Detecting Stochasticity in Discrete Signals via Nonparametric Excursion Theorem

Sunia Tanweer[1,2*]   Firas A. Khasawneh[2]

1. Dept. of Mechanical Engineering, Michigan State University
2. Dept. of Computational Mathematics, Science and Engineering, Michigan State University

Abstract

We develop a practical framework for distinguishing diffusive stochastic processes from deterministic signals using only a single discrete time series. Our approach is based on classical excursion and crossing theorems for continuous semimartingales, which correlates number N_ε of excursions of magnitude at least ε with the quadratic variation $[X]_T$ of the process. The scaling law holds universally for all continuous semimartingales with finite quadratic variation, including general Ito diffusions with nonlinear or state-dependent volatility, but fails sharply for deterministic systems—thereby providing a theoretically-certfied method of distinguishing between these dynamics, as opposed to the subjective entropy or recurrence based state of the art methods. We construct a robust data-driven diffusion test. The method compares the empirical excursion counts against the theoretical expectation. The resulting ratio $K(\varepsilon) = N_\varepsilon^{\mathrm{emp}}/N_\varepsilon^{\mathrm{theory}}$ is then summarized by a log–log slope deviation measuring the ε^{-2} law that provides a classification into diffusion-like or not. We demonstrate the method on canonical stochastic systems, some periodic and chaotic maps and systems with additive white noise, as well as the stochastic Duffing system. The approach is nonparametric, model-free, and relies only on the universal small-scale structure of continuous semimartingales.

Keywords: stochastic diffusion, discrete timeseries, quadratic variation, semimartingales, excursion statistics

1   Introduction

The problem of determining whether a discrete-time signal arises from deterministic chaos or stochastic dynamics has been studied for decades [1–3]. Earlier approaches, such as correlation dimension, Lyapunov exponents and classical surrogate tests, provided theoretical foundations but are now recognized as fragile under short data, noise and nonstationarity [3, 4]. Modern work therefore emphasizes finite-sample, noise-robust diagnostics operating directly on observed sequences [5, 6].

Recent contributions can be grouped into several methodological families. *Ordinal-pattern and permutation entropy methods* quantify the statistics of rank orderings and typically separate chaotic and stochastic dynamics in entropy-complexity representations. This includes multiscale extensions, fuzzy permutation entropy, and ANN-assisted permutation-entropy mapping [3, 5, 7]. *Recurrence-based methods* exploit geometric or temporal asymmetries in recurrence structures, including recent work using recurrence plots

---

for testing stochasticity [8]. *Network-based approaches* map time series to graphs-such as ordinal transition networks and visibility/horizontal visibility graphs-and detect deterministic signatures via nonuniform transition motifs or deviations from universal stochastic graph statistics [9–11]. *Information-theoretic and high-dimensional extensions* include Local Structure Entropy (LSE), which scales ordinal analysis to correlated or multivariate systems and has shown strong performance in distinguishing chaos from structured stochastic processes [12]. Finally, *machine-learning-based classifiers*, including ANN models built on ordinal features and deep neural networks trained directly on raw signals, have demonstrated excellent discrimination accuracy for short series, though often at the cost of interpretability and dependence on representative training corpora [5, 6].

Despite their empirical success, the existing approaches for distinguishing chaos from stochasticity share a common characteristic in their methodology: they are fundamentally data-driven and diagnostic rather than theory-determined. Ordinal-pattern [3], permutation-entropy [5], recurrence [1], and network-based [9, 10] methods all rely on auxiliary constructions—such as time-delay embeddings, symbolic partitions, recurrence thresholds, or graph representations—that introduce multiple hyperparameters whose selection is subjective and often system dependent. In practice, classification outcomes may vary substantially with embedding dimension, delay time, scale, or threshold choice, particularly for short, noisy, or nonstationary signals. Symbolic and graphical methods further require subjective interpretation of plots or clustering tendencies, while machine-learning-based classifiers depend on representative training data, architectural choices, and the assumption that test signals belong to the same distributional class as the training corpus. As a result, even highly accurate empirical methods typically lack universality, interpretability, and guarantees of transferability across systems.

These limitations motivate the development of complementary approaches grounded directly in stochastic process theory. In contrast to empirical pattern-based diagnostics, the method proposed here is informed by rigorous results on continuous semimartingales and their excursion statistics [13]. It operates directly on the observed time series without embedding, symbolic discretization, graphical representations, or model training, and its theoretical predictions are determined by intrinsic pathwise quantities such as quadratic variation. This framework therefore provides a principled, system-agnostic criterion for detecting stochastic diffusion, offering a theoretically interpretable alternative to existing empirical methodologies. The theory of excursions in random processes is mathematically rich [14, 15], with classical results describing the expected number of level crossings or excursions under various structural assumptions. Foundational contributions include Rice's formula for the expected number of level crossings in stationary Gaussian processes [16], its extensions by Ito [17] and Ylvisaker [18], and the generalizations by Leadbetter and Cryer for both stationary and non-stationary settings [19, 20]. Further developments address other stochastic processes, such as the work of Ivanov [21], and discrete-time first-order Markov sequences, for which Zekai Sen derived a recursive characterization of the distribution of up-crossings [22]. These classical excursion results establish the statistical backbone underlying level-crossing behavior in stochastic systems. In this work, we will be using the excursion results recently published for Ito processes [13].

2 Mathematical Background

2.1 Stochastic Dynamical systems

Regular differential equations (ODEs) are the workhorse for modeling continuous change. They express the rate of change of a system's state at any given time, allowing us to predict its deterministic evolution. However, many real-world systems are influenced by random fluctuations, making their behavior inherently stochastic, necessitating the use of stochastic differential equations.

Stochastic differential equations (SDEs) bridge the gap between ODEs and randomness by incorporating terms that introduce randomness into the model. Unlike ODEs with a single solution trajectory, SDEs

have a family of possible solutions. Each solution represents a path the system might take with a certain probability. We denote the state of the system at time t by a vector $X_t \in \mathbb{R}^n$. A drift term μ_t captures the deterministic component of the system's evolution, similar to the right-hand side of an ODE, while a diffusion term Ω_t multiplying a scalar Brownian motion W_t (a standard Wiener process $W_t \sim N(0, t)$) injects stochasticity into an equation.

The introduction of the randomness term into the equation leads to two main mathematical interpretations of the resulting SDE—Ito and Stratonovich—which are given according to

$$dX_t = \mu_t \, dt + \Omega_t \, dW_t \quad \text{(Ito)}; \qquad\qquad dX_t = \mu_t \, dt + \Omega_t \circ dW_t \quad \text{(Stratonovich)}.$$

These two are merely different approaches with different mathematical properties to solve a mathematical equation and generally give different results. For example, the Stratonovich interpretation leads to ordinary chain rule formulas under change of variables, while the Ito formula leads to second order terms. On the other hand, in contrast to Ito integrals, Stratonovich integrals are not martingales (see Section 2.2), so the Ito interpretation enjoys important computational advantages. This work adopts the Ito interpretation only.

2.2 Local Martingales and Semimartingales

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a filtered probability space and let $\mathcal{F}^* := (\mathcal{F}_t)_{(t \geq 0)}$ be the filtration of $\mathcal{F}$. An $\mathcal{F}^*$-adapted process $M : [0, \infty] \times \Omega \to \mathbb{R}$ is an $\mathcal{F}^*$-local martingale, if there exists a sequence of stopping times τ_k such that

1. τ_k are a.s. increasing, i.e. $\mathbb{P}(\tau_k < \tau_{k+1}) = 1$;

2. The τ_k are a.s. divergent, i.e. $\mathbb{P}(\lim_{k \to \infty} \tau_k = \infty) = 1$;

3. The process $M_{t \wedge \tau_k}$ is an $\mathcal{F}^*$-martingale, i.e. for all $s < t$, $\mathbb{E}[M_{t \wedge \tau_k} \mid \mathcal{F}_s] = M_{s \wedge \tau_k}$.

On the other hand, a semimartingale X_t is a process which can be decomposed into the sum of a local martingale M_t and a finite variation process A_t. Semimartingales are the largest class of processes with respect to which the Ito and Stratonovich integrals can be defined. All Ito processes are semimartingales but a Stratonovich may or may not be one.

2.3 Excursion Theory for Continuous Semimartingales

Let $(X_t)_{t \in [0,T]}$ be a real-valued continuous semimartingale with canonical decomposition

$$X_t = X_0 + M_t + A_t,$$

where X_0 is the initial value, M_t is a continuous local martingale and A_t is a finite-variation drift. Its quadratic variation is $[X]_T = \langle M \rangle_T$

Let $N_\varepsilon(X)$ denote the number of excursions or oscillations of amplitude at least ε. Then the following result holds.

Theorem 2.1 (Excursion Law for Continuous Semimartingales [13]). *Let X be a continuous semimartingale with finite quadratic variation $[X]_T < \infty$. Then*

$$\lim_{\varepsilon \to 0} \varepsilon^2 N_\varepsilon(X) = \frac{[X]_T}{2} \qquad \text{in } L^1.$$

Equivalently,

$$N_\varepsilon(X) = \frac{[X]_T}{2\varepsilon^2} \big(1 + o(1)\big).$$

This theorem holds for all continuous Ito diffusions

$$dX_t = \mu(X_t, t)\,dt + \sigma\Omega(X_t, t)\,dW_t,$$

including nonlinear, multiplicative, and state-dependent noise. Deterministic systems do *not* satisfy the ε^{-2} law, as their quadratic variation is zero in the continuous limit.

3 Methods

3.1 Quadratic Variation Estimation

Given discrete observations X_{t_i}, we estimate the quadratic variation and process-dependent theoretical curve as

$$\widehat{[X]}_T = \sum_{i=0}^{n-1}(X_{t_{i+1}} - X_{t_i})^2 \quad \text{and} \quad N_\varepsilon^{\text{theory}} = \frac{\widehat{[X]}_T}{2\varepsilon^2}.$$

Likewise, given the signal, we can also compute the empirical value of the excursion count as $N_\varepsilon^{\text{emp}}$.

3.2 Excursion Invariant Statistics

For a variety of ε, we compute $K(\varepsilon) = N_\varepsilon^{\text{emp}}/N_\varepsilon^{\text{theory}}$. A diffusion must satisfy $K(\varepsilon) \approx 1$ across scales. We choose a range of ε where this is satisfied, and then compute s—the slope obtained by fitting $\log N_\varepsilon \sim s \log \varepsilon$. An Ito driven stochastic signal must satisfy $s \approx -2$, while a deterministic signal will (typically) take smaller values due to having fewer excursions than a diffusion-driven signal. Fig. 1 shows the distribution of the slope for monte carlo simulations for pure Brownian motion. The histogram confirms that the slope is largely concentrated around a value of -2, demonstrating that the proposed method can correctly identify Ito diffusions. From the histograms, we also identify upper and lower thresholds, as -2.5 and -1.0, for the slope to cope with the errors arising due to the discretized and digital nature of real-world signals.
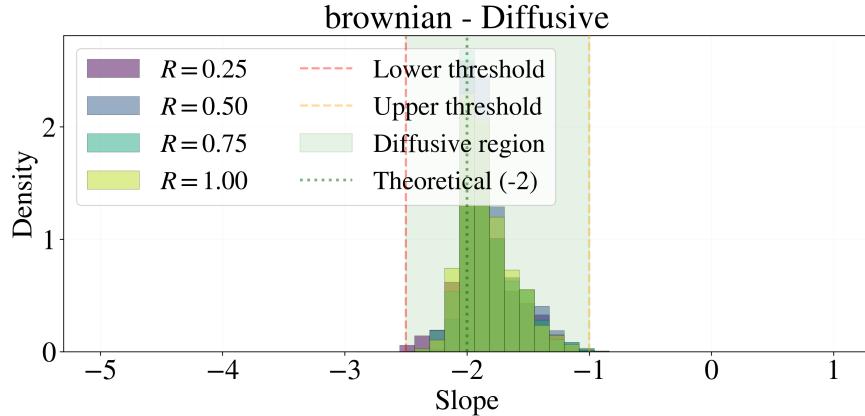

Figure 1: Histograms of the fitted log-log slope of the excursion count N_ε versus ε for Brownian process.

4 Results & Discussion

We evaluate the proposed excursion-based diffusion test across a broad class of deterministic and stochastic dynamical systems, with some real-world signals. For each synthetic system, ensembles of discrete time

---

Algorithm 1 Algorithm for Estimation and Classification

1: Compute the quadratic variation estimate $\widehat{[X]}_T$
2: Choose a logarithmic grid of ε
3: **for** each ε **do**
4:     Compute the empirical $N_\varepsilon^{\text{emp}}$ via the excursion algorithm
5: Compute the theoretical curve $N_\varepsilon^{\text{theory}}$ using quadratic variation
6: Form the excursion ratio $K(\varepsilon)$ and choose a range of ε with $K(\varepsilon) \sim 1$
7: Compute s in that range of ε
8: Classify into diffusion if $s \in [-2.5, -1]$; classify non-diffusive otherwise

---

series were generated over a range of observation lengths T, sampling intervals Δt, and signal-to-noise ratios $(SNR = 10 \log(P_{sig}/P_{noise})$ and $R = \langle \int_0^T \mu(X_t, t) \, dt \rangle / \langle \int_0^T \sigma \, \Omega(X_t, t) \, dt \rangle$ for deterministic and stochastic systems, respectively). Classification accuracy is defined as the fraction of realizations correctly identified as diffusion-like or non-diffusive relative to the known ground truth. Definitions of all systems and parameters used are given in Appendix A.

4.1   Validation on Canonical Diffusion Processes

We begin with classical diffusion processes, including the Ornstein-Uhlenbeck (OU) process, and the Cox-Ingersoll-Ross (CIR) model. Across all tested parameter regimes, these systems are classified as diffusion-like with near-perfect accuracy given sufficient discretization and sample length. Fig. 2 and 3 show the heatmaps of accuracy for each of these systems. For OU processes, the classification accuracy is largely 100%. For CIR, we see the accuracy increasing with R, discretization dt and signal length T.

4.2   Deterministic Periodic System

Purely deterministic oscillators, such as the simple harmonic oscillator, are consistently classified as non-diffusive at low and moderate noise levels. In these regimes, excursion counts grow sub-quadratically in ε^{-1}, yielding slope estimates substantially larger than -2 and systematic deviations from the quadratic-variation-based prediction. As additive noise increases, these systems exhibit a gradual transition toward diffusion-like behavior. At low signal-to-noise ratios (typically SNR ≤ 15), the noise component dominates the local structure of the signal, and the excursion statistics approach those of a diffusion. This transition is visible in the excursion slope distributions and accuracy heatmaps plotted against noise level and sampling resolution, shown in Fig. 4.

4.3   Deterministic Chaotic Maps

We next consider deterministic chaotic maps including Henon, logistic, and the notoriously difficult Linear Congruential Generator (LCG) maps. In the absence of noise, these systems are robustly classified as non-diffusive despite their apparent irregularity. Their excursion statistics fail to satisfy the ε^{-2} scaling law, reflecting the absence of finite quadratic variation in the continuous–time limit. When additive noise is introduced, these systems display sharp transitions from deterministic-like to diffusion-like classification. Notably, it is interesting that for chaotic maps, the accuracy increases as dt and T decrease instead of the other way around. See Fig. 5- 7 for heatmaps of accuracy.
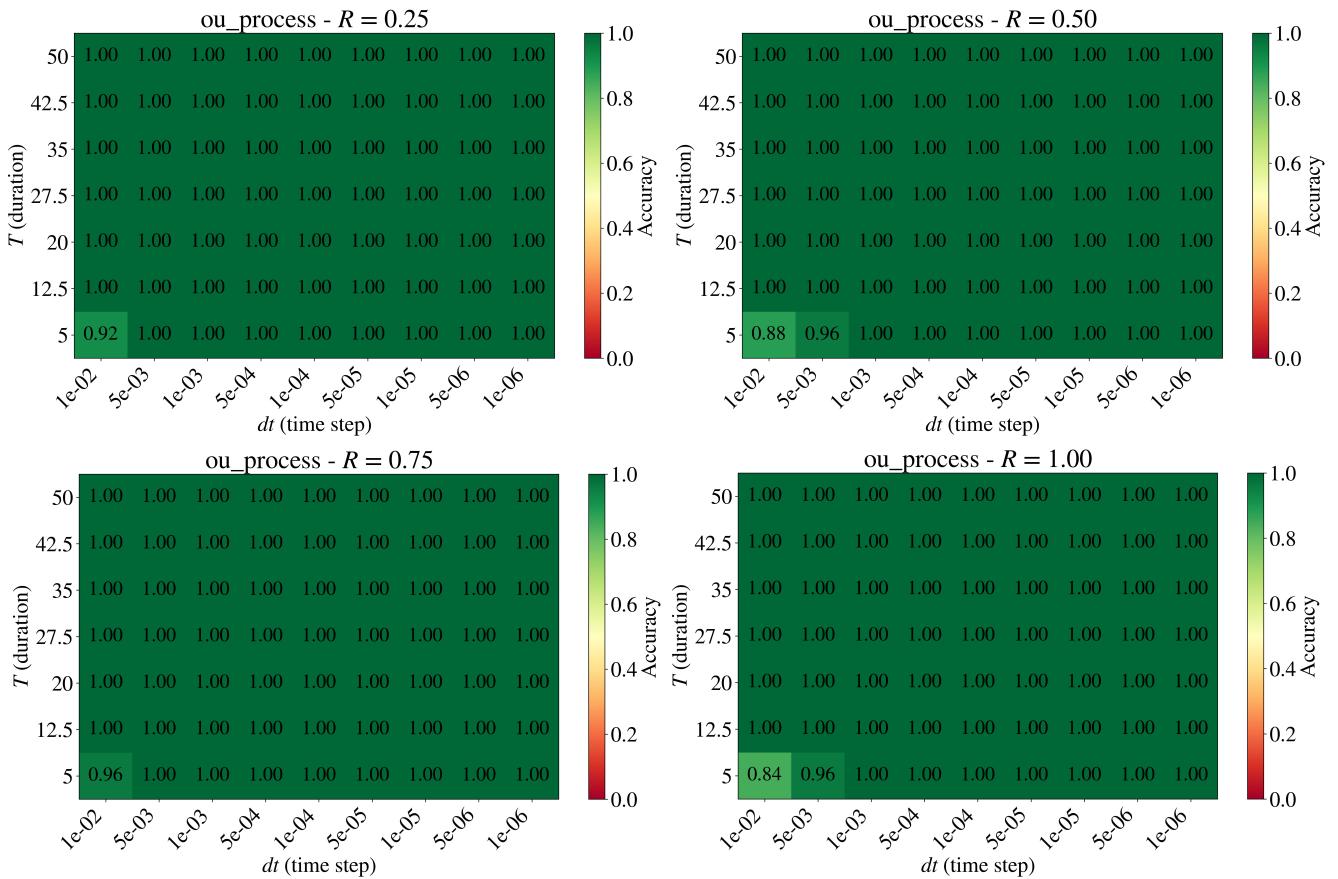
Figure 2: Heatmaps of accuracy for OU process, across various R.
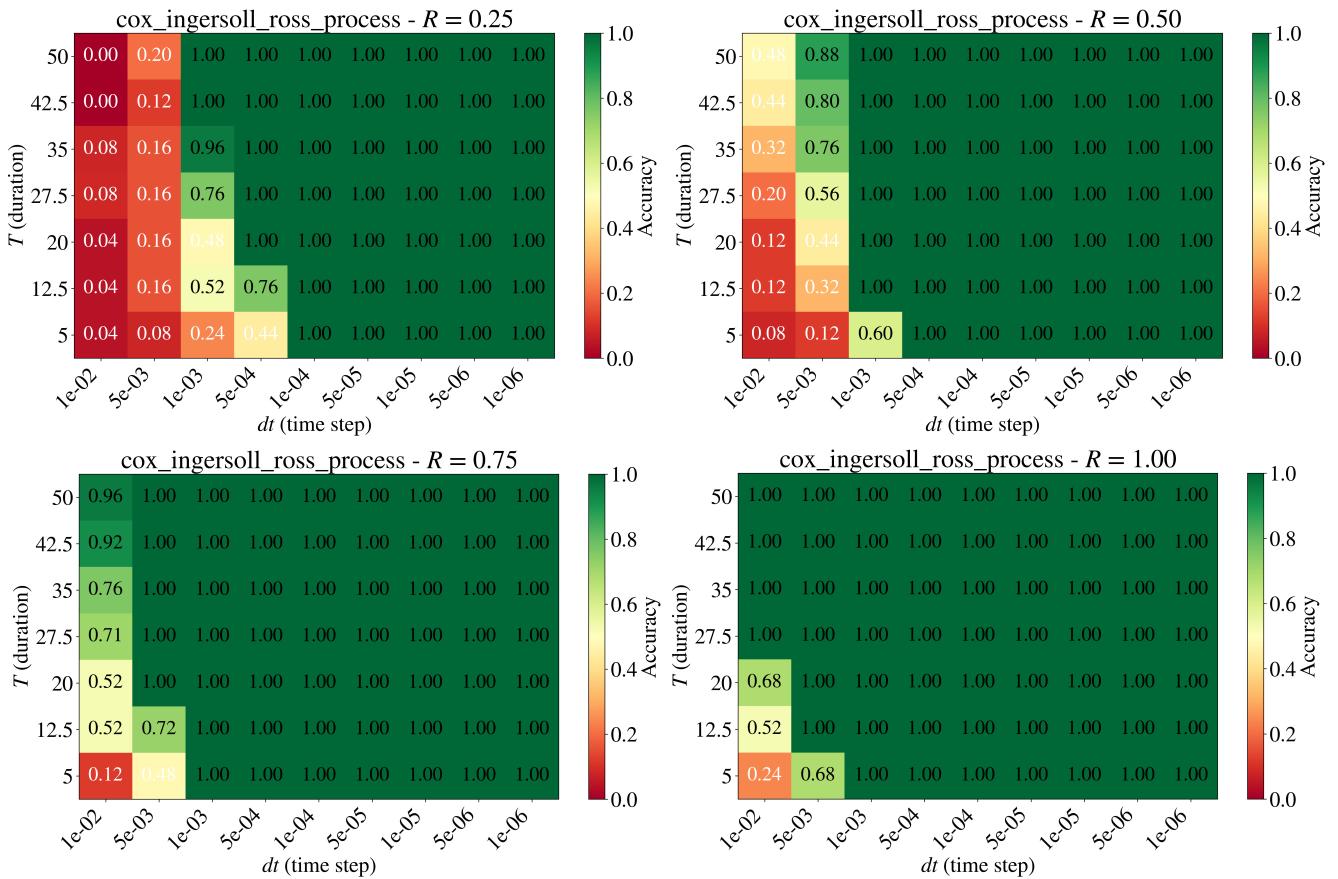
Figure 3: Heatmaps of accuracy for CIR process, across various R.
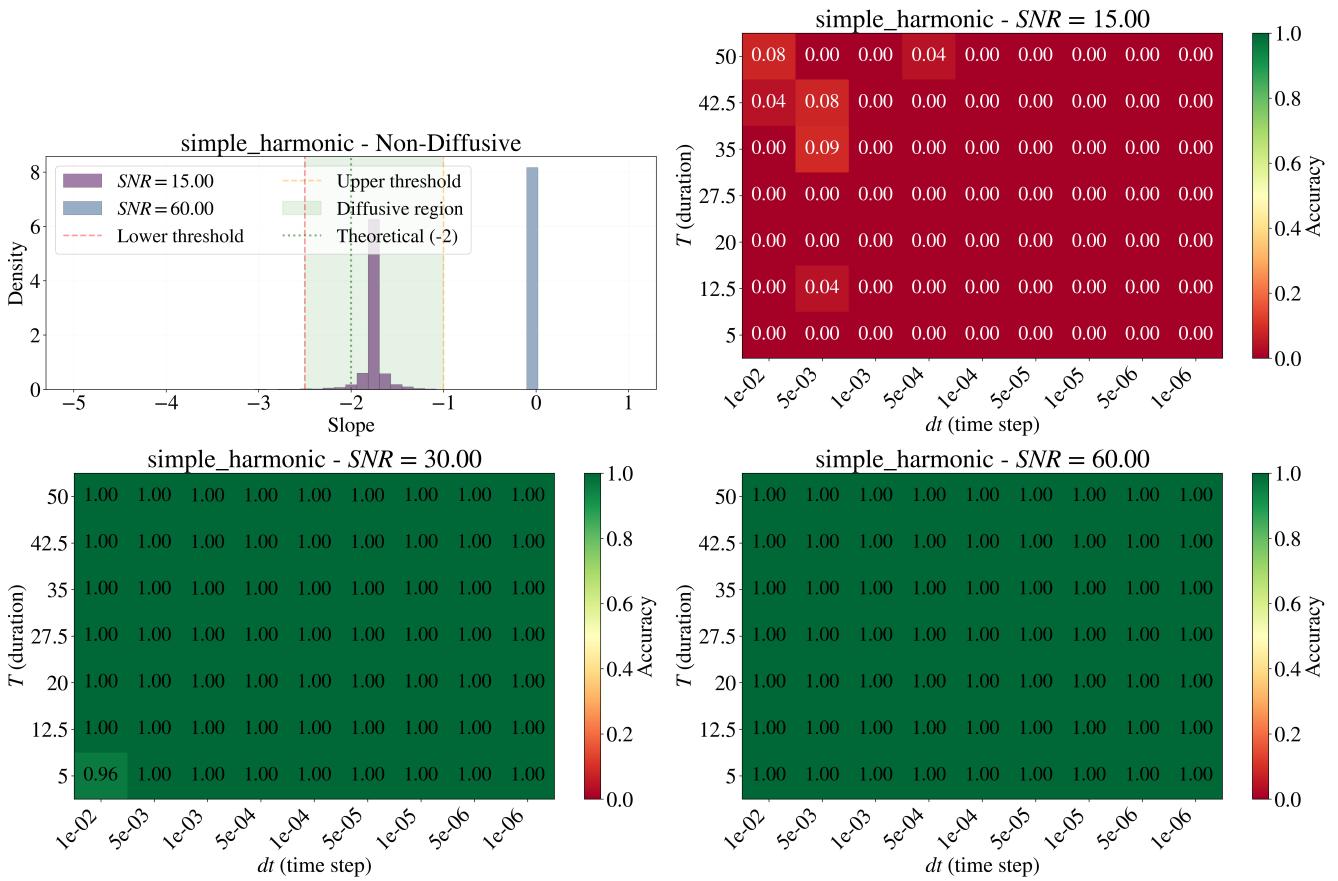
Figure 4: Heatmaps of accuracy for simple harmonic motion, across various SNR.
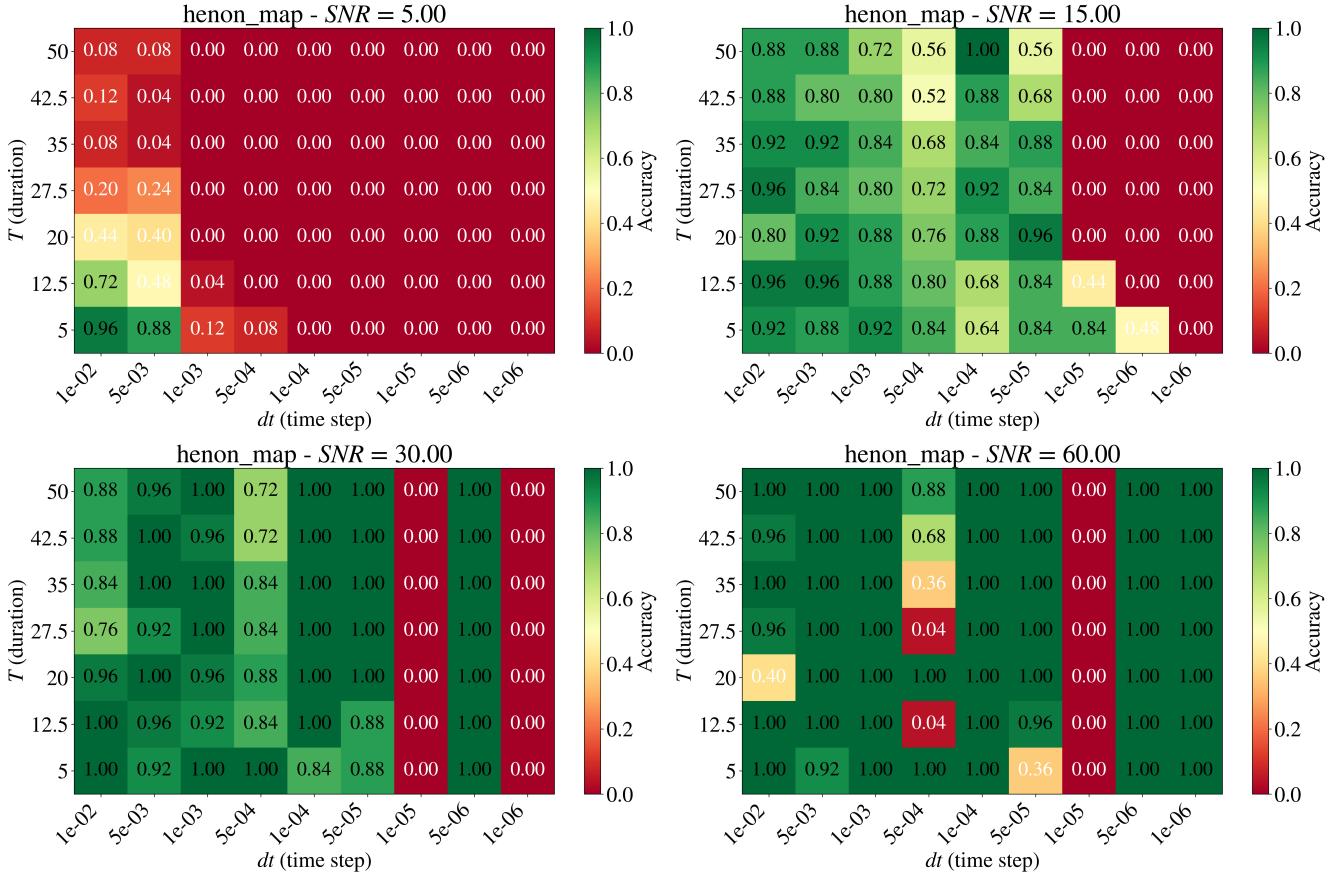
Figure 5: Heatmaps of accuracy for Henon map, across various SNR.
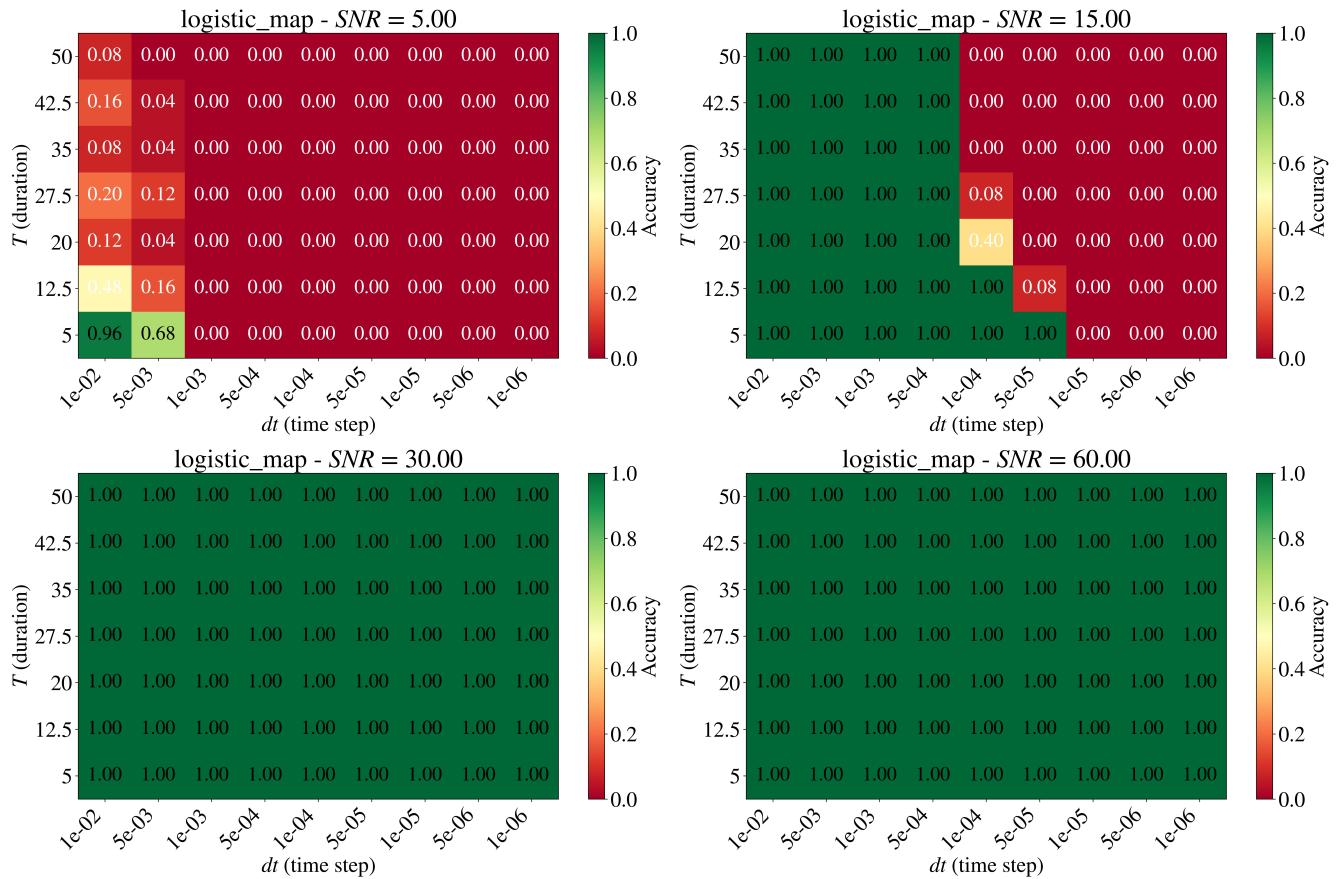
Figure 6: Heatmaps of accuracy for logistic map, across various SNR.

lcg - $SNR = 5.00$

T (duration)	1e-02	5e-03	1e-03	5e-04	1e-04	5e-05	1e-05	5e-06	1e-06
50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
42.5	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
35	0.08	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
27.5	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
20	0.08	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12.5	0.44	0.16	0.04	0.00	0.00	0.00	0.00	0.00	0.00
5	0.96	0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00

dt (time step)

lcg - $SNR = 15.00$

T (duration)	1e-02	5e-03	1e-03	5e-04	1e-04	5e-05	1e-05	5e-06	1e-06
50	0.88	0.36	0.16	0.04	0.00	0.00	0.00	0.00	0.00
42.5	0.80	0.52	0.32	0.24	0.00	0.00	0.00	0.00	0.00
35	0.72	0.68	0.40	0.36	0.00	0.00	0.00	0.00	0.00
27.5	0.80	0.72	0.32	0.64	0.00	0.00	0.00	0.00	0.00
20	0.80	0.64	0.40	0.42	0.00	0.00	0.00	0.00	0.00
12.5	0.96	0.88	0.36	0.60	0.12	0.00	0.00	0.00	0.00
5	1.00	0.88	0.72	0.52	0.40	0.17	0.00	0.00	0.00

dt (time step)

lcg - $SNR = 30.00$

T (duration)	1e-02	5e-03	1e-03	5e-04	1e-04	5e-05	1e-05	5e-06	1e-06
50	0.96	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00
42.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
35	1.00	0.92	1.00	1.00	1.00	1.00	1.00	1.00	1.00
27.5	0.92	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00
20	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00
12.5	0.88	0.96	0.92	1.00	1.00	1.00	1.00	1.00	1.00
5	1.00	0.96	1.00	1.00	0.96	1.00	1.00	1.00	1.00

dt (time step)

lcg - $SNR = 60.00$

| T (duration)	1e-02	5e-03	1e-03	5e-04	1e-04	5e-05	1e-05	5e-06	1e-06
50	1.00	0.84	1.00	1.00	1.00	1.00	1.00	1.00	1.00
42.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
35	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00
27.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
20	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
12.5	1.00	1.00	0.44	1.00	1.00	1.00	1.00	1.00	1.00
5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

dt (time step)

Figure 7: Heatmaps of accuracy for LCG map, across various SNR.

4.4 Deterministic Chaotic Systems

We next consider deterministic chaotic systems including duffing, rayleigh-duffing, chen and lu. Fig. 8 shows heatmaps of accuracy for chen and lu systems. Both systems behave similarly, and their classification is reasonably accurate even for low SNRs—granted dt and T are sufficiently large. Fig. 9 shows the accuracy heatmaps for duffing and rayleigh-duffing systems; with both systems observing the same trend—high classification accuracy beyond a critical SNR.
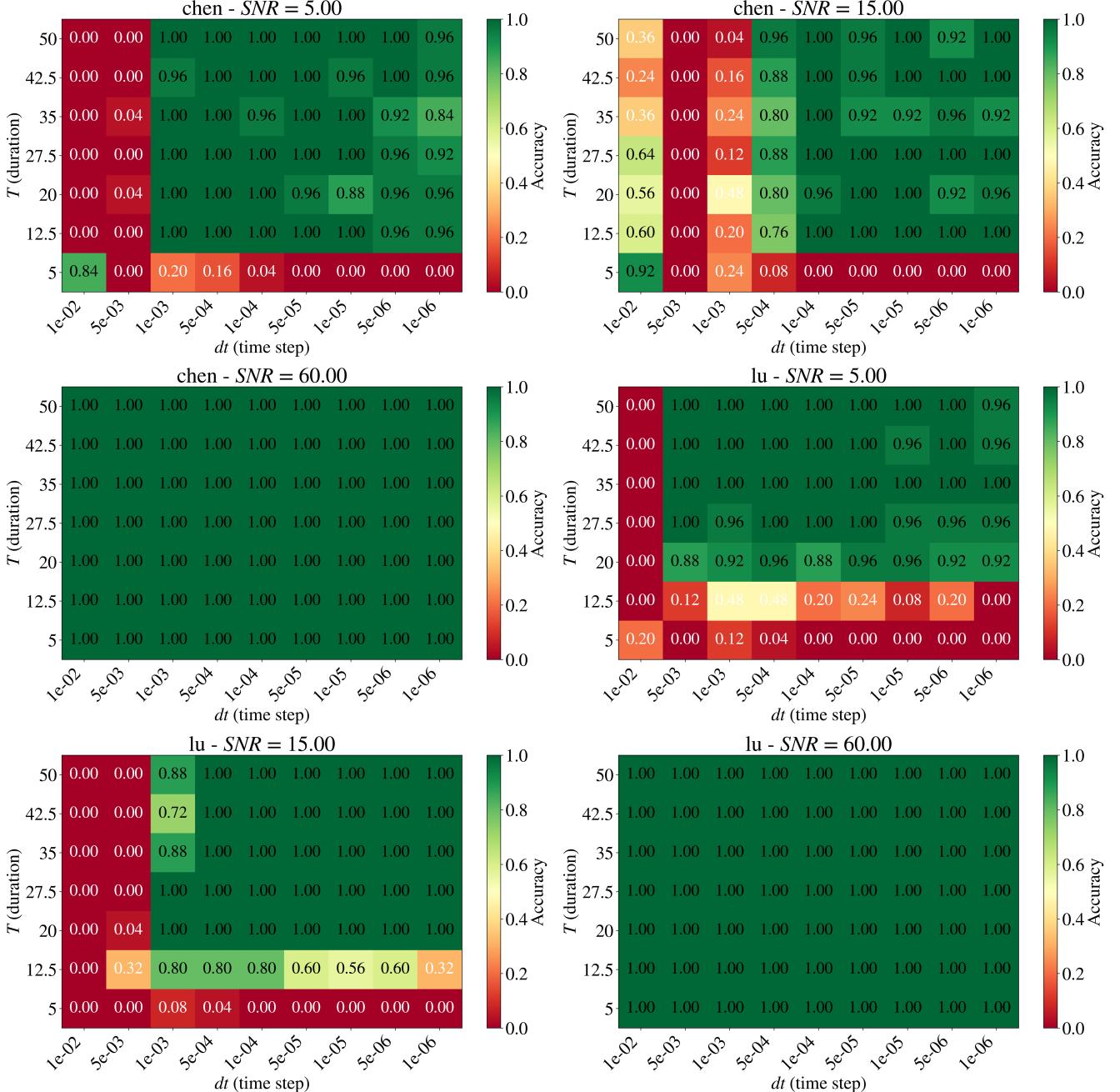
Figure 8: Heatmaps of accuracy for Chen and Lu systems, across various SNR.
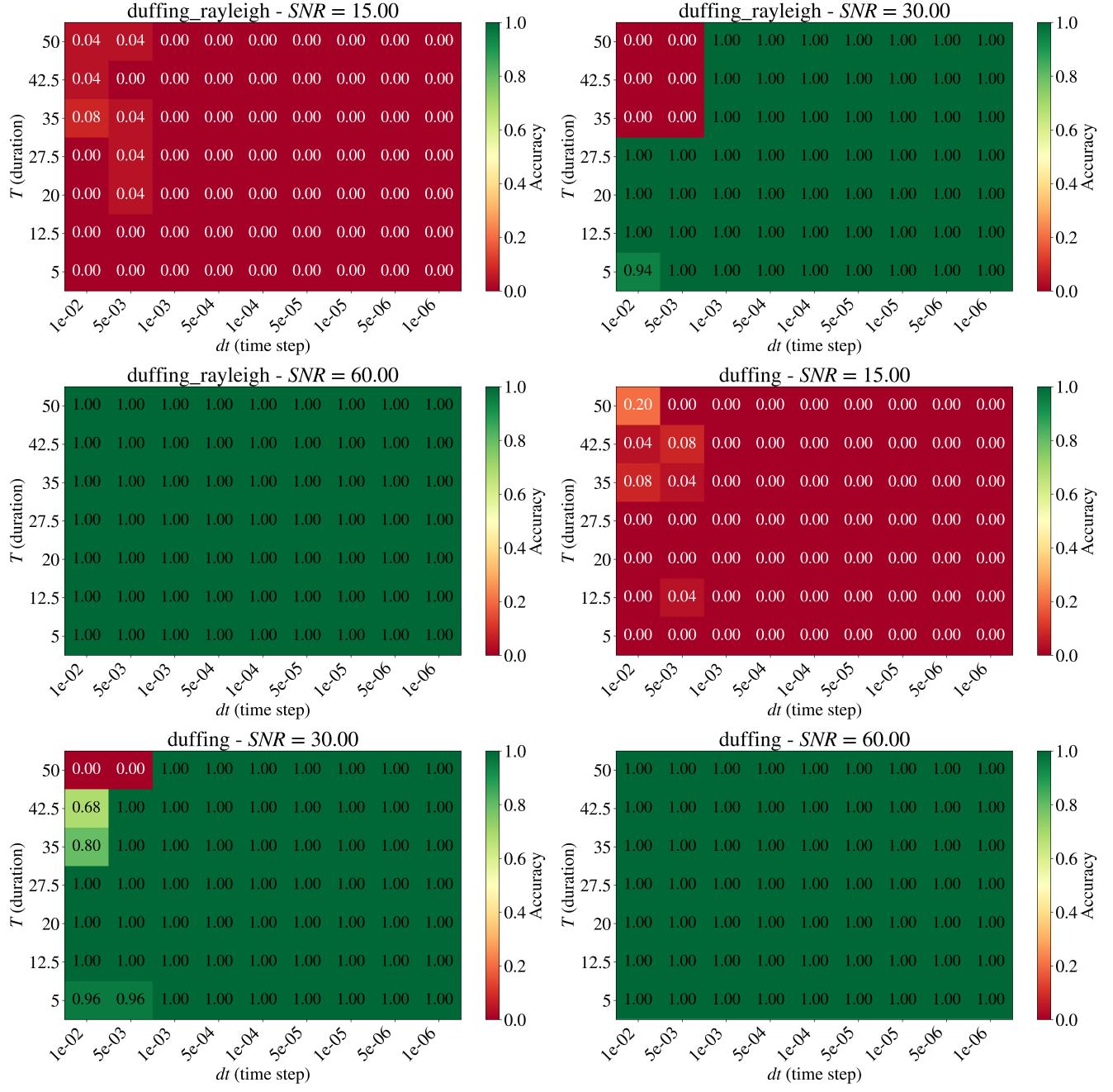
Figure 9: Heatmaps of accuracy for Duffing and Rayleigh-Duffing systems, across various SNR.

4.5 Stochastic Chaotic Systems

Further, we examine a deterministic chaotic system with intrinsic stochastic forcing—the stochastic Duffing oscillator. The stochastic Duffing system exhibits a clear monotonic increase in diffusion classification accuracy as noise intensity increases. Correspondingly, slope histograms narrow toward -2, indicating convergence toward diffusive behavior. See Fig. 10 for slope distributions and heatmaps of accuracy.
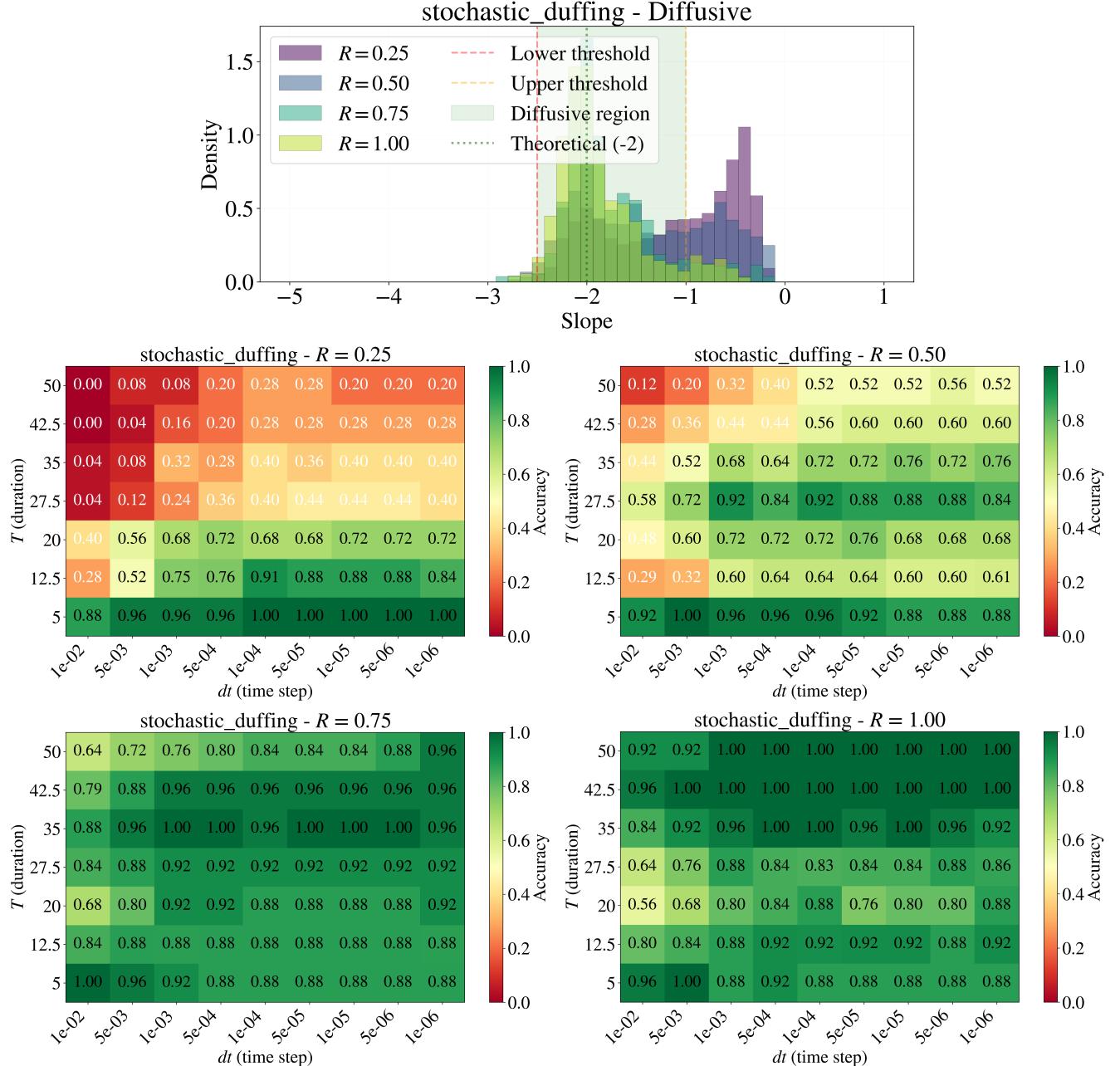
Figure 10: Slope distribution and heatmaps of accuracy for stochastic Duffing system, across various R.

4.6 Real-World Datasets

4.6.1 Financial Time Series

To demonstrate practical utility in real observational settings, we also applied the proposed method to financial market data, where differentiating diffusion-driven price variability from deterministic predictability is of fundamental interest. We analyzed minute-resolution returns of Bitcoin (BTC-USD) over a 30-day period, and daily returns of the S&P 500 exchange-traded fund (SPY) over the past ten years, see signals in Fig. 11. For each dataset, empirical returns were computed as

$$r_t = \frac{P_t - P_{t-\Delta t}}{P_{t-\Delta t}},$$

and treated as discrete observations of a stochastic process with sampling interval $\Delta t = 1$ minute (BTC) or $\Delta t = 1$ day (SPY).
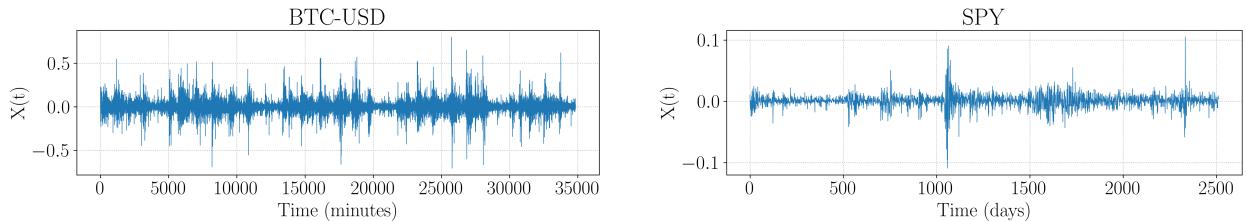
Figure 11: BTC-USD and SPY financial time series retrieved from yfinance on December 25, 2025.

For both assets, the fitted log–log slopes of the excursion count scaling law fell within the diffusion regime. Specifically, BTC: $\hat{s} \approx -2.11$, and SPY: $\hat{s} \approx -1.61$. Both values fall within $-2.5 < \hat{s} < -1.0$, indicating the presence of a nontrivial local martingale component. The Bitcoin data shows behavior closely aligned with Brownian scaling, while SPY exhibits mild deviations possibly due to the lower discretization.

Importantly, our method achieves these results without any assumptions of stationarity, absence of drift, or Gaussian increments. Unlike model-dependent diffusion diagnostics, the excursion test observes path characteristics—enabling a direct conclusion: at the scales observed, both BTC-USD and SPY returns possess locally diffusive structure with finite quadratic variation.

4.6.2 Real-world Audio Signals

We further evaluated the proposed excursion-based diffusion detector on a diverse real-world dataset of short audio signals [23]. The dataset contained 50 commonly studied sound classes (environmental, human, mechanical, and animal), each represented by 40 independent samples of 5 seconds duration each, sampled at uniform audio rate of 44.1 kHz. Amplitude envelopes of each were treated as the observed trajectory X_t, and each of the independent samples was classified as either diffusive or non-diffusive using our algorithm. Since there is no known ground truth for any of these audio signals, we compute and report the percentage of samples which got classified as diffusive for each sound type.

Diffusive behaviour was detected in a majority of sounds, including natural ambience such as rain (for which 82.5% of samples were classified diffusive), thunderstorm (90.0%), wind (77.5%), sea waves (97.5%), and biological vocalizations such as sheep bleats (92.5%), pig grunts (95.0%), and frogs (80.0%). Anthropogenic noise sources also frequently get classified: vacuum cleaner (95.0%), engine running (92.5%), chainsaw (100.0%), train (95.0%), and helicopter (87.5%). These results align with the presence of broadband stochastic components in aerodynamic and combustive processes.

More structured or transient-dominated signals show reduced diffusion classifications: mouse clicks (52.5%), clock tick (50.0%), can opening (50.0%), glass breaking (70.0%), and church bells (97.5%) occupy a broad continuum between deterministic periodicity and random excitation. Vocal sounds such as laughing (82.5%), crying baby (95.0%), and coughing (90.0%) are also predominantly classified diffusive, reflecting turbulent airflow in speech production.

Interestingly, simple and rhythmic event sequences like clock alarm (45.0%) and door knock (37.5%) exhibit less diffusive signatures likely due to repetitive deterministic structure. Overall, the results demonstrate that the proposed method is capable of distinguishing fine-grained differences in physical processes. Even across complex, real-world acoustic environments, signals exhibiting turbulent or noisy dynamics systematically trigger the characteristic $N(\varepsilon) \sim \varepsilon^{-2}$ scaling behavior, while impulsive or organized sounds do not. Therefore, this test provides a model-free pathway to partition high-dimensional multimedia signals into diffusion-driven versus deterministic regimes, offering potential applications in many fields.

5   Limitations

The proposed excursion-scaling test applies broadly to continuous semimartingales with finite quadratic variation. However, several practical considerations limit its universal deployment. First, reliable excursion statistics require fine resolution: large sampling intervals dt suppress small excursions and bias the estimate toward deterministic behavior. Similarly, low signal-to-noise environments can cause a truly deterministic signal to be misclassified. The method does not apply to jump processes, impulsive dynamics, or Levy-type models with infinite-activity discontinuities, since their excursion structure falls outside the theoretical regime. Fractional Brownian motion also lacks the finite quadratic variation needed for correct scaling. These constraints highlight that the method performs best when the underlying dynamics are continuous in time, well-resolved by sampling, and maintain sufficient variance to reveal the scaling law predicted by semimartingale theory.

6   Conclusion

We introduced a principled excursion-asymptote law for continuous semimartingales of Ito type, and demonstrated its use as a robust diagnostic to distinguish stochastic diffusion from deterministic/chaotic dynamics. Unlike many existing chaos-versus-noise discrimination methods that require extensive hyper-parameter tuning, multiple trained models, or subjective interpretations of surrogate metrics, our approach is fully self-contained and leverages a theoretically guaranteed scaling law for excursion counts. Across a broad benchmark of systems, the method exhibits strong performance—achieving near-perfect accuracy on canonical stochastic processes such as Brownian motion, Ornstein-Uhlenbeck dynamics, and Cox-Ingersoll-Ross diffusion. Deterministic oscillators and discrete maps are reliably classified when the time resolution is sufficiently fine, the time horizon is large enough to capture attractor dynamics, and the signal-to-noise ratio remains moderate to high. Stochastic chaotic systems are also correctly identified over a wide range of noise intensities when temporal resolution and duration are adequate. On real-world financial data, including minute-level cryptocurrency returns and daily equity returns, the method perfectly identified diffusive behavior consistent with market theory. Results on short environmental and human-generated audio clips further reveal interesting distinctions in how natural signals manifest stochasticity. Overall, these findings highlight the strong practical potential of the proposed diagnostic as a simple, theory-driven tool for separating stochasticity from chaotic determinism in physical, biological, and engineered systems.

7 Acknowledgements

References

[1]   N. P. Subramaniyam, J. F. Donges, and J. Hyttinen. "Signatures of chaotic and stochastic dynamics uncovered with epsilon-recurrence networks". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 471.2183 (Nov. 2015), p. 20150349. ISSN: 1471-2946. DOI: 10.1098/rspa.2015.0349. URL: http://dx.doi.org/10.1098/rspa.2015.0349.

[2]   D. M. Mateos, L. E. Riveaud, and P. W. Lamberti. "Detecting dynamical changes in time series by using the Jensen Shannon divergence". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.8 (Aug. 2017), p. 083118. ISSN: 1089-7682. DOI: 10.1063/1.4999613. URL: http://dx.doi.org/10.1063/1.4999613.

[3]   Massimiliano Zanin and Felipe Olivares. "Ordinal patterns-based methodologies for distinguishing chaos from noise in discrete time series". In: *Communications Physics* 4.1 (Aug. 2021). ISSN: 2399-3650. DOI: 10.1038/s42005-021-00696-z. URL: http://dx.doi.org/10.1038/s42005-021-00696-z.

[4]   Thiago Lima Prado, Bruno Rafael Reichert Boaretto, Gilberto Corso, Gustavo Zampier dos Santos Lima, Jürgen Kurths, and Sergio Roberto Lopes. "A direct method to detect deterministic and stochastic properties of data". In: *New Journal of Physics* 24.3 (Mar. 2022), p. 033027. ISSN: 1367-2630. DOI: 10.1088/1367-2630/ac5057. URL: http://dx.doi.org/10.1088/1367-2630/ac5057.

[5]   B. R. R. Boaretto, R. C. Budzinski, K. L. Rossi, T. L. Prado, S. R. Lopes, and C. Masoller. "Discriminating chaotic and stochastic time series using permutation entropy and artificial neural networks". In: *Scientific Reports* 11.1 (Aug. 2021). ISSN: 2045-2322. DOI: 10.1038/s41598-021-95231-z. URL: http://dx.doi.org/10.1038/s41598-021-95231-z.

[6]   Massimiliano Zanin. "Can Deep Learning distinguish chaos from noise? Numerical experiments and general considerations". In: *Communications in Nonlinear Science and Numerical Simulation* 114 (Nov. 2022), p. 106708. ISSN: 1007-5704. DOI: 10.1016/j.cnsns.2022.106708. URL: http://dx.doi.org/10.1016/j.cnsns.2022.106708.

[7]   L. Zunino, M. C. Soriano, and O. A. Rosso. "Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach". In: *Physical Review E* 86.4 (Oct. 2012). ISSN: 1550-2376. DOI: 10.1103/physreve.86.046210. URL: http://dx.doi.org/10.1103/PhysRevE.86.046210.

[8]   Yoshito Hirata. "Recurrence plots for characterizing random dynamical systems". In: *Communications in Nonlinear Science and Numerical Simulation* 94 (Mar. 2021), p. 105552. ISSN: 1007-5704. DOI: 10.1016/j.cnsns.2020.105552. URL: http://dx.doi.org/10.1016/j.cnsns.2020.105552.

[9]   F. Olivares, M. Zanin, L. Zunino, and D. G. Pérez. "Contrasting chaotic with stochastic dynamics via ordinal transition networks". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 30.6 (June 2020). ISSN: 1089-7682. DOI: 10.1063/1.5142500. URL: http://dx.doi.org/10.1063/1.5142500.

[10] Reik V Donner, Yong Zou, Jonathan F Donges, Norbert Marwan, and Jurgen Kurths. "Recurrence networks—a novel paradigm for nonlinear time series analysis". In: *New Journal of Physics* 12.3 (Mar. 2010), p. 033025. ISSN: 1367-2630. DOI: 10.1088/1367-2630/12/3/033025. URL: http://dx.doi.org/10.1088/1367-2630/12/3/033025.

[11] Lucas Lacasa, Bartolo Luque, Fernando Ballesteros, Jordi Luque, and Juan Carlos Nuño. "From time series to complex networks: The visibility graph". In: *Proceedings of the National Academy of Sciences* 105.13 (Apr. 2008), 4972–4975. ISSN: 1091-6490. DOI: 10.1073/pnas.0709247105. URL: http://dx.doi.org/10.1073/pnas.0709247105.

[12] Zelin Zhang, Jun Wu, Yufeng Chen, Ji Wang, and Jinyu Xu. "Distinguish between Stochastic and Chaotic Signals by a Local Structure-Based Entropy". In: *Entropy* 24.12 (Nov. 2022), p. 1752. ISSN: 1099-4300. DOI: 10.3390/e24121752. URL: http://dx.doi.org/10.3390/e24121752.

[13] Daniel Perez. "On the persistent homology of almost surely C^0 stochastic processes". In: *Journal of Applied and Computational Topology* 7.4 (July 2023), pp. 879–906. ISSN: 2367-1734. DOI: 10.1007/s41468-023-00132-x. URL: http://dx.doi.org/10.1007/s41468-023-00132-x.

[14] *Applications of Random Process Excursion Analysis*. Elsevier, 2013. ISBN: 9780124095014. DOI: 10.1016/c2012-0-06107-8. URL: http://dx.doi.org/10.1016/C2012-0-06107-8.

[15] Robert M. Blumenthal. *Excursions of Markov Processes*. Birkhäuser Boston, 1992. ISBN: 9781468494129. DOI: 10.1007/978-1-4684-9412-9. URL: http://dx.doi.org/10.1007/978-1-4684-9412-9.

[16] S. O. Rice. "Mathematical Analysis of Random Noise". In: *Bell System Technical Journal* 23.3 (July 1944), 282–332. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1944.tb00874.x. URL: http://dx.doi.org/10.1002/j.1538-7305.1944.tb00874.x.

[17] Kiyosi Ito. "The expected number of zeros of continuous stationary Gaussian processes". In: *Kyoto Journal of Mathematics* 3.2 (Jan. 1963). ISSN: 2156-2261. DOI: 10.1215/kjm/1250524817. URL: http://dx.doi.org/10.1215/kjm/1250524817.

[18] N. Donald Ylvisaker. "The Expected Number of Zeros of a Stationary Gaussian Process". In: *The Annals of Mathematical Statistics* 36.3 (June 1965), 1043–1046. ISSN: 0003-4851. DOI: 10.1214/aoms/1177700077. URL: http://dx.doi.org/10.1214/aoms/1177700077.

[19] M. R. Leadbetter. "On crossings of arbitrary curves by certain Gaussian processes". In: *Proceedings of the American Mathematical Society* 16.1 (Feb. 1965), 60–68. ISSN: 1088-6826. DOI: 10.1090/s0002-9939-1965-0170382-6. URL: http://dx.doi.org/10.1090/S0002-9939-1965-0170382-6.

[20] M. R. Leadbetter and J. D. Cryer. "On the Mean Number of Curve Crossings by Non-Stationary Normal Processes". In: *The Annals of Mathematical Statistics* 36.2 (Apr. 1965), 509–516. ISSN: 0003-4851. DOI: 10.1214/aoms/1177700160. URL: http://dx.doi.org/10.1214/aoms/1177700160.

[21] V. A. Ivanov. "On the Average Number of Crossings of a Level by the Sample Functions of a Stochastic Process". In: *Theory of Probability & Its Applications* 5.3 (Jan. 1960), 319–323. ISSN: 1095-7219. DOI: 10.1137/1105031. URL: http://dx.doi.org/10.1137/1105031.

[22] Zekai Sen. "Probabilistic modelling of crossing in small samples and application of runs to hydrology". In: *Journal of Hydrology* 124.3-4 (May 1991), pp. 345–362. ISSN: 0022-1694. DOI: 10.1016/0022-1694(91)90023-b. URL: http://dx.doi.org/10.1016/0022-1694(91)90023-B.

[23] Karol J. Piczak. "ESC: Dataset for Environmental Sound Classification". In: *Proceedings of the 23rd Annual ACM Conference on Multimedia*. Brisbane, Australia: ACM Press, Oct. 13, 2015, pp. 1015–1018. ISBN: 978-1-4503-3459-4. DOI: 10.1145/2733373.2806390. URL: http://dl.acm.org/citation.cfm?doid=2733373.2806390.

A Benchmark System Definitions

All benchmark signals used in this study are generated as one-dimensional observables from standard deterministic chaotic systems, continuous-time stochastic differential equations (SDEs), and hybrid stochastic-chaotic dynamics. Deterministic ODEs are integrated using adaptive Bogacki-Shampine RK23 with fixed output grid dt, while all SDEs are discretized using Euler-Maruyama. Maps are iterated directly. In every case, only the first state component $x(t)$ is recorded as the output. Here, W_t denotes a standard Wiener process, and $R > 0$ scales noise intensity via $\sigma = 1/R$.

Simple harmonic oscillator:
$$\ddot{x} + \omega^2 x = 0, \qquad \omega = 1.0. \tag{2}$$

A.1 Chaotic Attractors

Duffing oscillator:

$$\dot{x} = v, \tag{3}$$
$$\dot{v} = -\delta v - \alpha x - \beta x^3 + \gamma \cos(\omega t), \tag{4}$$

with $\delta = 0.2$, $\alpha = -1$, $\beta = 1$, $\gamma = 0.3$, $\omega = 1.2$.

Rayleigh-Duffing oscillator:

$$\dot{x} = v, \tag{5}$$
$$\dot{v} = -\delta v - \varepsilon \left(v - v^3\right) - \alpha x - \beta x^3 + \gamma \cos(\omega t), \tag{6}$$

with parameters $\delta = 0.2, \alpha = -1, \beta = 1, \gamma = 0.3, \omega = 1.2, \varepsilon = 0.1$.

Chen attractor:

$$\dot{x} = a(y - x), \tag{7}$$
$$\dot{y} = (c - a)x - xz + cy, \tag{8}$$
$$\dot{z} = xy - bz, \tag{9}$$

with $(a, b, c) = (35, 3, 28)$.

Lu attractor:

$$\dot{x} = a(y - x), \tag{10}$$
$$\dot{y} = -xz + cy, \tag{11}$$
$$\dot{z} = xy - bz, \tag{12}$$

with $(a, b, c) = (36, 3, 20)$.

A.2 Chaotic Maps

For all maps, x_{k+1} is observed at discrete times $t_{k+1} = t_k + dt$.

Logistic map:

$$x_{k+1} = rx_k(1 - x_k), \qquad r = 4.0. \tag{13}$$

Henon map:

$$x_{k+1} = 1 - ax_k^2 + y_k, \tag{14}$$
$$y_{k+1} = bx_k, \tag{15}$$

with $(a, b) = (1.4, 0.3)$.

Linear Congruential Generator (LCG):

$$x_{k+1} = (ax_k + c) \bmod m, \qquad a = 1664525, \ c = 1013904223, \ m = 2^{32}. \tag{16}$$

A.3 Stochastic Diffusions

These are simulated as Ito SDEs,

$$dX_t = f(X_t, t) \, dt + g(X_t, t) \, dW_t, \qquad g = \sigma = \frac{1}{R}.$$

Brownian motion:

$$dX_t = \sigma \, dW_t. \tag{17}$$

Ornstein-Uhlenbeck (OU) process:

$$dX_t = \theta(\mu - X_t) \, dt + \sigma \, dW_t, \qquad \theta = 1, \ \mu = 0. \tag{18}$$

Cox-Ingersoll-Ross (CIR) process:

$$dX_t = \kappa(\theta - X_t) \, dt + \sigma \sqrt{X_t} \, dW_t, \qquad X_t \geq 0. \tag{19}$$

Stochastic Duffing:

$$dX = V \, dt + \sigma \, dW_t, \tag{20}$$
$$dV = \left(-0.2V + X + X^3 + 0.3\cos(1.2t)\right)dt + \sigma \, dW_t, \tag{21}$$

using the same scalar Wiener increment in both equations (rank-1 diffusion).