett-Str. 40, D-34132 Kassel, Germany.*

(*Electronic mail: koepf@mathematik.uni-kassel.de)

(*Electronic mail: kurths@pik-potsdam.de)

(*Electronic mail: marwan@pik-potsdam.de)

(*Electronic mail: fouda@mathematik.uni-kassel.de)

(Dated: 25 August 2026)

Knowledge of the dynamics nature is essential for efficient system modeling, whereas the assessment of its complexity is useful for quantifying predictability. Depending on the nature of the information source, complexity can be ordered or disordered, although most existing methods do not specify this nature. In this paper, we present a novel determinism index as an efficient measure to determine the nature of data through the complexity of the ordinal pattern slope sum (sCOPSS). Combining the above index with the sCOPSS yields a comprehensive tool for estimating the complexity and disorder of short observations. Our results provide a negative complexity for deterministic data and a positive complexity for stochastic data and mixtures of stochastic and deterministic data, thus allowing identification among them. Mixing chaotic and stochastic data reveals a resonance phenomenon in sCOPSS that allows further reduction of the requirement for data length and improves the separability of the method. Applications to real-world data are given for the identification of audio signals and the classification of ECG beats. The results obtained confirm the effectiveness of the signed complexity for discriminating disorders and their complexity from short time series, thus separating sources in mixed data.

The choice of encoding method is crucial in defining efficient complexity measures. Choosing a coding procedure with known physical meaning or adapted to the dynamics under investigation ensures a good assessment of the complexity. We introduced the use of ordinal pattern slopes because they are directly related to the period of the dynamics under study. This approach has allowed us to define efficient complexity measures, enabling us to distinguish between different types of dynamics. In the same vein, we present the sCOPSS approach to determine both the nature of the dynamics and its complexity. It applies even to short-length data and optimizes the computation time and the implementation cost.

Complexity measures are essential tools for quantifying disordered data. They were originally developed to distinguish between types of deterministic chaos and time-correlated or uncorrelated stochastic signals. An example is the circulance measure that quantifies irregularity based on the topology of Ordinal Pattern Transition Networks (OPTNs) and positions a system on a continuous spectrum from regularity to stochasticity¹.

Most complexity measures are based on information entropy^{2–8} evaluated by encoding the time series into a set of symbols and determining the corresponding probability distribution. Complexity measures usually range between 0

and 1, making it difficult to accurately assess the complexity of all dynamics without a multi-scale consideration, especially in the case of high-dimensional systems and short-length data^{9–11}. Furthermore, the definition of the sole disorder level to distinguish between different natures of time series and assess their complexity is problematic, as sources of different nature could present similar complexity or disorder levels. This ambiguity can be raised using graphical representations such as the complexity-entropy plane¹², but this approach also complicates the interpretation of the temporal analysis of the disorder. A recurrence-based microstate analysis method was presented as an alternative for direct analysis, without requiring free parameters¹¹. However, a range of the recurrence threshold should be tested with adequate resolution to optimally predict the disorder. The authors claim that an optimal threshold can be found by the maximum entropy of the microstates. Although the complexity of the deterministic systems investigated with the so-defined method does not exceed 0.8, there could exist systems whose complexity equals stochastic systems, hence the need to distinguish the nature of complexity.

An efficient complexity measure should specify whether the complexity is ordered or disordered, or make a difference between organized and disorganized dynamics¹³. In organized dynamics, it is possible to determine a relationship between the different states of the phase space, while in disorganized systems, such a relationship is difficult to establish. For instance, deterministic tools like recurrence or differential equations efficiently organize dynamics. In contrast, predicting the behavior of disorganized systems strictly re-

^{a)}Also at Institute of Mathematics, University of Kassel, Heinrich-Plett-Str. 40, D-34132 Kassel, Germany.

quires stochastic models based on probability distributions. The transition from organized to disorganized nature is a continuum, and both types of dynamics present a complexity that is a measure of their unpredictability.

In this paper, we make a difference between the complexity and the nature (deterministic or stochastic) of the dynamics. We define disorder as a measure to estimate the nature of the dynamics, to distinguish between ordered and disordered complexity. We introduce the ordinal pattern slope sum (OPSS) encoding technique combined with the multi-lag approach in an OPTN to define the short complexity of OPSS (sCOPSS). The sCOPSS includes two complexity indicators, namely the mean rate r_s of s -type links (groups of $s + 1$ OPSS) and the normalized growth rate λ_s of s -type links. Considering that determinism is reflected in the existence of a relationship between states of the phase space, with limited emergence of new relationships over time, we further propose the determinism index $-1 \leq D_s \leq 1$ based on r_s and λ_s to assess the nature of the dynamics. The depth s of OPTN^{14,15} acts as a separability or scalability factor to increase the distance between dynamics with close complexity. Since an important challenge in experimental system analysis is the limited data length, the proposed approach is validated under short data constraints.

I. SHORT COMPLEXITY OF ORDINAL PATTERN SLOPE SUM: SCOPSS

The sCOPSS method first involves the definition of OPSS patterns.

A. OPSS encoding

Let $\{x_t\}_{t=0,1,\dots,T-1}$ be a time series of length T samples. From this time series embedding vectors $\mathbf{x}_k = (x_{k\tau_0}, x_{k\tau_0+\tau}, \dots, x_{k\tau_0+(n-1)\tau})$, $k, l \in \mathbb{N}$, can be derived. $n \geq 2$ is the embedding dimension, $\tau_0 = 1, 2, \dots$ is the sampling time of the embedding vectors, while $\tau = 1, 2, \dots$ is the delay time of the samples or time lag. By reordering the values in the embedding vectors $\mathbf{x}_k$ in ascending order along with their respective positions, permutations π_k are generated as vectors of the reordered positions. An example is given for $\mathbf{x}_k = (4, 1, 3, 0)$ for which the ascending reordering of $(\frac{4,1,3,0}{0,1,2,3})$ leads to $(\frac{0,1,3,4}{3,1,2,0})$, thus $\pi_k = (3, 1, 2, 0)$.

The OP encoding generates $n!$ symbols π_k . For $n > 3$, this number of symbols is large enough to be considered as nodes of an ordinal pattern transition network (OPTN) with depth $s > 2$. The OPSS encoding aims to optimally reduce this number of nodes by considering symmetries in the π_k symbols. The slope vector was defined as the difference between consecutive values in π_k : $\vartheta_k = \Delta\pi_k$ ⁶. The length of ϑ_k is $n - 1$. Similarly to the OP encoding that groups embedding waveforms according to the order relationship, the OPSS encoding groups them according to the OP slopes (see Fig. 1). Although computed from the permutations, OPSS is an independent encoding technique with particular properties. It

has the advantage of particularly focus on the frequency distribution or variation of the underlying dynamics, rather than the fine permutation identity that is lost. Indeed, we showed in^{6,16,17} that the largest slope $S = \max\{\vartheta_k\}$ is unique for all embedding vectors and equal to L for a period- L dynamics. We further define OPSS Θ_k^0 by summing the values in ϑ_k as

$$\Theta_k^0 = \sum_{i=1}^{n-1} \vartheta_k(i). \quad (1)$$

Although the number of OPs is significantly reduced, the sen-

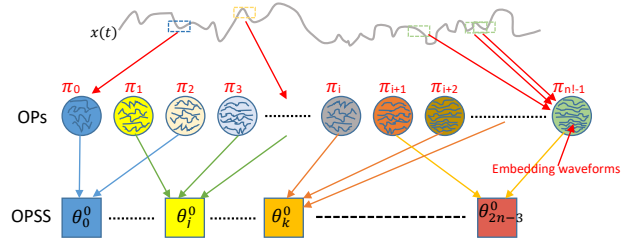


FIG. 1. Example of embedding waveforms clustering based on OP and OPSS encoding methods.

sitivity of OPSS to the frequency distribution of the time series could provide sufficient information to assess its complexity. We verified that $\Theta_k^0 \in \{-n+1, \dots, -1\} \cup \{1, \dots, n-1\}$, so the cardinality of the OPSS set is $\Lambda_0 = 2(n-1)$. For $n = 3$, there are four symbols $\Theta_j^0 \in \{-2, -1, 1, 2\}$ and six symbols π_k : $\pi_0 = (2, 1, 0)$; $\pi_1 = (2, 0, 1)$; $\pi_2 = (1, 2, 0)$; $\pi_3 = (1, 0, 2)$; $\pi_4 = (0, 2, 1)$; $\pi_5 = (0, 1, 2)$. π_1 and π_2 are symmetric, as well as π_3 and π_4 . Indeed, π_2 and π_3 can be obtained by a simple one step right circular shift of π_1 and π_4 , respectively. The corresponding probability distribution of $\{\Theta_j^0\}_{0 \leq j \leq 5}$ is $(\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6})$. The OPSS encoding thus transforms an equiprobable distribution of OP into a symmetric non-equiprobable distribution of OPSS that can be approximated by a gaussian distribution for large n .

The diversity in the OPSS patterns depends on the dynamics under investigation. Deterministic dynamics do not exhibit most of the OPSS with negative code as n increases. When considering, for example, the logistic map (LM) whose equation is $x_{t+1} = \mu x_t(1 - x_t)$ with $3.5 \leq \mu \leq 4$, t being the time index, we observed that it does not exhibit $-(n-1) \leq \Theta_k^0 \leq -2$, while the white Gaussian noise always exhibits the full range of OPSS, $\forall n$. This observation confirms that most of the missing patterns belong to the same clusters (OPSS), which makes characterization of the dynamics easier. Fig. 1 shows how the embedding waveforms are redistributed according to the encoding technique. The OPSS encoding determines further relationships between waveforms belonging to different OP clusters.

Similar to OP, the OPSS encoding focuses on the fundamental trajectory shape of a time series, capturing qualitative features of the underlying phase space. This approach renders the analysis robust against non-stationarities due to an invariant topological structure. While chaotic dynamics are characterized by forbidden or missing patterns, stochastic

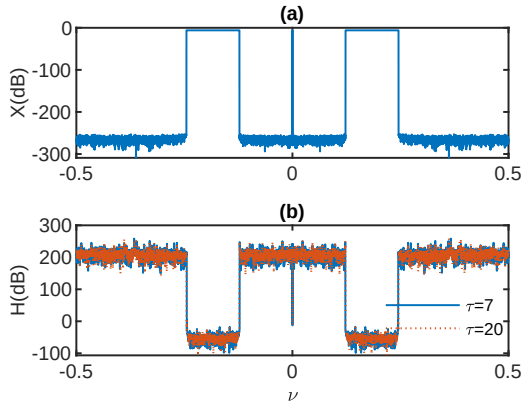


FIG. 2. Harmonic response of the OPSS encoding, highlighting the amplification of low-intensity frequencies and attenuation of high-amplitude ones: (a) spectrum ($X(\text{dB}) = 20\log_{10}(|X(\nu)|)$) of the input signal $x(t) = 0.5 + \sum_{k=1000}^{2000} \sin(2\pi k f_0 t)$, $f_0 = 1\text{Hz}$; (b) harmonic response ($H(\text{dB}) = 20\log_{10}(|H(\nu)|)$) of the OPSS encoding. OPSS encoding adaptively uncovers hidden information within the input signal by selectively amplifying (nonlinear amplification) its low-intensity components. The system performs a nonlinear amplification that extends the spectrum while attenuating high-amplitude frequency components.

noise yields equally likely patterns. This distinction facilitates the discrimination between deterministic and stochastic data through statistical distributions. Furthermore, encoding maps structurally diverse systems from varying domains into a uniform, comparable set of symbolic sequences, thus simplifying behavioral classification and segmentation, a property known as system invariance.

Operationally, the OPSS encoding acts as a nonlinear system that adaptively uncovers and enhances hidden features within the input signal by selectively amplifying weak frequency components that carry critical classification information. As illustrated by the harmonic response in Fig. 2, this transform amplifies weak frequencies and attenuates high-amplitude components simultaneously. Given that the amplification factor depends on both τ and the input signal, varying τ can be conceptualized as a multichannel decomposition framework with distinct channel responses. Integrating these outputs leverages complementary patterns across channels, thereby maximizing discriminative power and optimizing classification performance. This multi-response synergy constitutes the core motivation for the implementation of the sCOPSS algorithm.

B. Definition of sCOPSS

Clustering of missing patterns demonstrates the ability of OPSS encoding to effectively structure time-series sequences, which could foster the emergence of recurrences within OPSS series in the case of deterministic dynamics. Indeed, recurrences reveal the existence of an organization within the time series that can be modeled. Since we assume determinism to involve recurrences over time, the best approach to character-

ize it would be to consider groups of patterns. Therefore, we analyze the time evolution of the $s+1$ OPSS groups (group of s edges or s -type links). To gather information on their temporal distribution, we consider the construction of the network. $s \in \mathbb{N}$ is the depth of the network. The number of possible groups of $s+1$ OPSS is $\Lambda_s = \Lambda_0^{s+1}$, where Λ_0 is the number of possible network nodes and Λ_s , $s > 0$ the number of possible links of type s . By combining groups of s links generated from different time lags τ into a single set, we define the mean rate of links of type s as^{14,15}

$$r_s(n) = \frac{\zeta_s(n) - 1}{\Lambda_s(n) - 1}, \quad (2)$$

where $\zeta_s(n) = \#\left\{\bigcup_{\tau=1}^{\tau_{\max}} \Gamma_\tau\right\}$; $\Gamma_\tau = \{\Theta_j^s(\tau)\}$ is the set of s -type links $\Theta_j^s(\tau)$ of OPSS obtained for a given τ , $j \in \mathbb{N}_{\geq 1}$ (Fig. 3). Similarly, we define the normalized growth rate of the number of s -type links in terms of s as

$$\lambda_s(n) = \frac{\zeta_{s+1}(n) - 1}{\zeta_s(n) - 1}, \quad (3)$$

ζ_s and ζ_{s+1} being the numbers of links of type s , respectively, type $(s+1)$, effectively generated by time series.

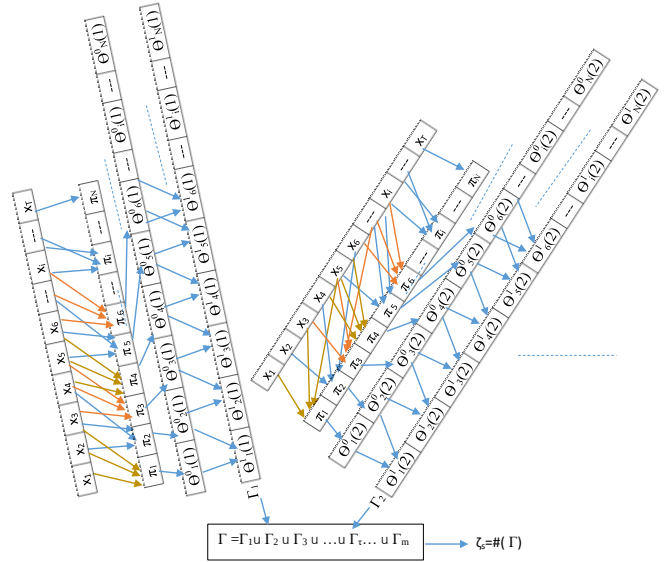


FIG. 3. Principle of the sCOPSS, $n = 3$, $s = 1$. For a given time lag τ , embedding vectors of $n = 3$ samples x_j (input layer) are considered to derive OP π_j (2nd layer), OPSS Θ_j^0 are then derived from each OP (3rd layer), finally, $s+1 = 2$ OPSS are grouped to form type- $s = 1$ links Θ_j^1 (4th layer). The process is repeated for all time lags $1 \leq \tau \leq \tau_{\max}$. The resulting sequences Γ_τ are merged to form a set Γ of s -type links with no redundancy. The cardinality of Γ is used as $\zeta_s(n)$.

C. Observation length requirement

The complexity is usually assessed using entropy measures, which is computationally expensive and should be avoided.

For Λ_s possible states, assuming that each of the states could occur Λ_s times in an observation of length T , $\Lambda_s \log_2(\Lambda_s)$ bits are required to save the subsequent probability distribution, instead of only Λ_s bits with the sCOPSS approach (Fig. 4). Moreover, the observation time T required according to the

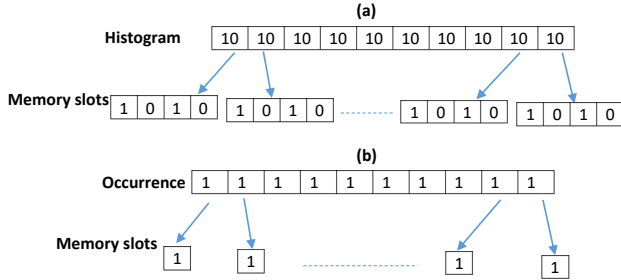


FIG. 4. Example of memory allocation requirement for (a) entropy-based complexity and (b) state occurrence-based complexity. The second case considers a single bit per possible state, while the first one considers binary words whose length depends on the duration of the observation.

sCOPSS definition is directly related to the number $\Omega(n, \tau_{max})$ of the embedding vectors screened. For a given delay time τ , the number of embedding vectors obtained by screening a time series of length T is the arithmetic progression $N_\tau = T - (n-1)\tau$. When varying $1 \leq \tau \leq \tau_{max}$, the total number of embedding vectors screened is an arithmetic series given by

$$\Omega(n, \tau_{max}) = \sum_{\tau=1}^{\tau_{max}} N_\tau, \quad (4)$$

hence

$$\Omega(n, \tau_{max}) = \left(T - \frac{1}{2}(n-1)(\tau_{max} + 1) \right) \tau_{max} \quad (5)$$

Calculating λ_s requires us to consider links of types s and $s+1$. For the set of embedding vectors to be statistically representative, its cardinality should therefore satisfy the condition $\Omega(n, \tau_{max}) \gg \Lambda_{s+1}$. For this purpose, we suggest $\Omega(n, \tau_{max}) > \alpha \cdot \Lambda_{s+1}$, $\alpha \in \mathbb{R}_{\geq 1}$ being a multiplicative factor to satisfy the above statistical requirement, leading to¹⁵

$$T > \alpha \frac{\Lambda_{s+1}}{\tau_{max}} + \frac{1}{2}(n-1)(\tau_{max} + 1), \quad (6)$$

with $\Lambda_{s+1} = \Lambda_0^{s+2}$, $\Lambda_0 = 2(n-1)$. The optimal value of τ_{max} is obtained by solving $\frac{\partial T}{\partial \tau_{max}} = \frac{-\alpha \Lambda_{s+1}}{\tau_{max}^2} + \frac{1}{2}(n-1) = 0$, which yields

$$\tau_{opt} = \left\lceil 2\sqrt{\alpha(2(n-1))^{s+1}} \right\rceil. \quad (7)$$

The corresponding optimal observation length is obtained by replacing τ_{opt} in Eq. (6), hence¹⁵

$$T_{opt} = \left\lceil \frac{n-1}{2} + \sqrt{\alpha(2(n-1))^{s+3}} \right\rceil. \quad (8)$$

α should be chosen as large as possible to refine the nature of the dynamics, while s should be chosen as large as possible for better resolution or separability of the method. The trade-off between the two lies in the observation time, which can be short in practice. In the conventional OPTN approach, the number of embedding vectors also considered as the required data length is expressed as $\Omega = T - (n-1)\tau < T$, which means that it cannot exceed the observation length T . In multilag consideration, the required data length is determined as $\Omega(n, \tau_{opt}) = \frac{(T_{opt} - \Lambda_0)^2}{\Lambda_0} = \alpha \Lambda_{s+1}$, where T_{opt} represents the effective observation length necessary to cover $\Omega(n, \tau_{opt})$. In the first case, $\Omega < T$ while $\Omega > T_{opt}$ in the multilag approach, thus confirming a reduction in the observation time requirement.

II. DETERMINISM INDEX D_s

In order to determine the nature of dynamics, we consider that a purely random (stochastic) dynamics is characterized by the point $A(\lambda_{s-1} = 1, r_{s-1}, r_s = 1)$ contained in the plane $(\mathcal{P}) : -x + z = 0$. Assuming that a system is as stochastic as its number of distinct groups of patterns is sensitive to s , we define the determinism index $-1 \leq D_s(n) \leq 1$ as the signed distance between $A(1, r_{s-1}, 1)$ and the plane $(\mathcal{P}_1) : -r_s(n) \cdot x + \lambda_{s-1}(n) \cdot z = 0$ containing the point $M(\lambda_{s-1}(n), r_{s-1}(n), r_s(n))$ characterizing the underlying dynamics

$$D_s(n) = \frac{r_s(n) - \lambda_{s-1}(n)}{\sqrt{r_s^2(n) + \lambda_{s-1}^2(n)}}. \quad (9)$$

As shown in Fig. 5, $(\mathcal{P})$ -determinism plane divides the space into $\lambda_{s-1} > r_s$ and $\lambda_{s-1} \leq r_s$ areas. Assuming that for $\alpha \rightarrow +\infty$, the number of patterns varies with s as $\eta_s = \varepsilon \Lambda_0(n) \eta_{s-1}$, where $\frac{1}{\Lambda_0} \leq \varepsilon \leq 1$ with $\varepsilon = 1$ for purely random sources and $\varepsilon = \frac{1}{\Lambda_0(n)}$ for periodic dynamics, λ_{s-1} and r_s could be approximated as $\lambda_s(n) \rightarrow \frac{\varepsilon \Lambda_0 - 1}{\Lambda_0 - 1}$ and $r_s(n) \rightarrow \frac{\varepsilon^s \eta_0}{\Lambda_0}$, where $\lim_{n \rightarrow +\infty} \eta_0 < \Lambda_0$ for deterministic dynamics and $\eta_0 = \Lambda_0$ for purely random sources. The significance of the two areas of the space is determined by solving $\lambda_{s-1} = r_s$, with $\eta_0 = \Lambda_0$, resulting in

$$\varepsilon^s(\Lambda_0 - 1) - \varepsilon \Lambda_0 + 1 = 0, \quad (10)$$

for which $\varepsilon_0 = 1$ is a solution $\forall s \geq 0$. For $s \geq 2$, only one second solution ε_1 exists in $[\frac{1}{\Lambda_0}, 1]$, with $\lim_{s \rightarrow +\infty} \varepsilon_1 = \frac{1}{\Lambda_0}$. For example, $\varepsilon_1 = \frac{1}{\Lambda_0 - 1}$ for $s = 2$ and $\varepsilon_1 = \frac{1}{2} \left(\sqrt{\frac{\Lambda_0 + 3}{\Lambda_0 - 1}} - 1 \right)$ for $s = 3$. In practice, the difficulty of considering $\alpha \rightarrow +\infty$ could lead to $\varepsilon_1 < \varepsilon_0 < 1$ for certain stochastic dynamics. When $\varepsilon_1 = \frac{1}{\Lambda_0}$, two useful intervals are considered: $\varepsilon \geq \varepsilon_0$ where $r_s \geq \lambda_{s-1}$, including stochastic dynamics; and $\varepsilon \in [\varepsilon_1, \varepsilon_0]$ where $r_s < \lambda_{s-1}$, indicating deterministic dynamics. $\varepsilon_{0,1}$ are equilibrium points, with $\varepsilon_1 = \frac{1}{\Lambda_0}$ attracting dynamics with a deterministic trend, and $\varepsilon_0 = 1$ attracting dynamics with a stochastic trend.

From the above, we conclude that $D_s(n) < 0$ is reserved for deterministic dynamics and the opposite side for stochastic dynamics. The dynamics is as stochastic as $D_s(n) \rightarrow 0$ and periodic dynamics are well detected as $\epsilon_1 \rightarrow \frac{1}{\Lambda_0}$.

The determinism index is thus defined as a measure of disorder. We adopt the determinism-complexity plane (D_s, λ_{s-1}) to assess both the nature and the complexity of the data. The interpretation is simplified by defining the signed complexity ζ_s as

$$\zeta_{s-1}(n) = \begin{cases} \lambda_{s-1} - 1, & \text{if } D_s(n) < 0; \\ \lambda_{s-1}, & \text{otherwise.} \end{cases} \quad (11)$$

Except for period-1 cycle dynamics for which $\lambda_{s-1}(n) = r_s(n) = 0$, periodic dynamics are characterized by $\lambda_{s-1}(n) = 0$ and $r_s(n) > 0$, hence $D_s(n) = 1$. For consistency purposes, we also set $D_s(n) = 1$ for period-1 dynamics. $D_s(n) = 1$ is an isolated point that can be reached exclusively by periodic dynamics. For purely random dynamics, $\lambda_{s-1}(n) = r_s(n) = 1$, leading to $D_s(n) = 0$. Mixtures exhibit intermediate characteristics ranging from regular to stochastic. Deterministic-stochastic mixtures could exhibit $D_s > 0$, while chaotic dynamics exhibit $D_s < 0$. Consequently, $\zeta_s = 0$ for periodic dynamics, $\zeta_s = 1$ for purely random dynamics (stochastic), and $\zeta_s < 0$ for deterministic dynamics (chaotic). $\zeta_s \rightarrow -1$ in the particular case of quasi-periodic dynamics, but is not equal to -1 by definition. The mixtures exhibit both $\zeta_s < 0$ and $\zeta_s > 0$. In the plane (D_s, ζ_{s-1}), $M(1, 0)$ or $\zeta_{s-1} = 0$ attracts periodic dynamics. Stochastic dynamics are close to $M(0, 1)$ or $\zeta_s \rightarrow 1$, as shown in Fig. 5. For classification purposes, other distances can be considered.

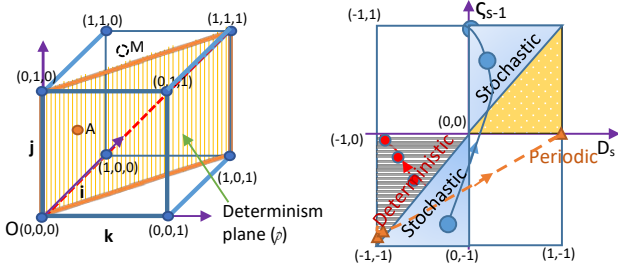


FIG. 5. Visualizing deterministic versus stochastic dynamics. Left: Geometric separation of deterministic dynamics (behind the determinism plane) and stochastic dynamics (in front). Right: Trajectory mapping in the (D_s, ζ_{s-1}) plane governed by α . Stochastic paths move from bottom-left to top-right; chaotic paths move from bottom-right to top-left.

III. VALIDATION

A. Simulation models

In this section, we evaluate the efficiency of our determinism index in measuring disorder and the ability of sCOPSS to discriminate dynamics with close complexity. For this purpose, we used systems with well-known properties.

1. Discriminating stochastic and deterministic dynamics

Correlated stochastic data are modeled as power-law noises with components at all frequencies f and a power spectral density $P \propto 1/f^\beta$ ($\beta \in \mathbb{R}$). Specifically, we consider three basic model types: (i) Gaussian noises (GN), including red ($\beta = -2$: RGN), pink ($\beta = -1$: PGN), white ($\beta = 0$, uncorrelated: WGN), blue ($\beta = 1$), and violet ($\beta = 2$); (ii) auto-regressive (AR) processes governed by $\eta_{t+1}(\gamma) = \gamma\eta_t + \text{WGN}_t$; and (iii) chaotic logistic map (LM) with $\mu = 4$. To minimize computational costs, we set $n = 3$, which corresponds to 4 nodes (OPSS).

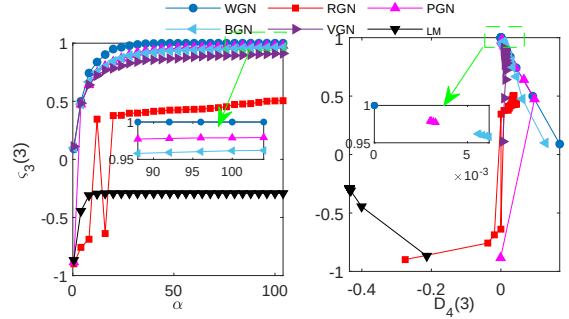


FIG. 6. Noise classification and dynamics characterization via sCOPSS. Left panel: Discrimination of dynamics with close properties using sCOPSS. Right panel: Quantification of the disorder using $D_s(n)$. Low-disorder deterministic dynamics (LM) require a shorter observation length (α) for complexity assessment than highly disordered systems (WGN). Deterministic dynamics are attracted left by the line $D_s(n) = -1$, while stochastic are attracted by $D_s(n) = 0$.

sCOPSS discriminates colored noise with good resolution (Fig. 6). The WGN shows the largest disorder because it has no correlation and the disorder decreases with an increase in the amount of temporal correlation (increase in γ).

When considering AR models, Fig. 7 shows that the disorder decreases with an increase in the amount of temporal correlation (increase in γ). To ensure that sCOPSS effectively distinguishes dynamics with close complexity, we ran sCOPSS 100 times with various s and the observation length factor α for each value of γ to determine the average disorder. The frequency of obtaining $\lambda_3(3) = 1$ with $\gamma = 0$ in the case $s = 4$, is equal to 92% and 99% for $\alpha = 26$ and $\alpha = 30$, respectively. For the closest dynamics ($\gamma = 0.25$), the probability of obtaining $\lambda_3(3) = 1$ is 36% and 47% for $\alpha = 26$ and $\alpha = 30$, respectively. When setting $s = 5$ with $\gamma = 0$, the probability of obtaining $\lambda_4(3) = 1$ is 50% and 96% for $\alpha = 30$ and $\alpha = 42$, respectively. There is no $\lambda_4(3) = 1$ in the case $\gamma = 0.25$ for $\alpha = 30$ and only 3 times $\lambda_4(3) = 1$ for $\alpha = 42$. The mean and standard deviation values are, respectively, $\bar{\lambda}_4(3) = 0.9999 \pm 0.0001$ for $\gamma = 0$ and $\bar{\lambda}_4(3) = 0.9981 \pm 0.0009$ for $\gamma = 0.25$, which clearly shows the importance of s in discriminating short-length dynamics with close properties. The data length requirement for $\alpha = 42$, $n = 3$, and $s = 5$ is $T_{opt} = 1661$, with $\tau_{opt} = 829$. Nevertheless, choosing large α is more necessary to determine the nature of dynamics rather than to discriminate them, while increasing s is helpful

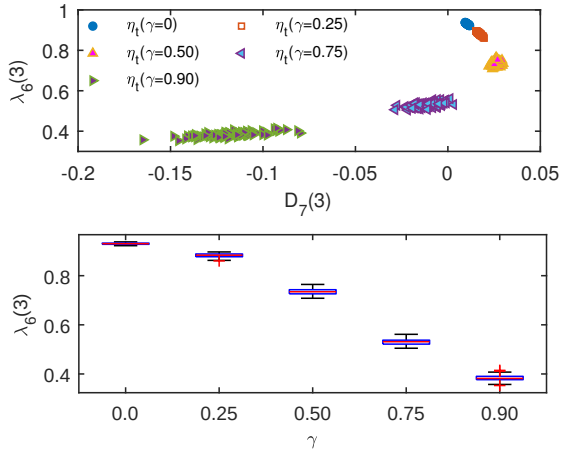


FIG. 7. Discrimination of AR noise with closer properties. We executed the sCOPSS algorithm over 100 independent realizations of white Gaussian noise (WGN) to evaluate the metrics $\lambda_6(3)$ and $D_7(3)$. Setting the threshold parameter to $\alpha = 10$ proved insufficient for detecting $\eta_t(0.75)$ and $\eta_t(0.9)$ as purely stochastic with $D_7(3) > 0$. However, this value was entirely sufficient to successfully distinguish between all considered processes.

in improving the separability of the method.

In the plane (D_s, ζ_{s-1}) , decision-making based solely on the sign of D_s is highly sensitive to the duration of the observation. Furthermore, cases involving periodic dynamics where $D_s > 0$ can be ambiguous. To improve the interpretation, we used the reduced complexity $\zeta_{s-1} - D_s$ and instead considered the plane $(D_s, \zeta_{s-1} - D_s)$. Indeed, we observed that certain stochastic dynamics occupy the region where $\zeta_{s-1} < D_s$ and $D_s < 0$ for short observation durations, while they occupy the region where $\zeta_{s-1} > D_s$ and $D_s > 0$ when the observation duration is long enough. We could then conclude that the dynamics is stochastic as $D_s \cdot (\zeta_{s-1} - D_s) \geq 0$ and deterministic if $D_s \cdot (\zeta_{s-1} - D_s) < 0$.

We further classified the dynamics from the logistic map with $3.5 \leq \mu \leq 4$ by step size 10^{-3} ; the sine-circle map (SM) $\theta_{t+1} = \rho + \theta_t + \frac{\mu}{2\pi} \sin(2\pi\theta_t)$, with $\rho = \frac{-1+\sqrt{5}}{2}$ and $0 \leq \mu \leq 2.5$ by step size 10^{-2} ; the rotation map (RM) $\theta_{t+1} = \rho + \theta_t$ with $0 \leq \rho \leq 0.5$ by step size 10^{-3} ; WGN; PGN; RGN and VGN, using $n = 3$, $s = 8$, and $\alpha \in \{8, 10\}$. The sine-circle map exhibits dynamics close to stochastic behavior for some values of μ , while the rotation map dynamics remain close to the periodic domain (Fig. 8). D_s is a continuum and is indeed negative for deterministic data. Despite the increase in s , the logistic map is fully detected as a deterministic system, while the WGN is detected as stochastic.

2. Scalability of the disorder in terms of s

We verified that s plays the same role as temperature in the distillation process, since increasing it gradually makes $\lambda_{s-1}(n)$ progressively decrease, starting from the lowest level of disorder (Fig. 9). This property makes it possible to increase the distance between the complexity values of even dy-

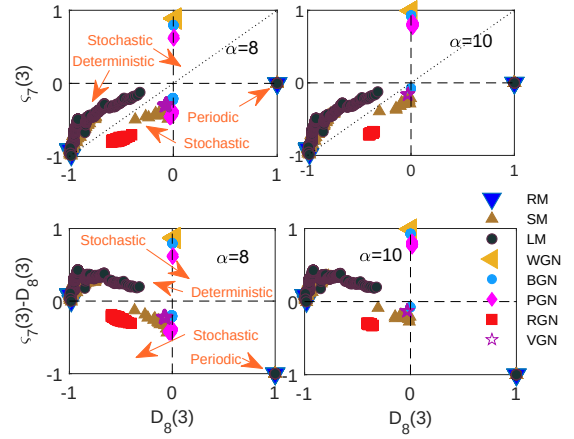


FIG. 8. Classification of dynamical systems using complexity and determinism index. Left panels correspond to $\alpha = 8$; right panels correspond to $\alpha = 10$. Bottom-row panels display the reduced complexity representation in the $(D_s(n), \zeta_{s-1}(n) - D_s(n))$ plane to isolate deterministic and stochastic regions. Quasi-periodic dynamics generated by the sine-circle map cluster are near $(-1, -1)$. Although certain correlated noise profiles and chaotic dynamics exhibit identical complexity $\zeta_s(n)$, they are distinctly differentiated by their determinism index.

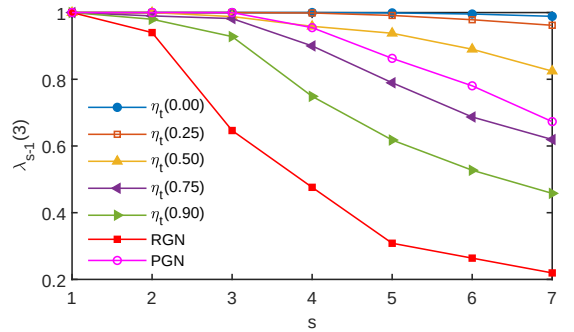


FIG. 9. Resolution enhancement as a function of the scale parameter s . Increasing s enables the separation of closely related dynamics. For a fixed parameter of $\alpha = 20$, the figure demonstrates that at $s = 7$, even the complexity of white Gaussian noise (WGN) begins to decrease, thereby validating the scalability of $\lambda_s(n)$ with respect to s . Furthermore, this example illustrates the capacity of the sCOPSS framework to execute multiclass classification.

namics with close properties, simply by increasing the parameter s .

3. Stochastic/chaotic resonance

Chaotic resonance is known as a phenomenon in which a nonlinear system uses intrinsic or external chaotic fluctuations to amplify weak input signals, allowing it to cross a detection threshold. When the external fluctuation is stochastic, the resonance is said to be stochastic.

We considered the noisy logistic map contaminated with $\eta_t(\gamma)$ as $y_t(\gamma) = x_t + \eta_t(\gamma)$, or with RGN $s(t) = x_t + RGN_t$

or PGN $z(t) = x_t + \text{PGN}_t$ to demonstrate the resonance phenomenon. The analysis of the noisy logistic map as a function of the signal-to-noise ratio reveals a maximum in complexity of the mixture. We consider this maximum to reflect a resonance between chaos and correlated noise. Indeed, adding noise at a low-level contributes to decreasing the complexity of the chaotic dynamics (Fig. 10); then further adding noise will increase the complexity until a maximum (stochastic resonance) before converging towards the complexity of noise.

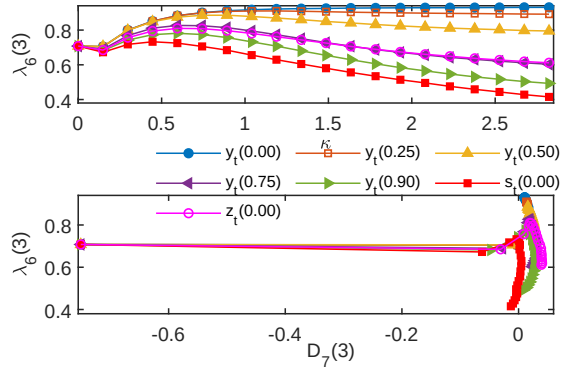


FIG. 10. Stochastic resonance and complexity analysis in a noisy logistic map ($\alpha = 40$). Top panel: Complexity versus κ . Bottom panel: Complexity versus determinism index. Adding chaos to stochastic processes can improve their classification.

To show the chaotic resonance, we considered adding a small amount of chaos to $\eta_t(\gamma)$ and setting $s = 7$. In the case where no chaos is added, sCOPSS failed to detect $\eta_t(0.75)$ and $\eta_t(0.90)$ as stochastic, even with $\alpha = 40$ (Fig. 7). The addition of chaos with $\sigma_x = 0.4\sigma_\eta$ improves the distance between the clusters of $\eta_t(\gamma)$, as well as the determination of their nature as stochastic, with only $\alpha = 10$ (Fig. 11). Chaos therefore can help in easily crossing the stochasticity threshold. Furthermore, the chaotic/stochastic resonance property further reduces the observation time required for a good classification of dynamics.

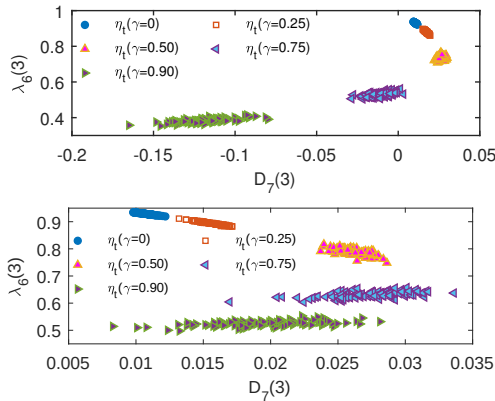


FIG. 11. Discrimination of AR noise with close properties using chaotic resonance. Top and bottom panels correspond to chaos injection rates of $\sigma_x = 0.0\sigma_\eta$ and $\sigma_x = 0.4\sigma_\eta$, respectively.

4. Speed performance

Unlike the graphical complexity-entropy plane, the proposed determinism-complexity plane does not consider probability distributions and differentiates deterministic and stochastic dynamics with similar complexity using the single signed complexity measure. It not only improves the implementation, but also makes it possible to realize a temporal analysis of disorder. Combined with the multi-lag approach, the OPSS encoding improves the observation time and memory constraints of OPTN-based classification methods. Indeed, in the case of OP encoding, the number of possible $s + 1$ -type links is $\Lambda_s^\pi = (n!)^{(s+2)}$, which requires a minimal observation time $\alpha(n!)^{(s+2)} + (n-1)\tau$ for efficient analysis. In our case, the observation time is given as in Eq. (8), which leads to an improvement factor

$$v_L = \frac{\alpha(n!)^{s+2} + (n-1)\tau}{\lceil \frac{n-1}{2} + \sqrt{\alpha\Lambda_{s+2}} \rceil}. \quad (12)$$

This improvement in observation length is important in practice, where data length is usually short.

Similarly, for the same number of embedding vectors screened Ω , the memory requirement of an entropy-based OPTN using OP encoding is $\Lambda_s^\pi \log_2(\Omega)$, while it is only Λ_s in our case, also leading to an improvement factor

$$v_M = \frac{\Lambda_s^\pi}{\Lambda_s} \log_2(\Omega). \quad (13)$$

The improvement in the memory requirement is particularly useful for implementing the algorithm on embedded targets. The above two improvements inevitably imply a reduction in computation time and implementation cost. Compared to OP encoding combined with entropy-based approaches, we estimate the computation time gain of sCOPSS as the ratio of the respective minimal data length (number of embedding vectors) requirements

$$v_T \geq \frac{\Lambda_{s+1}^\pi}{\Lambda_{s+1}} = \left(\frac{n(n-2)!}{2} \right)^{s+2}, \quad n \geq 2 \quad (14)$$

which is quite large. In the case $n = 3, s = 4$, the minimal time gain as in Eq. (14) is $v_T = 1.5^6 = 11.39$. Using a not yet optimized algorithm with OP and OPSS encoding and using the same parameter setting, we found $v_T = 18.70$ by averaging running times of 10 experiments, which is effectively greater than the expected gain of 11.39. The experiment was carried out on an Intel Core i5 computer with 8 GB RAM, running Windows 11.

B. Application to real-world data

1. Exploratory application to audio waveforms

We also applied our algorithm to real-world data such as audio waveforms and electrocardiogram (ECG) data. We first considered MATLAB built-in audio files "Handel", "Splat",

"Chirp", "Train", and "Laughter" and applied our method to the corresponding waveform signals. The LM, RGN and WGN are used to evaluate the relative disorder of these signals. For discriminating them, we set $n = 3$, $s = 8$, and $\alpha = 2$ and used a sliding window of length T_{opt} , with a step size of $\delta_T = \frac{T_{opt}}{3}$. The corresponding result is shown in Fig. 12 (first panel, from left) where distinct clusters can be observed to discriminate signals. Some resulting clusters are sparse, especially for the "Splat" signal. This sparsity reflects the diversity in the nature of the different sequences of real-world signals, unlike the LM, WGN and RGN models, whose clusters are more compact. This compactness of the LM, WGN and RGN clusters confirms that their statistical properties do not significantly vary over time. Although these audio signals are less complex than the LM, they appear mainly to be less deterministic and thereby to contain stochastic components.

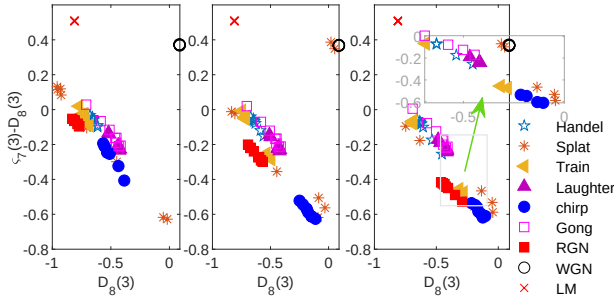


FIG. 12. Analysis of MATLAB built-in audio waveforms under parameter settings $n = 3$, $s = 8$, and $\alpha = 2$. From left to right, chaos is injected at rates of $\sigma_x = 0.0$, $\sigma_x = 0.05\sigma_y$, and $\sigma_x = 0.1\sigma_y$, where x_t represents the logistic map and y_t denotes the analyzed audio file.

To assess the content of each signal, we injected chaos with $\sigma_x = 0.05\sigma_y$ and $\sigma_x = 0.1\sigma_y$, where y_t is the audio waveform or WGN, or RGN, and x_t the logistic map. The corresponding results are shown in Fig. 12 (second and third panels). Fig. 17 shows the magnitude spectrum $|Y_t(\nu)|$ of each sequence analyzed in terms of normalized frequencies ν . "Splat" has two main tones: a hiss (high frequency) and a tap (low frequency). This could explain the two main observed sub-clusters. The hiss corresponds to the block close to $D_8(3) = -1$ where $|Y_t(\nu)|$ appears smoother, hence easy to model, which justifies its detection as deterministic. $|Y_t(\nu)|$ is more complex for tap sequences and appears to be more stochastic than high frequency tones. "Laughter" appears to be a recording of laughter from mainly three to four people. This could be confirmed by the corresponding spectrum $|Y_t(\nu)|$ dominated by three to four peaks. The splitting of its cluster into three closer sub-clusters is consistent with the assumption of the presence of three main laughter. "Train" also features two tones, including a firm two-beat horn blast separated by a short pause. This behavior is illustrated by two sub-clusters observed in Fig. 12 and dominant peaks in $|Y_t(\nu)|$. The "Gong" corresponds to a sound propagation that continually fades away, which could explain the spread spectrum observed. "Handel" and "Chirp" are characterized by compact clusters probably corresponding to a repeating sound pattern. Nevertheless, "Handel" can be

considered to exhibit two to three close tones when considering $\sigma_x = 0.1\sigma_y$. By adding chaos to these audio waveforms, the nature of the stochastic sequences does not change, but even previous deterministic sequences cross the determinism limit, which let us assume that they are initially mixtures of deterministic and stochastic data. This observation also suggests the chaotic resonance phenomenon, which would appear to exacerbate the disorder.

We plot the first return maps $x_{t+1} = f(x_t)$ of the above dynamics to visualize their describability or disorder. It appears that, except for the logistic map that shows a parabolic shape, the other return maps are mainly scattered throughout the phase space (Fig. 13). This observation confirms their stochastic trend.

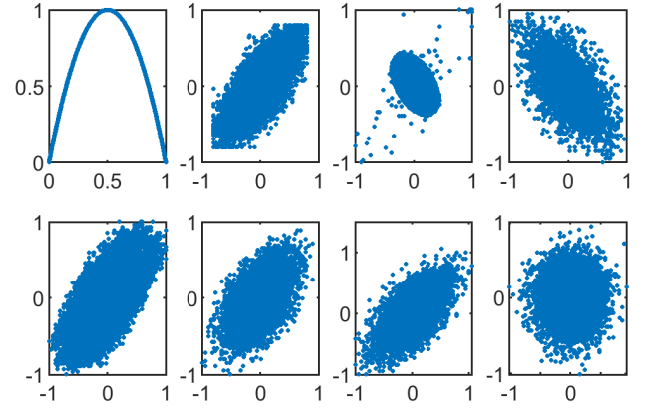


FIG. 13. First return map of analyzed dynamics. From left to right and top to bottom: LM, "Handel", "Splat", "Chirp", "Train", "Laughter", "Gong" and WGN.

2. Application to ECG data

Finally, sCOPSS was applied to the MIT-BIH ECG database from physionet.org. The dataset features several heartbeat categories, including normal sinus rhythm (N) and abnormal types—specifically Premature Ventricular Contractions (PVC) and Premature Atrial Contractions (PAC)—both labeled here as A-type beats. To highlight chaotic/stochastic resonance, we analyzed a normal beat of length $T_N = 294$ (beat #1) and an abnormal beat of length $T_A = 236$ (beat #7) from the recording #100. The entire recording was contaminated with LM and WGN at noise levels $0 \leq \sigma_x \leq 2$. Using algorithm parameters $n = 3$, $s = 8$, and $\alpha = 2$, the extracted beats were processed with sCOPSS, the performance results are illustrated in Fig. 14. Each signal has a distinct complexity value governed by its intrinsic nature and duration, which is itself determined by the respective lengths of the ECG beats. For LM, these baseline values are $LM_N = 0.1162$ (for T_N) and $LM_A = 0.076$ (for T_A). Under WGN, the corresponding baseline values are $WGN_N = 0.0724$ and $WGN_A = 0.0497$ for T_N and T_A , respectively. The complexity of the noisy ECG is denoted as $Y_{ECG+X}(\sigma_x)$, where $X \in \{LM, WGN\}$, $Y \in \{N, A\}$, and σ_x represents the contamination level. In the noise-free

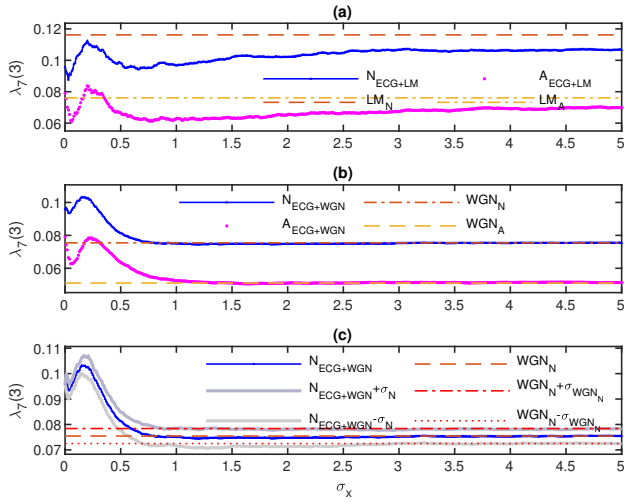


FIG. 14. Dependence of ECG beat complexity on noise contamination levels: (a) case of LM; (b) case of WGN; and (c) average behavior under WGN contamination. Both LM- and WGN-contaminated datasets exhibit complexity extrema. The presence of maxima indicates resonance, which is denoted as chaotic resonance for LM and stochastic resonance for WGN.

case ($\sigma_x = 0$), the baseline complexities are $N_{\text{ECG}+X}(0) = 0.0958$ for the normal beat and $A_{\text{ECG}+X}(0) = 0.079$ for the abnormal beat. For both LM and WGN, the complexity profiles display a local minimum ($Y_{\min}$) and a local maximum ($Y_{\max}$). To ensure statistical reliability for the stochastic case, the WGN experiment was repeated 50 times, consistently yielding the same pattern shown in Fig. 14(c), where the error σ_Y around the mean value of $Y_{\text{ECG}+X}$ is plotted. Crucially, $Y_{\max} \geq Y(0)$ holds regardless of the type of noise or beat category. This phenomenon indicates the presence of a resonance $\mathcal{R}$ —specifically, a chaotic resonance between the ECG signal and LM, and a stochastic resonance between the ECG signal and WGN.

The resonance point marks a fundamental transition in the composition of the signal mixture. Initially dominated by the ECG signal, the complexity of the mixture shifts asymptotically toward that of the noise background as the contamination level (σ_x) increases, with convergence occurring more rapidly under WGN than LM. This behavior effectively compensates for the short duration of the observed signals. Theoretically, given an optimal minimum duration T_{opt} , the complexity of pure WGN should approach 1; however, for a finite length like $T_N = 294$, it evaluates to $\text{WGN}_N = 0.076$. In resonance, the embedding of the ECG component increases the mean complexity to $N_{\text{ECG}+X} = 0.0837$, mirroring the effect of extending the duration of the noise signal. This hypothesis is verified by the complexity gap between WGN_N and WGN_A , which is directly due to the duration difference between $T_N = 294$ and $T_A = 236$. In contrast, injection of noise increases the complexity of ECG beats at resonance, mitigating the limitations of a short observation window. Crucially, the emergence of this resonance is related to the sCOPSS configuration, where the parameters n and s must be explicitly

selected such that $T < T_{\text{opt}}$.

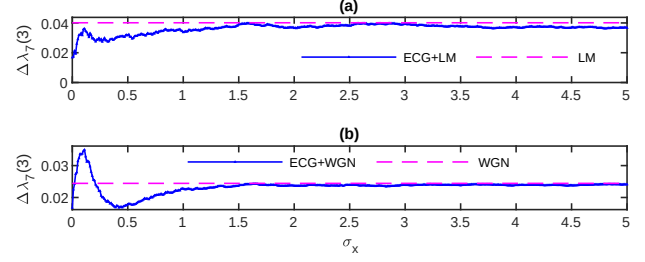


FIG. 15. Dependence of the distance between ECG beat on noise contamination levels: (a) case of LM and (b) case of WGN.

The slow convergence of the complexity of the mixture with LM suggests that high rates of chaos contamination could be considered to increase the separability of the beats, as observed in Fig. 14(a). Furthermore, the larger baseline discrepancy between LM_N and LM_A relative to that between WGN_N and WGN_A indicates that LM noise facilitates better differentiation between normal and abnormal beat classes. Figure 15 quantifies this distance between the complexities of normal and abnormal beats in both LM and WGN contamination regimes. At a noise level of $\sigma_x = 0.11$, both metrics reach a local maximum ($\Delta_{\mathcal{R}}\lambda^{\text{LM}} = 0.0361$ and $\Delta_{\mathcal{R}}\lambda^{\text{WGN}} = 0.035$), significantly outperforming the noise-free baseline ($\Delta_0\lambda = 0.0168$). Ultimately, the absolute maximum separation is recorded at $\sigma_x = 1.58$ under LM contamination, validating its exceptional efficacy for this specific case study.

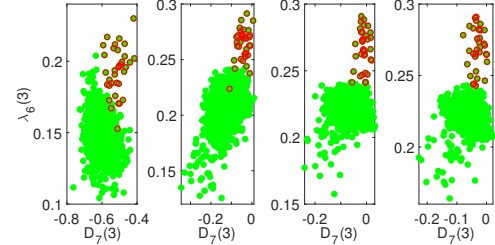


FIG. 16. ECG beat classification using data deficit compensation via chaotic injection ($n = 3$, $s = 7$, and $\alpha = 2$). From left to right, chaos is injected at rates of $\kappa = 0.0$, $\kappa = 0.4$, $\kappa = 0.9$, and $\kappa = 2.0$. Green dots represent normal beats, while red dots indicate abnormal beats. Noise could also be introduced, albeit at a lower proportion.

Following the previous experiment, we applied sCOPSS to the entire recording #100, where the beat length varies between 294 and 408. The parameters were set as $n = 3$, $s = 7$, and $\alpha = 2$, which dictates a minimal data length of $T_{\text{opt}} = 1450$ and $\tau_{\text{opt}} = 724$. Because the maximum ECG beat length falls short of this threshold, we compensated for the data deficit by introducing chaotic noise. Fig. 16 illustrates the classification performance when contaminating the ECG recording with x_t ($\mu = 4$) in various noise-to-signal ratios, $\kappa = \frac{\sigma_x}{\sigma_{\text{ECG}}}$. Cluster discrimination improves as κ increases. Although complexity initially peaks for both groups, the complexity of normal beats decreases before that of abnormal beats, thereby widening the gap between the two clusters.

Although the optimal choice of κ depends on the individual beat, larger values of κ compress the clusters into more compact structures, shifting the classification of the ECG beats towards stochastic data. Ultimately, while large values of s are necessary for robust clustering, the maximum effective limit of κ increases proportionally to the deficit between the data length and T_{opt} .

For validation purposes, we chose a pilot test and validation set of 20 recordings ranging from #100 to #124 of the MIT database and classified the beats into normal and abnormal types. We set $n = 3$, $s = 8$, $\alpha = 2$ and considered the first return map of $r_8(3)$ to evaluate the efficiency of our method (Figs. 19-21). We set threshold lines with equations $y_m = r_s(n) + a_1$ and $x_m = a_2$ (gray lines), where a_1 and a_2 are training parameters. Abnormal beats are therefore those points with the y -coordinate r_s^y verifying the condition $r_s^y < y_m$, and the x -coordinate r_s^x the condition $x > a_2$. A manual selection of abnormal beat clusters in the case of chaos contaminated data with $\sigma_x = 2$ (sCOPSS₂) led to $a_1 = [4.0, 3.11, 3.48, 3.1, -0.98, 3.32, 2.62, 4.26, 7.89, 7.25, 1.56, 4.33, 1.95, 3.81, 0.83, 0.63, 3.19, 2.62, 3.57, 3.01]$ and $a_2 = [0.06, 0.09, 0.07, 0.05, 0.06, 0.08, 0.05, 0.06, 0.05, 0.05, 0.11, 0.08, 0.05, 0.09, 0.08, 0.06, 0.08, 0.05, 0.11, 0.13]$ respectively for the different recordings. In the case of the original data corresponding to sCOPSS₀, we also manually set $a_1 = [3.10, 4.30, 7.22, 3.16, -1.04, 4.25, 2.77, 4.27, 4.13, 3.45, 3.17, 3.01, 3.03, 2.95, 1.79, 0.21, 4.61, 3.74, -1.11, 3.27]$ and $a_2 = [0.02, 0.02, 0.01, 0.01, 0.01, 0.02, 0.01, 0.01, 0.01, 0.01, 0.01, 0.02, 0.01, 0.01, 0.01, 0.01, 0.01, 0.01, 0.02]$. The corresponding results are summarized in Table I and compared to COPPS and sCOPPS methods. The comparison of sCOPSS₀ and sCOPSS₂ confirms the usefulness of adding chaos to the recordings to improve the effectiveness of sCOPSS. Indeed, the addition of chaos allows abnormal beats to stand out from normal beats by exceeding the average complexity by a wide margin, whereas normal beats retain substantially the same average complexity. A comparative analysis shows that sCOPSS consistently

TABLE I. Comparison of COPPS, sCOPPS and sCOPSS classification results for 20 patients from the MIT-BIH database.

Measure	COPPS	sCOPPS _{Post-R}	sCOPPS _{R-R}	sCOPSS ₀	sCOPSS ₂
$Acc(\%)$	98.08	99.46	99.62	99.37	99.80
$S_e(\%)$	99.07	88.79	91.78	86.63	96.24
$S_p(\%)$	70.71	99.85	99.90	99.83	99.93
$P^+(\%)$	73.42	95.58	97.05	94.82	98.01
$P^-(\%)$	98.94	99.56	99.70	99.52	99.86

outperforms COPPS in nearly all metrics evaluated, except sensitivity, where COPPS exhibits a higher baseline. However, this higher sensitivity comes at the expense of lower performance in all other measures. Furthermore, equivalent sensitivity thresholds can still be achieved within the other frameworks. The performance of sCOPSS could be further enhanced by optimizing the noise profile—either by choosing an alternative noise type or by fine-tuning the noise level to maximize beat discrimination. Adopting alternative thresholding strategies could also provide additional performance gains. In Table II, we considered signal-dependent param-

eters as $a_1 = a_{01} \cdot \bar{r}_s$ and $a_2 = \bar{r}_s - a_{02} \sqrt{\bar{r}_s \cdot \sigma_{r_s}}$ to determine whether a given beat belongs to the normal or abnormal beat group, $0 \leq a_{01}, a_{02} \leq 1$. We ran the fitting algorithm with step sizes $\delta_{a_{01}} = 0.001$ and $\delta_{a_{02}} = 0.001$ to find the optimal values. Three levels of LM contamination are considered as $\sigma_x = 0$ (sCOPSS₀), $\sigma_x = 2$ (sCOPSS₂), and $\sigma_x = 3$ (sCOPSS₃). The corresponding results obtained with particular values of a_{01} and a_{02} are summarized in Table II. The complete sensitivity of the classification metric to the optimization parameters a_{01} , a_{02} , and the chaos contamination rate σ_x is shown in Fig. 25.

A comparative evaluation against Permutation Entropy (PE), Approximate Entropy (ApEn), and Sample Entropy (SaEn) was conducted using the identical fitting framework. The corresponding parameterizations and results are given in Table II. For the SaEn formulation, the decision boundary model is inverted so that $y_m = r_s(n) - a_1$ and $x_m = a_2$, where a_1 and a_2 retain their original definitions. Consequently, abnormal heartbeats are classified and isolated based on joint conditions $r_s^y < y_m$ and $r_s^x > a_2$. The results in Table II confirm the superiority of the OPTN-based approaches. Detailed

TABLE II. Comparison of PE, ApEn, SaEn, and sCOPSS classification results for 20 patients from the MIT-BIH database.

Measure	PE	ApEn	SaEn	sCOPSS ₀	sCOPSS ₂	sCOPSS ₃
a_{01}	0.07	0.266	0.62	0.92	0.322	0.408
a_{02}	0.02	0.8	0.01	0.7	0.272	0.344
$Acc(\%)$	95.72	93.48	96.34	98.49	99.4	99.49
$S_e(\%)$	25.77	22.42	50.14	74.51	85.72	88.23
$S_p(\%)$	98.26	96.05	98.01	99.35	99.89	99.89
$P^+(\%)$	34.84	17.05	47.68	80.69	96.62	96.79
$P^-(\%)$	97.34	97.16	98.19	99.08	99.49	99.58

analysis results are shown in Figs. 19-21 for sCOPSS and Figs. 22-24 for PE, ApEn and SaEn, respectively.

C. Future work

In summary, while audio waveform applications demonstrate the feasibility of source separation, future validation requires annotated datasets. Prospective analysis of specific observational models will further consolidate mixture characterization and refine the deployment of sCOPSS on real-world data. In particular, leveraging variations in beat durations enhances the distinction between normal and abnormal beats as their complexities approach their respective chaotic references—an insight that provides promising prospects for optimizing heartbeat classification. Unlike uncorrelated dynamics, mixtures and correlated stochastic processes require extended observation periods to yield a positive determinism index. This reflects the physical reality that low-disorder random fluctuations require extended observation to accumulate and become difficult to model. Consequently, while short-term correlated stochastic dynamics can be modeled deterministically, their long-term trajectories diverge significantly¹⁸. Nevertheless, while brief observations of these stochastic dynamics yield a negative determinism index, they still exhibit a positive value of $(\zeta_{s-1} - D_s) \cdot D_s$,

which distinguishes them from the negative values inherent to deterministic data. Our Matlab code can be downloaded at www.mathematik.uni-kassel.de/~fouda for possible application to other systems.

IV. CONCLUSION

This paper introduces sCOPSS, a computationally efficient measure to analyze short simulations and experimental time series. By characterizing both system disorder and complexity, sCOPSS leverages a unique stochastic/chaotic resonance property to enhance classification performance on brief data. Its mathematical framework minimizes data length and memory overhead, resulting in faster processing times than traditional OPTN-based methods. Additionally, a novel determinism index (D_s) successfully differentiates between stochastic and chaotic dynamics of similar complexity by mapping them to positive and negative values, respectively. While short or low-disorder observation windows can occasionally blur the line between deterministic and stochastic mixtures—as demonstrated by the "Splat" audio waveform analysis—the strict mathematical limit $\lim_{\alpha \rightarrow +\infty} D_s(\zeta_{s-1} - D_s)$ provides a definitive boundary to separate them, although highly correlated dynamics require longer observation times.

- ¹M. Potratzki, M. Adams, T. Bröhl, and K. Lehnertz, "Quantifying the irregularity of a time series," *Phys. Rev. Res.* **8**, 013001 (2026).
- ²L. L. Trulla, A. Giuliani, J. Zbilut, and C. Webber, "Recurrence quantification analysis of the logistic equation with transients," *Physics Letters A* **223**, 255–260 (1996).
- ³B. Deka and D. Deka, "An improved multiscale distribution entropy for analyzing complexity of real-world signals," *Chaos, Solitons & Fractals* **158**, 112101 (2022).
- ⁴S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences* **88**, 2297–2301 (1991).
- ⁵C. Bandt and B. Pompe, "Permutation entropy: A natural complexity measure for time series," *Phys. Rev. Lett.* **88**, 174102 (2002).
- ⁶J. S. A. Eyebe Fouda and W. Koepf, "Detecting regular dynamics from time series using permutations slopes," *Communications in Nonlinear Science and Numerical Simulation* **27**, 216–227 (2015).
- ⁷C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal* **27**, 379–423 (1948).
- ⁸J. Kurths, A. Voss, P. Saparin, A. Witt, H. J. Kleiner, and N. Wessel, "Quantitative analysis of heart rate variability," *Chaos: An Interdisciplinary Journal of Nonlinear Science* **5**, 88–94 (1995).
- ⁹P. Grassberger and H. Kantz, "Generating partitions for the dissipative hénon map," *Physics Letters A* **113**, 235–238 (1985).
- ¹⁰F. Christiansen and A. Politi, "Guidelines for the construction of a generating partition in the standard map," *Physica D: Nonlinear Phenomena* **109**, 32–41 (1997), proceedings of the Workshop on Physics and Dynamics between Chaos, Order, and Noise.
- ¹¹J. V. V. Flauzino, T. L. Prado, N. Marwan, J. Kurths, and S. R. Lopes, "Quantifying disorder in data," *Phys. Rev. Lett.* **135**, 097401 (2025).
- ¹²L. Zunino, M. Zanin, B. M. Tabak, D. G. Pérez, and O. A. Rosso, "Complexity-entropy causality plane: A useful approach to quantify the stock market inefficiency," *Physica A: Statistical Mechanics and its Applications* **389**, 1891–1901 (2010).
- ¹³C. Letellier, I. Leyva, and I. Sendiña-Nadal, "Dynamical complexity measure to distinguish organized from disorganized dynamics," *Phys. Rev. E* **101**, 022204 (2020).
- ¹⁴J. S. A. Eyebe Fouda, W. Koepf, N. Marwan, J. Kurths, and T. Penzel, "Complexity from ordinal pattern positioned slopes (coppss)," *Chaos, Solitons & Fractals* **181**, 114708 (2024).

- ¹⁵J. S. A. Eyebe Fouda, "Short complexity of ordinal pattern positioned slope measure for short-length data analysis," *Information Sciences* **735**, 123095 (2026).
- ¹⁶J. S. A. Eyebe Fouda, J. Y. Effa, M. Kom, and M. Ali, "The three-state test for chaos detection in discrete maps," *Applied Soft Computing* **13**, 4731–4737 (2013).
- ¹⁷J. S. A. Eyebe Fouda, W. Koepf, and S. Jacquir, "The ordinal kolmogorov-sinai entropy: A generalized approximation," *Communications in Nonlinear Science and Numerical Simulation* **46**, 103–115 (2017).
- ¹⁸A. Moujahid and F. Vardillo, "Comparing virus incubation time in sirc models: Deterministic versus stochastic approaches," *Infectious Disease Modelling* **11**, 16–28 (2026).

Appendix A: Calculation details

1. Reduced observation length requirement

The number of possible symbols that can be observed when considering an OPSS network with given n and s is $\Lambda_s = \Lambda_0^{s+1}$. For a given delay time τ , the number of embedding vectors obtained by screening a time series of length T is the arithmetic progression

$$N_\tau = T - (n-1)\tau. \quad (\text{A1})$$

When varying $1 \leq \tau \leq \tau_{\max}$, the total number of embedding vectors screened is an arithmetic series given by

$$\Omega(n, \tau_{\max}) = \sum_{\tau=1}^{\tau_{\max}} N_\tau = \left(T - \frac{1}{2}(n-1)(\tau_{\max} + 1)\right)\tau_{\max}. \quad (\text{A2})$$

To compute λ_{s-1} , the s -type and $s+1$ -type links are considered. The minimal observation length required is Λ_{s+1} related to the $s+1$ -type links. Therefore, we consider that this observation length should correspond at least to the total number of embedding vectors, hence the condition $\Omega > \alpha\Lambda_{s+1}$, where $\alpha \in \mathbb{R}_{\geq 1}$ is a multiplicative factor. This condition leads to the following relation:

$$\left(T - \frac{1}{2}(n-1)(\tau_{\max} + 1)\right)\tau_{\max} \geq \alpha\Lambda_{s+1}, \quad (\text{A3})$$

hence

$$T \geq \alpha \frac{\Lambda_{s+1}}{\tau_{\max}} + \frac{1}{2}(n-1)\tau_{\max} + \frac{1}{2}(n-1), \quad (\text{A4})$$

The optimal value T_{opt} is obtained by solving $\frac{\partial T}{\partial \tau_{\max}} = 0$, hence

$$-\alpha \frac{\Lambda_{s+1}}{\tau_{\max}^2} + \frac{1}{2}(n-1) = 0, \quad (\text{A5})$$

thus leading to

$$\tau_{opt} = \sqrt{2\alpha \frac{\Lambda_{s+1}}{n-1}}, \quad (\text{A6})$$

and

$$T_{opt} = \frac{n-1}{2} + \sqrt{\alpha\Lambda_{s+1}}, \quad (\text{A7})$$

with $\Lambda_s = \Lambda_0^{s+1}$ and $\Lambda_0 = 2(n-1)$. For $\tau_{\max} = \tau_{opt}$, the number of embedding vectors obtained with the effective observation length T_{opt} is $\Omega(n, \tau_{opt}) = \alpha\Lambda_{s+1}$.

2. Observation and computation time gains

The gain in computational efficiency directly depends on the number of embedding vectors screened, Ω , which is inherently governed by the required observation length, T . Because the encoding process of these vectors consumes most of the execution time, computational overhead escalates with their quantity. Furthermore, for any given encoding approach, the minimum number of required embedding vectors is strictly determined by the cardinality of the set of symbols, Λ_0 . For the OP encoding used in the PE algorithm, for example, this minimum number is $n!$, which gives each symbol the chance to appear at least once during observation. A good statistical analysis requires one to go beyond this minimal value. In the case of OPTN where groups of symbols are considered, this basic number simply exponentially increases, depending on the group length (or network depth); and the number of embedding vectors can be set as a multiple of the minimal number required. For a given OPTN, taking into account the depth of the network s , the required number of embedding vectors is thus given as

$$\Omega = \alpha \Lambda_0^{s+1}, \quad (\text{A8})$$

where $\alpha \in \mathbb{R}_{\geq 1}$ is a multiplicative factor. To avoid large values of α , we instead considered $\Omega = \alpha \Lambda_0^{s+2}$. Therefore, in the case of the OP encoding, the minimal number of embedding vectors is given as

$$\Omega_{OP} = \alpha (n!)^{s+2}, \quad (\text{A9})$$

In the case of the OPSS encoding, the basic number of symbols is $\Lambda_0 = 2(n-1)$ and the number of embedding vectors is

$$\Omega_{OPSS} = \alpha (2(n-1))^{s+2}, \quad (\text{A10})$$

The ratio between the above number of embedding vectors is considered as the computational time gain between OP and OPSS-based OPTN and is expressed as

$$v_T = \frac{\Omega_{OP}}{\Omega_{OPSS}} = \frac{(n!)^{s+2}}{(2(n-1))^{s+2}}, \quad (\text{A11})$$

which can be simplified as

$$v_T = \left(\frac{n(n-2)!}{2} \right)^{s+2}, \quad (\text{A12})$$

Depending on the algorithm, the requirement for the observation length is also improved. In the case of conventional OPTN, the relationship between the number of embedding vectors and the required observation length is given as

$$\Omega = T - (n-1)\tau, \quad (\text{A13})$$

where the delay time τ is usually set as $\tau = 1$. In the case of a multilag approach as in the sCOPSS algorithm, the number of embedding vectors is obtained as in Eq. (7). Let T_1 be the observation length required to generate Ω embedding vectors with the conventional OPTN and T_2 be the one to generate Ω with the multilag approach, the resulting observation time gain is

$$v_L = \frac{T_1}{T_2} = \frac{\Omega + (n-1)\tau}{\frac{\Omega}{\tau_{max}} + \frac{1}{2}(n-1)(1 + \tau_{max})}. \quad (\text{A14})$$

Setting $\tau_{max} = \tau_{opt}$ leads to $T_2 = T_{opt}$ as given in Eq. (A7), hence

$$v_L = \frac{\Omega + (n-1)\tau}{\frac{n-1}{2} + \sqrt{\alpha \Lambda_{s+2}}}. \quad (\text{A15})$$

Since $\Omega = \alpha \Lambda_{s+1}$, Eq. (A15) becomes

$$v_L = \frac{\alpha \Lambda_{s+1} + (n-1)\tau}{\frac{n-1}{2} + \sqrt{\alpha \Lambda_{s+2}}}. \quad (\text{A16})$$

In the particular case of OP encoding,

$$v_L = \frac{\alpha (n!)^{s+2} + (n-1)\tau}{\frac{n-1}{2} + \sqrt{\alpha (n!)^{s+3}}}. \quad (\text{A17})$$

In the case of OPSS encoding, $(n-1) = \frac{\Lambda_0}{2}$, hence

$$v_L = \frac{4\alpha \Lambda_s}{1 + 4\sqrt{\alpha \Lambda_s}} + \frac{2\tau}{1 + 4\sqrt{\alpha \Lambda_s}}. \quad (\text{A18})$$

The combined improvement due to OPSS encoding and multilag consideration compared to OP encoding used in conventional OPTN is obtained as

$$v_L = \frac{\alpha (n!)^{s+2} + (n-1)\tau}{\frac{n-1}{2} + \sqrt{\alpha (2(n-1))^{s+3}}}. \quad (\text{A19})$$

Appendix B: Details of the real-world data analysis

This section gives detailed results, as commented on in section III-B. The spectra along with the corresponding values of $\zeta_7(3)$ and $D_8(3)$ for the height sequences of each audio waveform, as well as the logistic map, red Gaussian, and white Gaussian noises. Fig. 17 shows the results obtained with the original waveforms, while Fig. 18 shows the results obtained with chaos contaminated waveforms.

A similar experiment is conducted with series #100 of ECG recordings from the MIT-BIH database. The first return map of $r_s(n)$ obtained with $n = 3$, $s = 8$ and $\alpha = 2$ is depicted in Fig. 19 for the original ECG recordings and in Fig. 20 for the chaos contaminated ECG, with $\sigma_x = 2$. The first return map is obtained by considering the time series $r_s(n)$ of consecutive ECG beats. The comparison of the two experiments shows that the addition of chaos improves the distance between the normal and abnormal clusters.

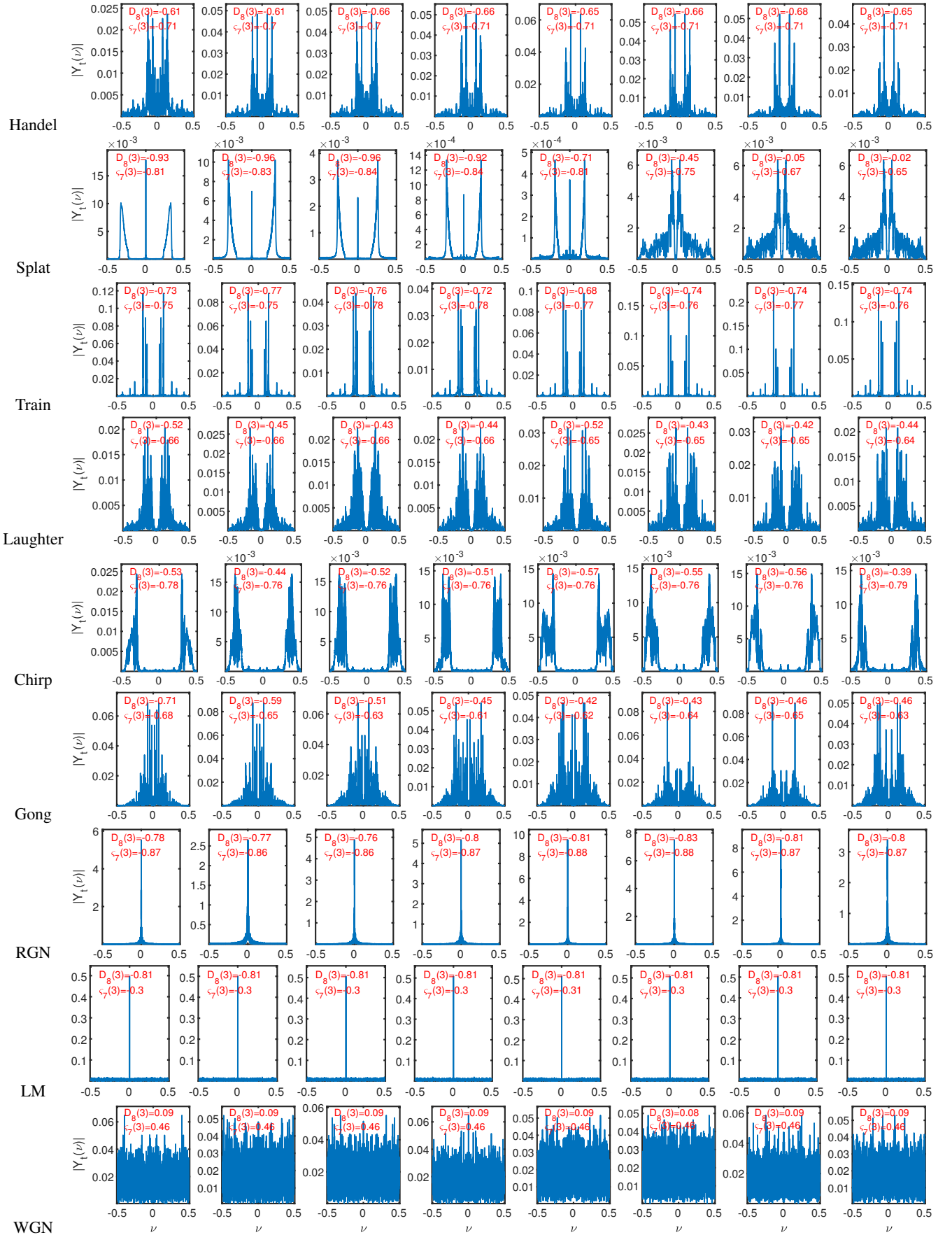


FIG. 17. Relationship between the descriptability of the magnitude spectrum $|Y_t(\nu)|$ and the nature of analyzed sequences of the MATLAB built-in audio waveforms. ν is the normalized frequency. The apparently ordered spectrum of the first four sequences of Splat are detected as deterministic.

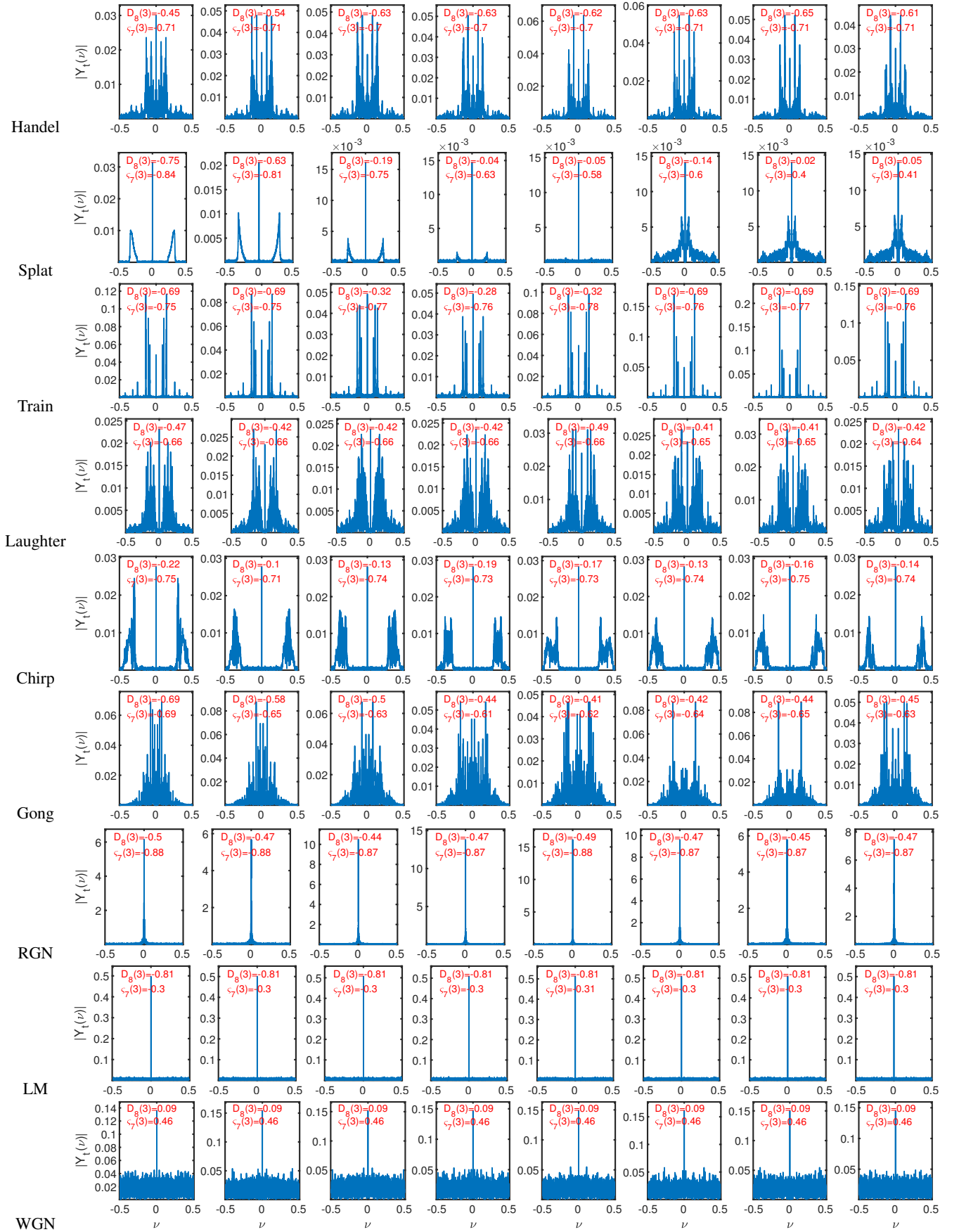


FIG. 18. Relationship between the descriptability of the magnitude spectrum $|Y_t(\nu)|$ and the nature of analyzed sequences of the MATLAB built-in audio waveforms. Audio waveforms are contaminated with chaos at a rate of $\sigma_x = 0.1$, x_t being the logistic map. ν is the normalized frequency.

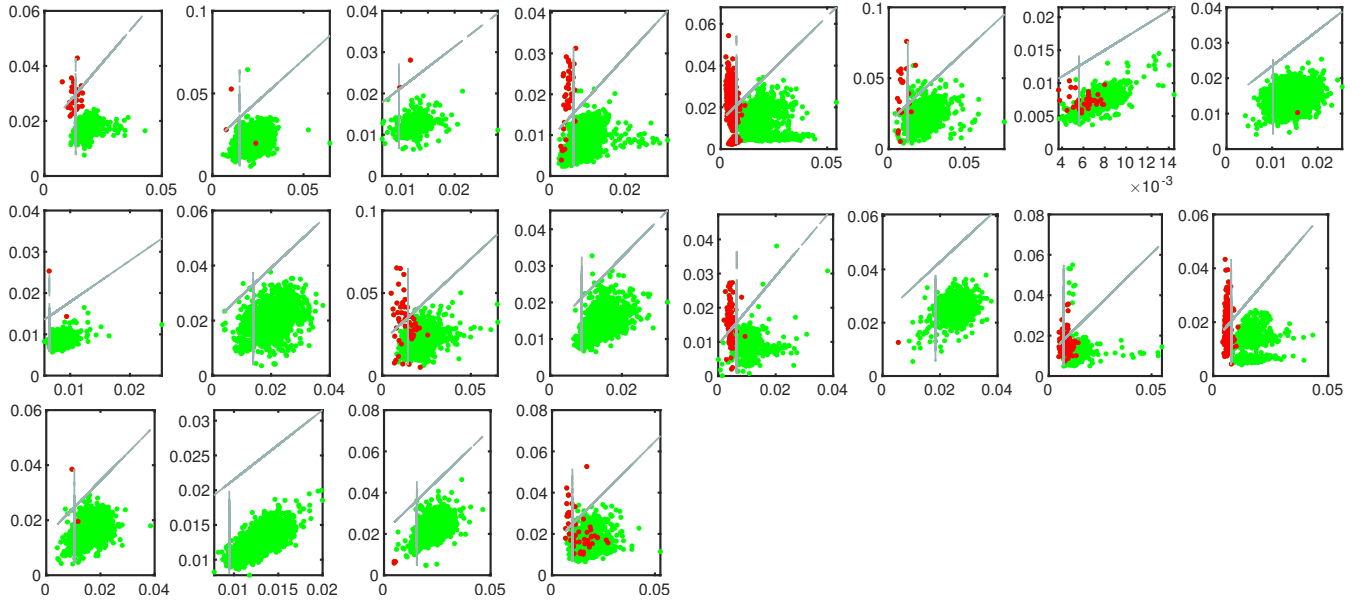


FIG. 19. First return map of $r_8(3)$ of noise free ECG recordings from the MIT database. Although abnormal beats (red) consistently cluster in the upper-left corner, they still partially overlap with normal beats (green).

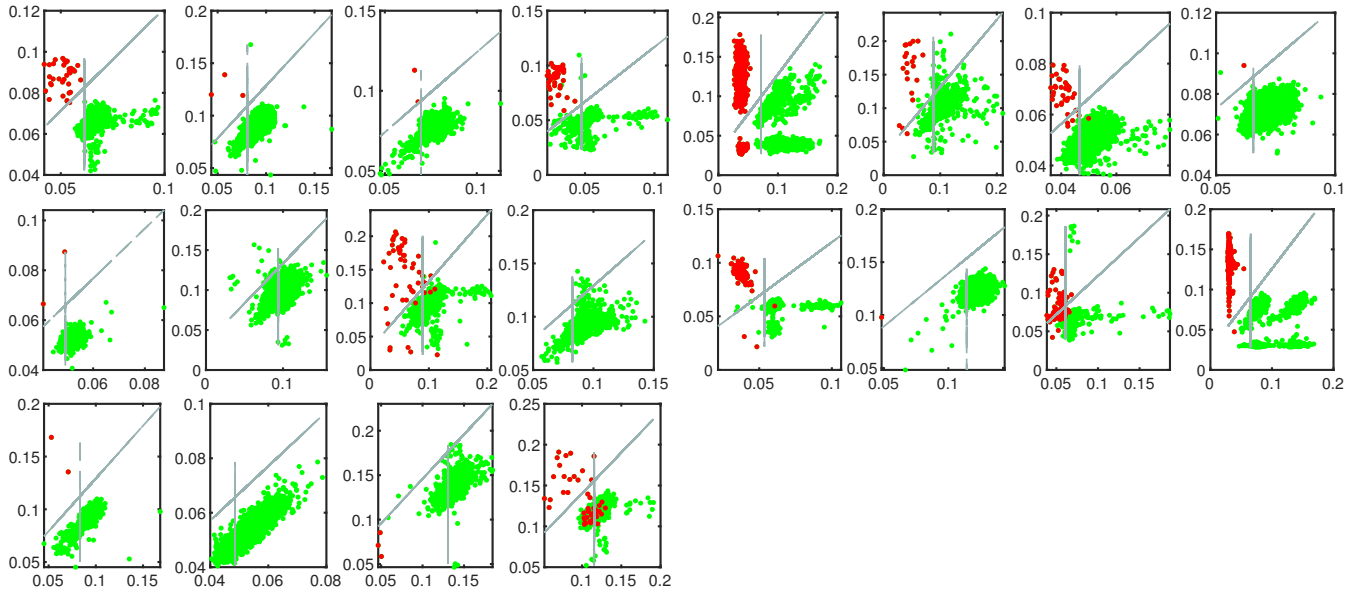


FIG. 20. First return map of $r_8(3)$ of ECG recordings contaminated with chaotic noise ($\sigma_x = 2$, using a logistic map with $\mu = 4$). The overlap between the abnormal (red) and normal (green) beat clusters is reduced, while abnormal beats continue to consistently cluster in the upper-left corner.

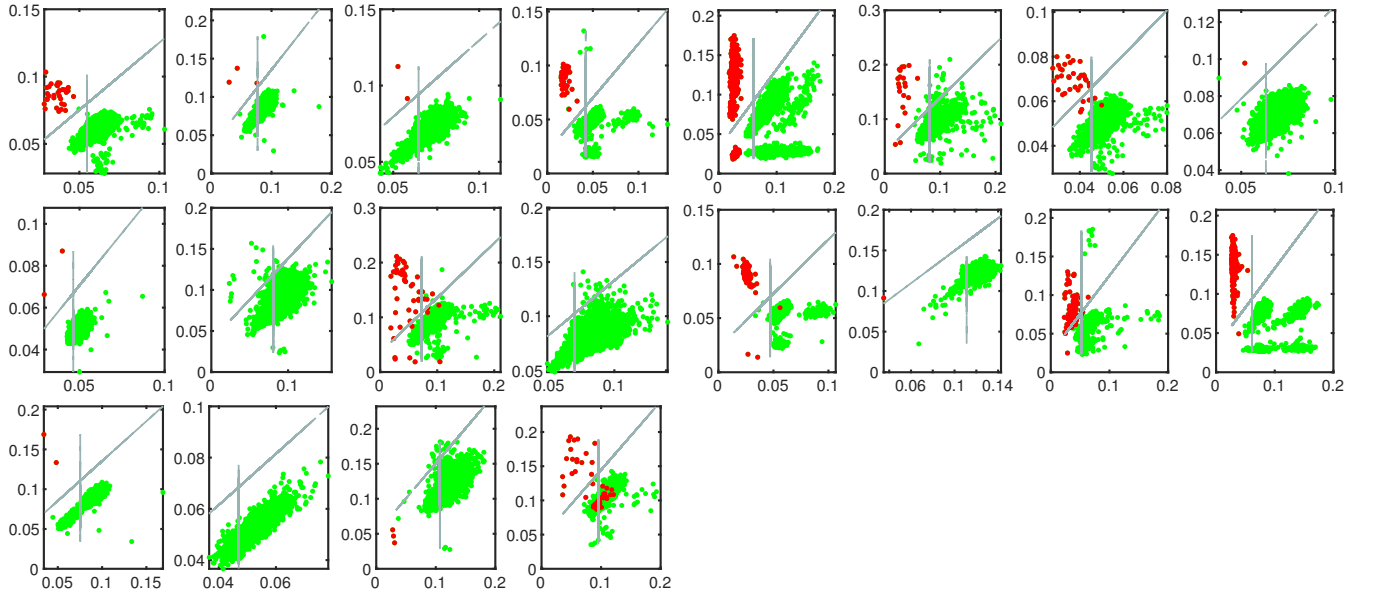


FIG. 21. First return map of $r_8(3)$ for ECG recordings contaminated with chaotic noise ($\sigma_\chi = 3$, using a logistic map with $\mu = 4$). The gap between the normal (green) and upper-left abnormal (red) beat clusters is further widened, maximizing the discriminative power between the two classes.

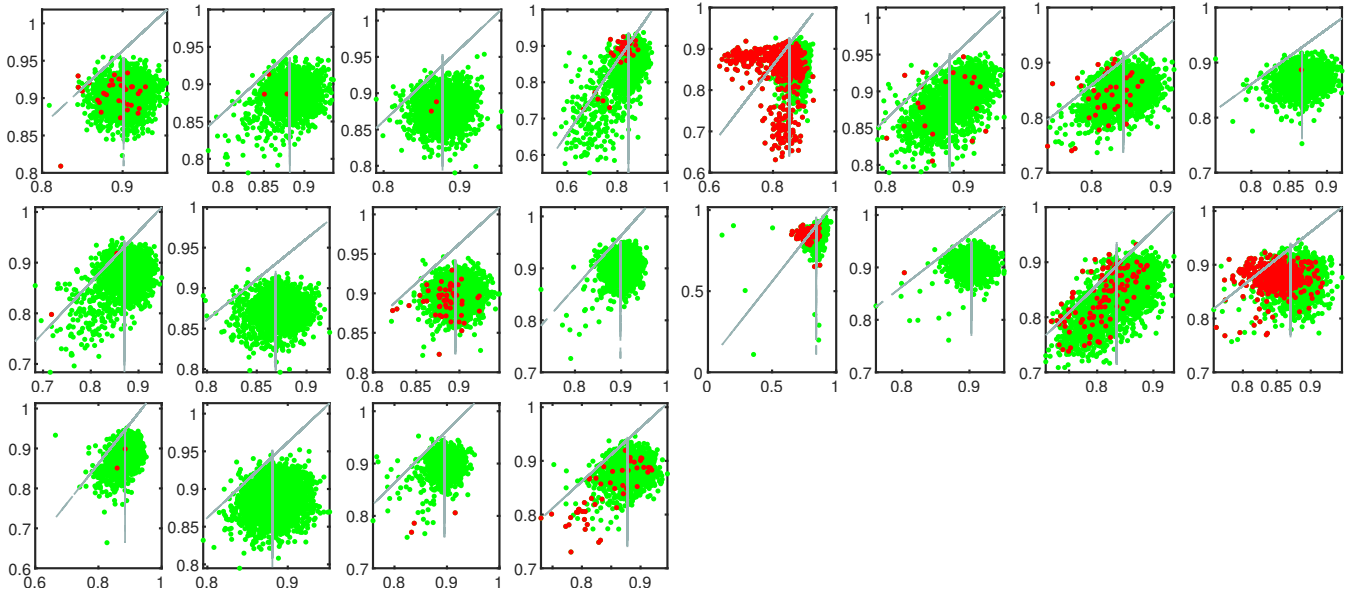


FIG. 22. First return map of PE for ECG recordings, illustrating the persistent overlap between abnormal (red) and normal (green) beats.

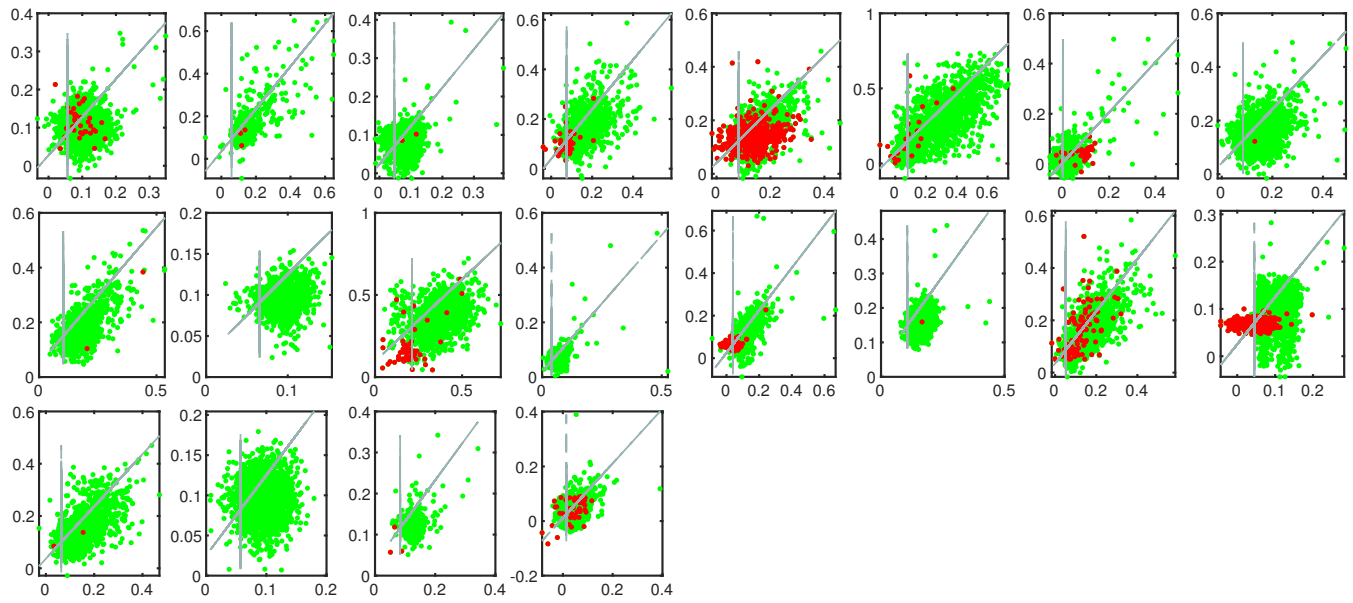


FIG. 23. First return map of ApEn of ECG recordings, showing abnormal beats (red) intermingled with normal beats (green).

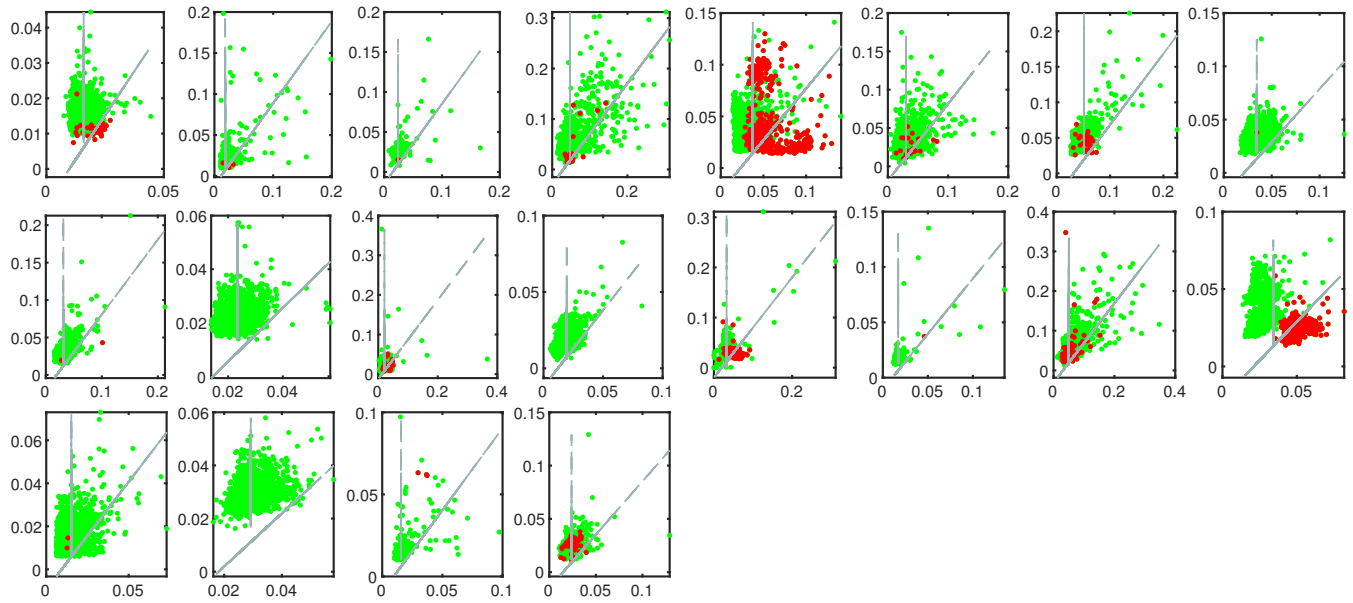


FIG. 24. First return map of SaEn for ECG recordings illustrating superior cluster separation between abnormal (red) and normal (green) beats relative to PE and ApEn.

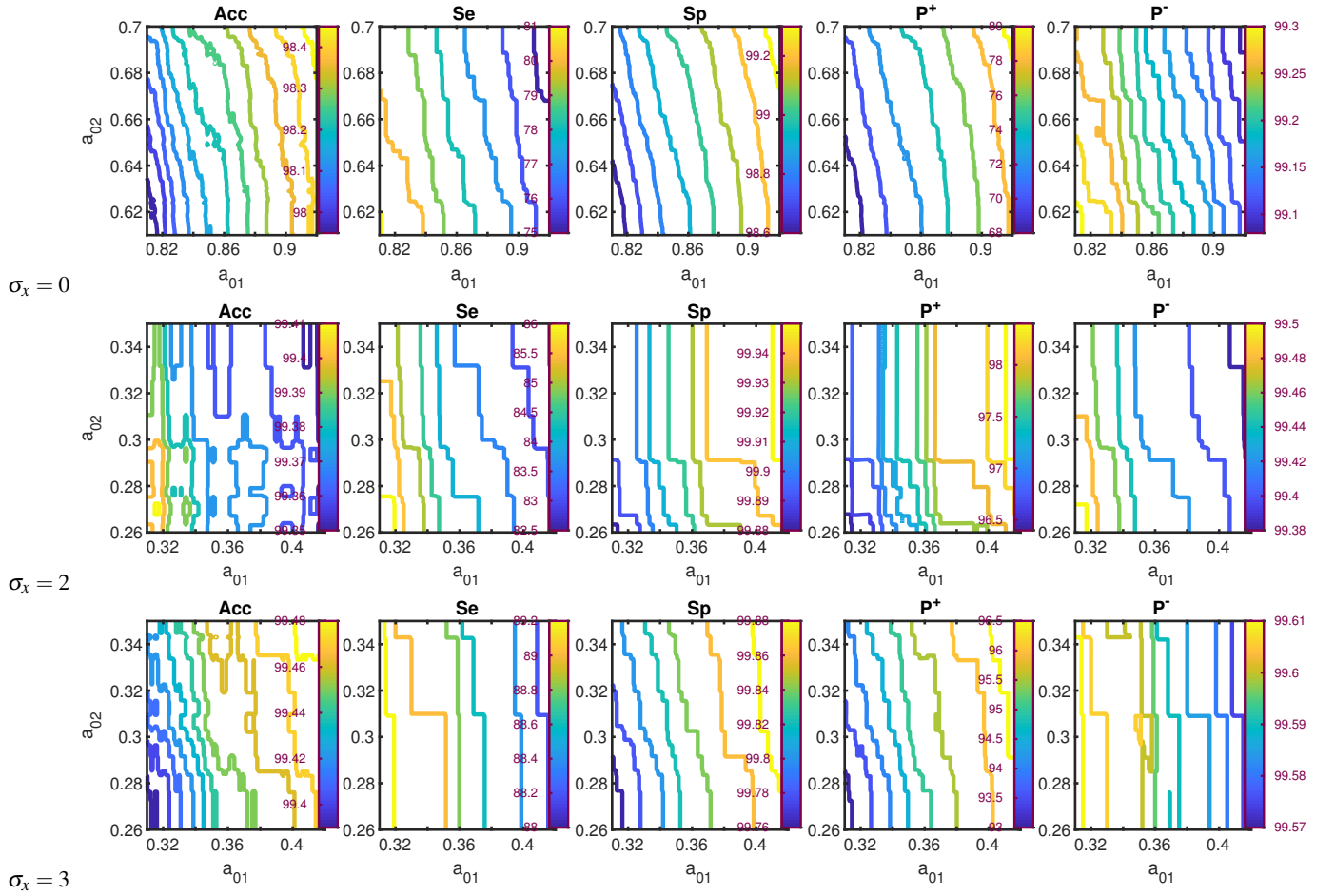


FIG. 25. Classification performance sensitivity to optimization parameters a_{01} , a_{02} , and chaos contamination rate σ_x for the sCOPSS algorithm: the overall performance improves as σ_x increases. Identifying a range of learning parameters that optimizes the trade-off among performance indicators becomes easier as the noise level increases. At $\sigma = 0$, no overlapping zone exists for a_{01} or a_{02} where all metrics simultaneously are close to their respective high percentage. Increasing the noise to $\sigma_x = 2$ widens the high-percentage zones for a_{02} , though a common intersection at the maximum level remains absent. In contrast, the $\sigma_x = 3$ case reveals a shared intersection zone for $a_{02} > 0.33$ where all metrics approach their respective highest percentage, thereby validating the efficiency gains brought by the introduction of chaos.