Title

Sufficient statistics for parameter inference in the stochastic Van der Pol oscillator model for thermoacoustic instabilities

Journal

Physical Review E, 114(1)

ISSN

2470-0045

Author

Ju, Y Sungtaek

Publication Date

2026-07-07

DOI

10.1103/k56x-bf4g

Sufficient statistics for parameter inference in the stochastic Van der Pol oscillator model for thermoacoustic instabilities

Y. Sungtaek Ju*

Department of Mechanical and Aerospace Engineering,

University of California, Los Angeles, California 90095, USA

(Dated: May 13, 2026)

Abstract

Thermoacoustic instabilities pose a critical challenge in the design and operation of many combustion systems. We report a low-dimensional statistics approach for inferring the parameters of the stochastic Van der Pol oscillator from trajectory observations. From the exponential family structure of the stationary amplitude distribution, we identify sufficient statistics that capture essential parametric information. To break the parameter degeneracy inherent to stationary distribution methods, we incorporate the amplitude autocorrelation time, which provides independent information about noise intensity through relaxation dynamics. The inverse mapping from sufficient statistics to parameters can be effectively approximated by quadratic polynomial regression, achieving comparable accuracy to neural networks while requiring orders of magnitude fewer parameters and providing interpretable closed-form expressions. Comparison with previous parameter identification approaches reveals complementary strengths. In the critical regime near the Hopf bifurcation, our method enables regime classification with reasonable accuracy. Under measurement noise, a periodogram-based moment correction yields graceful degradation across all three parameter estimates, with residual high-noise error dominated by autocorrelation-time bias rather than moment bias. The methodology provides interpretable symbolic expressions suitable for thermoacoustic stability monitoring, and can be extended to other stochastic oscillator systems.

I. INTRODUCTION

A. Thermoacoustic instabilities and reduced-order models

Thermoacoustic instabilities pose a critical challenge in the design and operation of modern gas turbines, rocket engines, and other combustion systems [1, 2]. These self-excited oscillations arise from the coupling between acoustic pressure fluctuations and unsteady heat release, leading to structural damage, operational constraints, and potentially catastrophic failure [3]. This has motivated a sustained research effort towards understanding, predicting, and controlling these phenomena.

The development of reduced-order models that capture essential thermoacoustic physics while remaining computationally tractable is central to this effort. Among such models, the

* sungtaek.ju@ucla.edu

stochastic Van der Pol (VdP) oscillator has proven valuable in describing noise-driven limit-cycle dynamics observed in turbulent combustors [4, 5]. In this framework, the amplitude of the acoustic pressure evolves according to a nonlinear oscillator equation with additive stochastic forcing that represents the turbulence-induced fluctuations inherent in practical combustion systems.

The stochastic VdP model captures several key features of thermoacoustic dynamics: the supercritical Hopf bifurcation from stable to unstable operation, amplitude saturation through nonlinear damping, and the noise-induced fluctuations around the deterministic limit cycle. The model parameters encode essential physics: the linear growth rate determines the proximity to the stability boundary, the nonlinear damping coefficient governs the limit-cycle amplitude, and the noise intensity reflects the turbulent forcing environment. Accurate identification of these parameters from experimental pressure measurements facilitates the evaluation of the stability margin, the design of the control system, and the validation of higher-fidelity computational models.

Beyond thermoacoustics, the stochastic VdP model finds applications across various domains including cardiac rhythm dynamics [6, 7], neural oscillations [8, 9], auditory hair cell mechanics [10, 11], aeroelastic flutter [12], and general relaxation oscillator systems [13]. The ubiquity of this model underscores the importance of robust parameter identification methods.

B. The parameter identification problem

Given time-series observations of a stochastic VdP oscillator, the inverse problem is to recover the governing parameters $\boldsymbol{\theta} = (\varepsilon, \alpha, d)$, representing linear damping, nonlinear saturation, and noise intensity, respectively. This problem is fundamentally challenging for several reasons.

First, finite observation times introduce statistical variability in any estimator. The problem is particularly acute for slowly decorrelating systems near the stability boundary, where the oscillation envelope evolves on timescales comparable to or exceeding typical measurement durations [14]. Second, methods requiring differentiation of noisy signals suffer accuracy degradation, as differentiation amplifies high-frequency noise while the signal of interest resides at lower frequencies [15, 16]. Third, as demonstrated by Boujo and Noiray [17], the

stationary probability density of the amplitude depends only on two combinations of the three physical parameters, creating a degeneracy that may not be resolved by collecting more data or applying more sophisticated optimization algorithms to amplitude statistics alone.

These challenges have motivated the development of various methodological approaches, which we review before presenting our contribution.

C. Fokker–Planck-based methods

The prevailing approach for thermoacoustic parameter identification operates in amplitude-phase coordinates via the Fokker–Planck equation. The foundational approach [4] recognizes that under stochastic averaging the amplitude evolves as an Itô diffusion with drift and diffusion coefficients that are explicit functions of the physical parameters. Parameters are then identified by fitting the stationary probability density function (PDF) or the Kramers–Moyal coefficients extracted from finite-time transition statistics.

Noiray and Denisov [14] refined this approach by developing extrapolation procedures for short-time Kramers–Moyal coefficients to estimate their instantaneous limits, mitigating finite-time bias that arises from discrete sampling. Related work on finite sampling interval effects in Kramers–Moyal analysis [18, 19] has developed methods to correct for systematic errors at low sampling rates.

Boujo and Noiray [17] developed an adjoint Fokker–Planck method that avoids finite-time extrapolation by directly optimizing parameters to match the predicted stationary distribution with observations. Their optimization-based framework iteratively adjusts parameters to minimize discrepancy between predicted and measured statistics, with gradients computed efficiently via adjoint equations. Critically, this work explicitly noted that the stationary PDF depends only on the ratios ε/d and α/d, and not the three parameters individually. This observation motivates our work.

Bonciolini *et al.* [20] extended the framework to coloured noise forcing, demonstrating that accurate identification requires appropriate band-pass filtering when the noise correlation time is non-negligible. Lee *et al.* [21, 22] applied adjoint-based optimization with gradient computation via automatic differentiation.

These methods have also been extended to more complex configurations. For azimuthal

thermoacoustic modes in annular combustors, Indlekofer *et al.* [23] employed Langevin-based regression to identify parameters from experimental pressure data. Their approach exploits the stochastic wave equation governing amplitude and spinning dynamics, fitting drift and diffusion coefficients to extract the linear growth rate, nonlinear saturation, and symmetry-breaking parameters.

D. Neural networks

Recent work has applied neural networks to parameter identification in stochastic oscillators. Son and Lee [24] proposed a Physics-Informed Neural Networks (PINNs) architecture that combines a Fokker–Planck residual loss with negative log-likelihood loss for the amplitude PDF. Their network learns to approximate the stationary PDF while satisfying the governing equation as a soft constraint. On benchmark problems spanning stable and unstable regimes, they report mean parameter errors of approximately 6% for ε and d, but very large errors ($\sim$50%) for α.

Xie *et al.* [25] reported the prediction of bifurcation characteristics and amplitude death phenomena from PINN-derived VdP parameters, demonstrating practical utility for combustor design and coupled-oscillator analysis. Mahesh *et al.* [26] applied PINNs to thermoacoustic interactions in bluff-body combustors, using a coupled acoustic-VdP model to reconstruct pressure fields from sparse measurements while simultaneously estimating model parameters.

Sugandi *et al.* [27] proposed an alternative neural approach using Deep Operator Networks (DeepONets) to learn the mapping from drift coefficients to finite-time Kramers–Moyal coefficients. This surrogate model enables rapid forward evaluation, accelerating Bayesian inference via Markov Chain Monte Carlo sampling. While achieving approximately two orders of magnitude speedup over finite-difference adjoint Fokker–Planck solutions, the method still requires iterative optimization.

Despite their successes, many of existing neural network-based approaches often face inherent limitations. Physics enters only through the loss function as soft constraints, requiring the network to learn the geometric structure of the solution manifold from scratch. This demands substantial training data and careful hyperparameter tuning, particularly for balancing multiple loss terms with competing gradients. The risk of converging to de-

generate solutions, such as probability densities that approach zero everywhere, requires additional regularization. On the other hand, the Fokker–Planck regularization provides strong constraints that enable accurate estimation even with limited trajectory data, as we demonstrate in our comparisons.

E. Data assimilation approaches

An alternative paradigm for thermoacoustic parameter identification uses data assimilation techniques that sequentially update model estimates as new observations become available. Nóvoa and Magri [28] introduced real-time thermoacoustic data assimilation using an ensemble square-root Kalman filter (EnSRKF) to infer both thermoacoustic states and parameters on the fly. Their method combines a physical low-order model forecast with noisy pressure observations through Bayesian updating, demonstrating successful learning across all nonlinear thermoacoustic regimes from limit cycles to chaos. This real-time sequential assimilation approach has been further extended to digital twin frameworks [29].

The data assimilation and sufficient statistics approaches involve different trade-offs suited to different operational contexts. Data assimilation requires a validated dynamical forward model and must propagate an ensemble of trajectories (typically 40 or more members) at each observation time. At these extra costs, it provides sequential parameter tracking and natural uncertainty quantification through ensemble spread, which the stationary statistics approaches do not directly provide. The data assimilation method, however, is sensitive to initial parameter estimates and dynamic models and requires careful tuning of filter parameters to avoid divergence in strongly nonlinear regimes.

In contrast, sufficient-statistics approaches require no forward model integration. The parameters are obtained by evaluating closed-form polynomial expressions from accumulated amplitude statistics. This yields a computationally negligible inference step (microseconds versus repeated ensemble integrations) and eliminates sensitivity to initialization. These approaches, however, assume quasi-stationary parameters over the observation window and do not inherently provide uncertainty estimates.

The two approaches are thus complementary: data assimilation excels when parameters drift in time and instantaneous tracking is required, while sufficient statistics provides an efficient diagnostic tool when parameters are quasi-steady and compact interpretable

expressions are desired for stability monitoring.

F. Contributions of the present work

We present a sufficient statistics approach that provides a principled framework for understanding and addressing the parameter identification problem. The main contributions are as follows.

First, we demonstrate that τ_{corr}, the amplitude autocorrelation time, provides independent information about d that breaks the parameter degeneracy. This information derives from the Fokker–Planck relaxation dynamics rather than the stationary distribution and thereby provides orthogonal constraints. Second, we show that quadratic polynomial regression with only 30 coefficients achieves performance matching that of neural networks, demonstrating that the inverse problem has low intrinsic dimensionality once the right features are identified. Third, we provide comparison with the previous PINN method at equal trajectory lengths, revealing complementary strengths: the previous PINNs excel at estimating ε (2.3× better) while our approach excels at estimating α (12× better). Fourth, we characterize behavior in the critical regime $|\varepsilon| \leq 0.05$, demonstrating robust performance and reliable regime classification. Fifth, we show that a periodogram-based moment correction yields graceful degradation across all three parameter estimates under measurement noise, with residual high-noise error attributable to autocorrelation-time bias rather than moment bias.

II. PROBLEM FORMULATION

A. The stochastic Van der Pol oscillator

The stochastic Van der Pol oscillator is governed by the second-order stochastic differential equation (SDE):

$$\ddot{x} + \omega_0^2 x = (\varepsilon + \alpha x^2)\dot{x} + \sqrt{2d}\,\xi(t), \tag{1}$$

where $x(t)$ is the oscillator displacement (acoustic pressure in thermoacoustic applications), ω_0 is the natural frequency, ε is the linear growth rate, α is the cubic nonlinear coefficient, d is the noise intensity, and $\xi(t)$ is Gaussian white noise with $\langle \xi(t)\xi(t') \rangle = \delta(t-t')$. This is the

classical stochastic Van der Pol form used in the thermoacoustic literature [4, 17, 20] and adopted by Son and Lee [24], our principal benchmark. The sign conventions follow Noiray and Schuermans [4]: $\varepsilon < 0$ corresponds to positive damping (stable fixed point), while $\varepsilon > 0$ corresponds to negative damping (unstable fixed point leading to limit cycle oscillations); $\alpha < 0$ provides cubic saturation of the limit cycle. Throughout this work we set $\omega_0 = 2\pi$ in keeping with the Son and Lee benchmark; the methodology is invariant under rescaling of ω_0.

Nondimensionalization and timescales. Time in Eq. (1) has been nondimensionalized by ω_0^{-1}, so that one oscillation period corresponds to $2\pi/\omega_0$ dimensionless time units. With $\omega_0 = 2\pi$ and our integration step $\mathrm{d}t = 0.01$, a trajectory of length $T = 100$ contains approximately 16 oscillation cycles, while $T = 500$ contains approximately 80. The amplitude autocorrelation time τ_{corr} ranges from roughly 3 to 12 dimensionless units across our parameter set, so trajectories of length $T = 100$ correspond to roughly 8–30 correlation times and $T = 500$ to 40–170 correlation times, ensuring statistical stationarity for moment estimation.

B. Amplitude equation via stochastic averaging

Under the assumption of weak nonlinearity and slow amplitude variation ($|\varepsilon|, |\alpha|a^2, d \ll \omega_0^2$), the method of stochastic averaging [30–33] reduces the two-dimensional system to a one-dimensional amplitude equation. Introducing amplitude-phase coordinates via $x = a\cos\phi$ and $\dot{x} = -a\omega_0 \sin\phi$, the averaged amplitude dynamics becomes:

$$\mathrm{d}a = \left(\frac{\varepsilon}{2}a + \frac{\alpha}{8}a^3 + \frac{d}{2\omega_0^2 a}\right)\mathrm{d}t + \frac{\sqrt{d}}{\omega_0}\,\mathrm{d}W_t, \tag{2}$$

where W_t is a standard Wiener process. This Itô SDE has drift coefficient $\mu(a) = (\varepsilon/2)a + (\alpha/8)a^3 + d/(2\omega_0^2 a)$ and diffusion coefficient $\sigma^2(a) = d/\omega_0^2$, matching Son and Lee [24] Eq. (2.3) after averaging. The drift coefficients $\varepsilon/2$ and $\alpha/8$ come from period-averaging $\sin^2\theta$ and $\cos^2\theta\sin^2\theta$; the $d/(2\omega_0^2 a)$ term is the Itô correction from amplitude–phase coordinates; the noise coefficient $\sqrt{d}/\omega_0$ arises from projecting the additive forcing onto the slow amplitude.

C. Stationary distribution and exponential family structure

The stationary probability density of the amplitude satisfies the time-independent Fokker–Planck equation [33, 34]:

$$0 = -\frac{\partial}{\partial a}[\mu(a)p_{\mathrm{s}}(a)] + \frac{1}{2}\frac{\partial^2}{\partial a^2}[\sigma^2(a)p_{\mathrm{s}}(a)]. \tag{3}$$

Solving the zero-current condition for Eq. (2), the stationary PDF takes the form:

$$p_{\mathrm{s}}(a) = \mathcal{N} \cdot a \cdot \exp\left[\frac{\varepsilon\omega_0^2}{2d}a^2 + \frac{\alpha\omega_0^2}{16d}a^4\right], \tag{4}$$

where $\mathcal{N}$ is the normalization constant ensuring $\int_0^\infty p_{\mathrm{s}}(a)\,\mathrm{d}a = 1$.

This distribution belongs to the exponential family with natural parameters $\eta_1 = \varepsilon\omega_0^2/(2d)$, $\eta_2 = \alpha\omega_0^2/(16d)$ and sufficient statistics $T_1 = a^2$, $T_2 = a^4$. By the factorization theorem, (T_1, T_2) are jointly sufficient for the parameter combinations $(\varepsilon\omega_0^2/d, \alpha\omega_0^2/d)$. For fixed ω_0, this reduces to sufficiency for the ratios $(\varepsilon/d, \alpha/d)$ as in the standard formulation.

D. The identifiability problem

The exponential family structure reveals a fundamental limitation: the stationary distribution depends on ε, α, d only through the natural parameters (η_1, η_2), which at fixed ω_0 are equivalent to the ratios ε/d and α/d. Consequently, any method based solely on fitting the stationary PDF, whether by moment matching, maximum likelihood, or neural density estimation, can identify at most two independent parameter combinations from amplitude data.

This parameter degeneracy was explicitly noted by Boujo and Noiray [17]. Identifying unambiguously the three parameters requires at least one additional constraint derived from dynamical information beyond the stationary distribution. They suggested velocity increment analysis as one possibility, but noted its noise sensitivity.

Methods that do achieve identification (such as adjoint Fokker–Planck with time-dependent data, or velocity increment analysis) implicitly or explicitly use dynamical information beyond the equilibrium distribution.

III. THEORETICAL FRAMEWORK

A. Equivalence of formulations and the adjoint operator

For clarity, we briefly state the relationships among the formulations used in the rest of this work. The Itô SDE (1) and the Fokker–Planck equation (3) describe the same stochastic process: under the standard Itô calculus correspondence, the SDE $\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t$ has a probability density $p(x,t)$ obeying the Fokker–Planck equation $\partial_t p = -\partial_x[\mu p] + \frac{1}{2}\partial_x^2[\sigma^2 p]$ [32–34]. The Langevin equation, in which the noise is written as a fluctuating force $\xi(t)$ with $\langle \xi(t)\xi(t')\rangle = \delta(t-t')$, is mathematically equivalent to the SDE form with the Itô interpretation.

Under the additional assumption of weak nonlinearity and slow amplitude variation ($|\varepsilon|, |\alpha|a^2, d \ll \omega_0^2$), the method of stochastic averaging maps the two-dimensional $(x, \dot{x})$ system to the one-dimensional amplitude SDE (2), whose associated Fokker–Planck equation governs the amplitude probability density. The averaging procedure introduces a controlled approximation whose accuracy improves as the timescale separation between the carrier oscillation ($\sim 2\pi/\omega_0$) and the amplitude relaxation (τ_{corr}) increases. Throughout this work, we operate with $\tau_{\mathrm{corr}} \gtrsim 2\pi/\omega_0$, well within the regime of validity.

The Fokker–Planck operator $\mathcal{L}$ governing the amplitude dynamics is

$$\mathcal{L}p = -\frac{\partial}{\partial a}\big[\mu(a)p\big] + \frac{1}{2}\frac{\partial^2}{\partial a^2}\big[\sigma^2(a)p\big], \qquad \sigma^2(a) = d/\omega_0^2. \tag{5}$$

Its L^2 adjoint, with respect to the stationary measure p_{s}, is

$$\mathcal{L}^\dagger f = \mu(a)\,\frac{\partial f}{\partial a} + \frac{1}{2}\sigma^2(a)\,\frac{\partial^2 f}{\partial a^2}, \tag{6}$$

and governs the time evolution of expectation values: $\partial_t \mathbb{E}[f(a_t)] = \mathbb{E}[(\mathcal{L}^\dagger f)(a_t)]$. The stationary density p_{s} lies in $\ker \mathcal{L}$, while time-correlation functions, including the amplitude autocorrelation $C(\tau)$ in Sec. III E, evolve under $\mathcal{L}^\dagger$. This distinction underlies our main theoretical observation: (T_1, T_2) are sufficient for the kernel of $\mathcal{L}$, while τ_{corr} encodes the leading nontrivial eigenvalue of $\mathcal{L}^\dagger$, providing strictly orthogonal information.

B. Sufficient statistics for the stationary distribution

From the exponential family representation (4), the sufficient statistics for the natural parameters $(\eta_1, \eta_2) = (\varepsilon\omega_0^2/2d, \alpha\omega_0^2/16d)$ are [35]:

$$T_1 = \langle a^2 \rangle, \quad T_2 = \langle a^4 \rangle, \tag{7}$$

where $\langle \cdot \rangle$ denotes expectation under the stationary distribution. These moments capture all information about the natural parameters contained in the amplitude observations. Higher moments $\langle a^{2k} \rangle$ for $k \geq 3$ do not provide additional parametric information; they are determined by T_1 and T_2 through the exponential family structure.

C. Breaking the degeneracy: the role of τ_{corr}

To recover all three parameters (ε, α, d) individually, we require additional information beyond the stationary distribution. We propose using the amplitude autocorrelation time τ_{corr}, defined as:

$$\tau_{\text{corr}} = \int_0^\infty C(\tau)\, \mathrm{d}\tau, \quad C(\tau) = \frac{\langle a(t)a(t+\tau)\rangle - \langle a \rangle^2}{\mathrm{Var}[a]}. \tag{8}$$

The correlation time characterizes how quickly the amplitude process forgets its initial condition. This information is encoded in the transient dynamics of the Fokker–Planck equation, not its stationary solution. The Fokker–Planck operator $\mathcal{L}$ governing the amplitude dynamics is given by Eq. (5) with $\sigma^2(a) = d/\omega_0^2$.

Spectral decomposition of the autocorrelation. The Fokker–Planck operator $\mathcal{L}$ for a diffusion in a confining potential has a discrete real spectrum with eigenvalues $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$. The corresponding eigenfunctions of $\mathcal{L}^\dagger$, $\{\phi_n\}$, form a complete basis, and any observable $f(a)$ with $\langle f \rangle_{\text{s}} = 0$ admits the expansion $f(a) = \sum_{n\geq 1} c_n \phi_n(a)$, giving the spectral decomposition of the autocorrelation

$$C(\tau) = \sum_{n\geq 1} \frac{c_n^2}{\mathrm{Var}[f]}\, e^{-\lambda_n \tau}. \tag{9}$$

For sufficiently large τ, the slowest mode dominates, so $C(\tau) \sim \exp(-\lambda_1 \tau)$, and we identify $\tau_{\text{corr}} = 1/\lambda_1$ (the relaxation time of the spectral gap). A pure single exponential is exact for an Ornstein–Uhlenbeck process ($\lambda_n = n\lambda_1$ collapses to a single mode for linear observables)

but is approximate for the full nonlinear amplitude dynamics, where sub-leading modes contribute a short transient before the leading mode dominates. In our nonlinear fits, this manifests as a small departure from a pure exponential at very short lags; we exclude the first lag from the fit and find that the residual is consistently below 5 % over the range used (see Sec. III E).

This relaxation time scale depends on all three parameters (ε, α, d) in a manner that is not reducible to the natural parameters (η_1, η_2) alone. Substituting the dimensionless time $\tau = (d/\omega_0^2)t$ into Eq. (2) yields a dimensionless form depending only on the ratios $\varepsilon\omega_0^2/d$ and $\alpha\omega_0^2/d$, so a dimensionless correlation time $\tau'_{\rm corr}(\eta_1, \eta_2)$ corresponds to a physical $\tau_{\rm corr} = (\omega_0^2/d)\,\tau'_{\rm corr}$. Hence the product $\tau_{\rm corr} \cdot d \propto \omega_0^2$ for fixed natural parameters; at fixed ω_0 this reduces to $\tau_{\rm corr} \propto 1/d$. For the linearized (Ornstein–Uhlenbeck) limit where $\alpha \to 0$, this scaling is exact since the relaxation rate is set by d alone when $\varepsilon\omega_0^2/d$ is fixed. For the full nonlinear system, we verify this scaling numerically; the persistence of the scaling reflects that the relaxation dynamics is dominated by fluctuations in the potential well, whose shape is determined by the natural parameters.

To examine the scaling $\tau_{\rm corr} \times d \approx$ constant for fixed ratios $(\varepsilon/d, \alpha/d)$, we simulate the stochastic VdP oscillator across parameter combinations at $\omega_0 = 2\pi$, holding the ratios fixed and varying d:

TABLE I. Testing of the scaling $\tau_{\rm corr} \times d \approx$ constant for fixed natural parameters $(\varepsilon/d, \alpha/d)$ at $\omega_0 = 2\pi$. CV denotes coefficient of variation across the three values of d.

ε/d	α/d	d values	$\langle\tau_{\rm corr} \cdot d\rangle$	CV
-2	-1	$\{0.05, 0.075, 0.10\}$	0.47	2.2 %
-1	-1	$\{0.05, 0.10, 0.15\}$	0.87	5.4 %
$+1$	-1	$\{0.05, 0.10, 0.15\}$	1.06	1.6 %
0	-1	$\{0.05, 0.10, 0.15\}$	3.13	16.4 %

As shown in Table I, the product $\tau_{\rm corr} \times d$ is well conserved away from the bifurcation (CV $\leq$ 5.4 %); near the critical point ($\varepsilon = 0$), the larger CV reflects the divergence of the relaxation time scale, where the exponential ACF fit becomes harder to constrain and the small-ε expansion underlying the scaling has corrections of order $\sqrt{d}/\omega_0^2$. The invariance implies that $\tau_{\rm corr}$, combined with T_1 and T_2, provides the third independent constraint needed

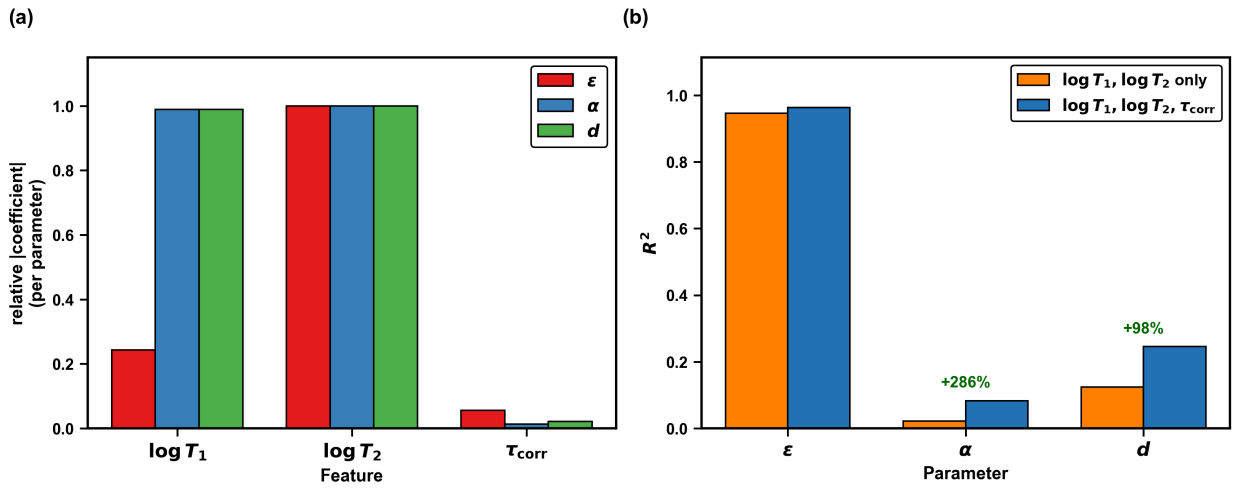

FIG. 1. Contribution of τ_{corr} to parameter identifiability. (a) Linear coefficient magnitudes for each feature. (b) R^2 comparison with and without τ_{corr}, showing substantial improvements for α and d and a small marginal gain for ε.

to identify (ε, α, d) separately.

D. Quantifying τ_{corr}'s contribution to identifiability

The triplet $(T_1, T_2, \tau_{\text{corr}})$ captures the essential information for identifying (ε, α, d): the amplitude moments T_1, T_2 determine the ratios ε/d, α/d through the stationary distribution, while τ_{corr} provides leading-order temporal information that scales these ratios to absolute values. Whereas the full transition density formally contains complete information, these three statistics provide an efficient low-dimensional summary. As we demonstrate empirically, they achieve performance comparable to methods using far richer representations.

To verify that τ_{corr} provides genuine information for breaking the degeneracy, we examine its marginal contribution when all three parameters vary. Using 90 test samples with $\varepsilon \in [-0.29, 0.29]$, $\alpha \in [-0.14, -0.06]$, and $d \in [0.06, 0.15]$, we fit polynomial regressions with and without τ_{corr}. Figure 1 presents this analysis visually.

Including τ_{corr} improves the estimation of α by nearly $4\times$ (R^2 from 0.021 to 0.083) and the estimation of d by approximately $2\times$ (R^2 from 0.124 to 0.246), consistent with τ_{corr} providing information not captured by the amplitude moments. The absolute R^2 values are modest because the test set spans a wide range of all three parameters; the relative gain is

the relevant quantity for assessing τ_{corr}'s informational contribution.

The small marginal gain for ε (R^2 from 0.946 to 0.963) reflects a distinct phenomenon: ε estimation is already well-constrained by T_1 and T_2 because $\varepsilon\omega_0^2/d$ and the moments are strongly related. Our benchmark tests (Sec. V B) use fixed $\alpha = -0.1$ and $d = 0.1$, which removes parameter degeneracy since d is known. In this setting, T_1 and T_2 directly constrain ε, and τ_{corr}'s role in breaking degeneracy is masked. The contribution of τ_{corr} emerges when d is unknown, as demonstrated by the generalization experiments (Sec. V G).

E. Amplitude extraction and the autocorrelation function

We briefly recall the operations used to compute the amplitude statistics. Given a real-valued signal $x(t)$, its *analytic signal* is $z(t) = x(t) + \mathrm{i}\,\mathcal{H}[x](t)$, where the *Hilbert transform*

$$\mathcal{H}[x](t) \;=\; \frac{1}{\pi}\,\mathrm{PV}\!\int_{-\infty}^{\infty} \frac{x(s)}{t-s}\,\mathrm{d}s \tag{10}$$

shifts each Fourier component by $-\pi/2$. The Hilbert transform is an *all-pass* filter: its frequency response has unit magnitude at all frequencies and therefore does not attenuate noise.

The *instantaneous amplitude* (or envelope) is the modulus of the analytic signal,

$$a(t) \;=\; |z(t)| \;=\; \sqrt{x(t)^2 + \mathcal{H}[x](t)^2}\,, \tag{11}$$

where $|\cdot|$ denotes the modulus of a complex number. For narrowband signals such as $x(t) \approx a(t)\cos(\omega_0 t + \phi(t))$ with $a(t)$ varying slowly relative to the carrier, $|z(t)|$ recovers the slowly varying envelope $a(t)$ to leading order; the modulus operation acts as a nonlinear demodulation that separates the envelope from the carrier.

The *normalized autocorrelation function* (ACF) of the amplitude is

$$C(\tau) = \frac{\big\langle (a(t) - \bar{a})\,(a(t+\tau) - \bar{a})\big\rangle}{\mathrm{Var}[a]}, \tag{12}$$

where $\bar{a} = \langle a \rangle$ and the average is over a stationary realization. $C(0) = 1$ by construction, and $C(\tau) \to 0$ as $\tau \to \infty$ for an ergodic stationary process. We define the amplitude correlation time as $\tau_{\text{corr}} = \int_0^\infty C(\tau)\,\mathrm{d}\tau$, which for a single-exponential $C(\tau) = e^{-\tau/\tau_{\text{corr}}}$ recovers the decay constant.

Three items are worthy of note. First, although the Hilbert transform does not filter noise, the envelope operation $|\cdot|$ followed by time-averaging in Eq. (12) is a nonlinear low-pass operation that partially suppresses high-frequency measurement noise. Second, white measurement noise contributes only to $C(0)$ and not to $C(\tau)$ for $\tau > 0$, because uncorrelated noise produces no temporal correlations; this is the origin of τ_{corr}'s noise robustness. Third, no differentiation of $x(t)$ is required, in contrast to velocity-based estimators.

Noise sensitivity analysis. The amplitude correlation time τ_{corr} exhibits different noise sensitivity than the amplitude moments. Under additive measurement noise $x_{\mathrm{obs}}(t) = x_{\mathrm{true}}(t) + \eta(t)$ with $\eta \sim \mathcal{N}(0, \sigma_\eta^2)$, the raw amplitude moments are biased: $\langle a_{\mathrm{obs}}^2 \rangle \approx \langle a_{\mathrm{true}}^2 \rangle + 2\sigma_\eta^2$ and similarly for $\langle a^4 \rangle$ with a larger bias. These biases can be estimated and removed by reading σ_η^2 off the high-frequency tail of the periodogram of x_{obs}, where the signal has negligible power. The autocorrelation function at lag $\tau > 0$ is inherently insensitive to white noise. Numerical experiments confirm that, with this noise correction applied, the residual relative bias in $\langle a^2 \rangle$ stays below 1.5 % and in $\langle a^4 \rangle$ below 0.5 % up to $\sigma_\eta/\sigma_x = 0.3$, while τ_{corr} exhibits a measurable downward bias (-43 % at 30 % noise) reflecting the suppression of the empirical ACF by the noise-induced amplitude fluctuations.

IV. POLYNOMIAL REGRESSION METHODOLOGY

A. Feature extraction

From a trajectory $\{x(t_k)\}_{k=0}^N$, we extract the three sufficient statistics. The amplitude is extracted via the Hilbert transform: $a(t) = |x(t) + \mathrm{i}\mathcal{H}[x](t)|$. The empirical second and fourth moments are computed as $T_1 = (1/N)\sum_k a(t_k)^2$ and $T_2 = (1/N)\sum_k a(t_k)^4$. The correlation time τ_{corr} is extracted by fitting an exponential decay model $C(\tau) = A\exp(-\tau/\tau_{\mathrm{corr}})$ to the amplitude autocorrelation function. The fit is performed over lags $\tau \in [\mathrm{d}t, 500\,\mathrm{d}t]$, explicitly excluding the lag-zero sample: including $C(0) = 1$ would force the exponential to bridge the noise-induced spike at $\tau = 0$ (see Sec. III E), producing a spuriously short τ_{corr} under measurement noise. With this lag-zero exclusion, τ_{corr} inherits the noise robustness of $C(\tau > 0)$. As discussed in Sec. III E, a pure single exponential is the leading-eigenmode approximation to the spectral decomposition of the FP autocorrelation and is exact only in the OU (linearized) limit; for the full nonlinear system it is a working approximation,

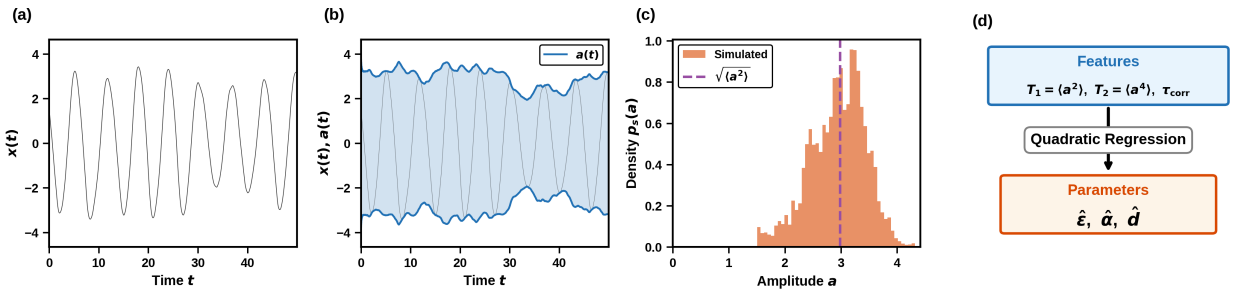

FIG. 2. Method schematic. (a) Observed noisy trajectory $x(t)$ over $T = 50$ time units. (b) Amplitude envelope $a(t)$ extracted via Hilbert transform, showing $a(t)$ as the envelope of the underlying oscillation. (c) Stationary amplitude distribution $p_{\mathrm{s}}(a)$ computed from $T = 500$ time units, with $\sqrt{\langle a^2 \rangle}$ indicated by the dashed line. (d) Quadratic polynomial regression mapping features to parameters.

with root-mean-square residual below 5 % of $C(0)$ over $\tau \in [0.5, 5]$. The fit is bounded by $\tau_{\mathrm{corr}} \in [0.1, 200]$, and the autocorrelation is computed via FFT. If the nonlinear fit fails to converge, we fall back to the lag at which the ACF first crosses $1/e$, providing a reasonable estimate even for atypical decay profiles. Figure 2 illustrates the complete inference pipeline.

B. Quadratic polynomial model

We approximate the inverse mapping from features to parameters using quadratic polynomials. For each physical parameter $\theta \in \{\varepsilon, \alpha, d\}$, we fit

$$\hat{\theta} = \beta_0 + \sum_{i=1}^{3} \beta_i \tilde{f}_i + \sum_{i=1}^{3} \beta_{ii} \tilde{f}_i^2 + \sum_{i<j} \beta_{ij} \tilde{f}_i \tilde{f}_j, \tag{13}$$

where $\tilde{f}_i$ are standardized features computed from the training data. The amplitude moments span a wide dynamic range across the training set, so we use $\tilde{f}_1 = (\log T_1 - \mu_{\log T_1})/\sigma_{\log T_1}$ and $\tilde{f}_2 = (\log T_2 - \mu_{\log T_2})/\sigma_{\log T_2}$; the correlation time has a narrower range and is standardized linearly, $\tilde{f}_3 = (\tau_{\mathrm{corr}} - \mu_\tau)/\sigma_\tau$. Standardization constants are listed in Appendix A.

The quadratic model has 10 terms per parameter (1 intercept + 3 linear + 3 squared + 3 cross terms), giving 30 total coefficients for all three parameters.

C. Training procedure

The coefficients are determined from the simulated training data. We sample $N = 3000$ parameter combinations uniformly from $\varepsilon \in [-0.35, 0.35]$, $\alpha \in [-0.15, -0.05]$, $d \in [0.05, 0.15]$. These ranges are not intrinsic to the Van der Pol model but are chosen to bracket the values typical of laboratory thermoacoustic data after non-dimensionalization by ω_0, and in particular to include the benchmark test cases of Son and Lee [24] ($\varepsilon = \pm 0.1, \pm 0.2, \pm 0.3$, $\alpha = -0.1$, $d = 0.1$) used in our comparisons. The constraint $\alpha < 0$ ensures supercritical Hopf behavior; the d range corresponds to turbulence-induced fluctuation levels reported in experimental combustor studies [4, 20]. The training set size is held constant across all trajectory lengths to ensure fair comparison. For each parameter set, we simulate Eq. (1) by integrating with fourth-order Runge–Kutta for the deterministic part combined with an Euler–Maruyama increment for the additive noise, a strong order-1.0 scheme appropriate for additive-noise SDEs and matching the integration method used by Son and Lee [24]. We use a time step $\mathrm{d}t = 0.01$, giving 100 integration steps per oscillation period $2\pi/\omega_0 = 1$, with a burn-in of 10 000 steps to reach the stationary distribution before recording. We confirmed numerical accuracy by comparing the empirical stationary moments to the analytical predictions from Eq. (4); agreement is within 2% in the supercritical regime. We then compute $(\langle a^2 \rangle, \langle a^4 \rangle, \tau_{\mathrm{corr}})$ from each trajectory and discard samples with $\tau_{\mathrm{corr}} \notin [0.1, 150]$ to exclude regimes where stochastic averaging breaks down. Finally, we solve the normal equations for each parameter:

$$\boldsymbol{\beta}_i = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\theta}_i, \tag{14}$$

where $\mathbf{X}$ is the design matrix of quadratic features.

D. Two-stage inversion via the moment relations

As a principled alternative to the polynomial regression, we implemented a two-stage inversion that separates the geometric part of the inverse problem from the model-dependent part.

Stage 1: moment inversion (model-agnostic). The stationary amplitude PDF belongs to an exponential family of the form

$$p_{\mathrm{s}}(a) \;\propto\; a\, \exp(\eta_1 a^2 + \eta_2 a^4), \qquad \eta_2 < 0, \tag{15}$$

whose natural parameters (η_1, η_2) are in one-to-one correspondence with the moments (T_1, T_2) via the implicit relations

$$T_k(\eta_1, \eta_2) = \frac{\int_0^\infty a^{2k+1}\, e^{\eta_1 a^2 + \eta_2 a^4}\, \mathrm{d}a}{\int_0^\infty a\, e^{\eta_1 a^2 + \eta_2 a^4}\, \mathrm{d}a}, \quad k = 1, 2. \tag{16}$$

Given the empirical moments (T_1, T_2), we solve Eq. (16) for (η_1, η_2) by Levenberg–Marquardt iteration with the polynomial prediction as warm start. Convergence is achieved in 5–10 iterations on all test cases considered. The natural parameters (η_1, η_2) are minimal sufficient statistics for the stationary distribution within the exponential family (15).

Stage 2: mapping to physical parameters. The translation $(\eta_1, \eta_2, \tau_{\mathrm{corr}}) \to (\varepsilon, \alpha, d)$ is learned from the same training set used by the polynomial regression. Motivated by the $\tau_{\mathrm{corr}} \times d$ scaling identified in Sec. III, we fit RBF interpolators for $\varepsilon\tau_{\mathrm{corr}}$, $\alpha\tau_{\mathrm{corr}}$, and $d\tau_{\mathrm{corr}}$ as functions of (η_1, η_2), and divide by τ_{corr} at prediction time.

Performance and comparison. On the Son and Lee [24] benchmark cases, the polynomial regression substantially outperforms the two-stage method by factors of up to roughly $10\times$ across all parameters at both $T = 100$ and $T = 500$ (Table II).

TABLE II. Polynomial regression vs. two-stage inversion on the Son and Lee benchmark (main regime $|\varepsilon| \geq 0.2$, mean over 5 seeds per case).

	$T = 100$			$T = 500$		
Method	ε	α	d	ε	α	d
Polynomial	13.8%	4.3%	5.8%	9.8%	3.9%	3.9%
Two-stage	37.4%	41.4%	34.5%	15.4%	16.6%	16.0%

The two-stage method is more sensitive to finite-sample noise in (T_1, T_2) because the moment-inversion step is geometrically ill-conditioned at large $|\eta_i|$, where the stationary PDF concentrates and the Jacobian $\partial(T_1, T_2)/\partial(\eta_1, \eta_2)$ has condition number exceeding 10^5 in the supercritical regime at $\omega_0 = 2\pi$. The polynomial regression averages over this conditioning implicitly. The two-stage method's principal advantage is conceptual rather than numerical: it separates the well-posed half of the inverse problem (the bijective map $(T_1, T_2) \leftrightarrow (\eta_1, \eta_2)$, exact within the exponential family) from the model-dependent half (the relationship between natural and physical parameters, which depends on averaging approximations and noise-structure assumptions). The polynomial regression collapses both halves into a single

fit. We therefore retain the polynomial regression as the primary method and present the two-stage inversion as a principled baseline.

E. Comparison methods

We compare against five alternative approaches. The first is k-Nearest Neighbours (k-NN), which stores the training set and predicts by averaging the $k = 5$ nearest neighbours in standardized feature space, representing a nonparametric baseline with approximately 18 000 effective parameters. The second is a Statistics-Based Neural Network (SBNN), which is a multilayer perceptron with hidden layers of 128, 128, 64 units, ReLU activations, and dropout regularization. It has approximately 25 000 trainable parameters. The third is the Volterra series approach following Belousov *et al.* [11], which estimates $|\varepsilon|$ from spectral linewidth and d from velocity increment variance but cannot estimate α independently. The fourth is the Son and Lee [24] PINN, using published results employing physics-informed neural networks solving the Fokker–Planck equation. The fifth is sparse identification of nonlinear dynamics (SINDy) [36], applied to the state vector $X = (x, \dot{x})$ with a polynomial library up to cubic order, finite-difference estimation of $\dot{x}$ and $\ddot{x}$, and sequential thresholded least squares with $\lambda = 0.005$ (chosen by Pareto analysis on a held-out validation case). As detailed in Sec. V D, SINDy correctly identifies the structural form of the deterministic dynamics (linear restoring force and cubic nonlinearity) but its parameter estimates are biased for the noise-driven regime of interest here: differentiation amplifies the stochastic forcing into the regression target; the small-magnitude growth-rate coefficient is routinely thresholded to zero by the sparsification step; and the noise intensity is not directly identified. Stochastic extensions of SINDy [37–39] that abandon pointwise OLS in favour of likelihood-based or weak-form identification may address these limitations.

V. RESULTS

A. Evaluation metrics and test structure

We evaluate using relative error for parameters bounded away from zero:

$$E_\theta = \frac{|\hat{\theta} - \theta|}{|\theta|} \times 100\%. \tag{17}$$

For the critical regime where $|\varepsilon| \leq 0.05$, relative error becomes ill-defined or misleading, so we report mean absolute error (MAE). Results are reported as mean $\pm$ standard deviation across 5 random seeds per test case.

Following Son and Lee [24], we consider benchmark test cases with $\varepsilon \in \{-0.3, -0.2, +0.2, +0.3\}$, $\alpha = -0.1$, and $d = 0.1$. Cases with $|\varepsilon| < 0.1$ are analysed separately in Sec. V F because their proximity to the bifurcation point leads to qualitatively different error characteristics that would otherwise dominate the variance in summary statistics.

B. Benchmark cases at $T = 100$

We first report the four benchmark test cases (excluding $|\varepsilon| < 0.1$, which falls in the critical regime; see Sec. V F) using trajectory length $T = 100$ time units (10 000 samples at $\mathrm{d}t = 0.01$). Table III presents per-case results for our quadratic regression method together with their previously reported PINN performance.

TABLE III. Per-case comparison at $T = 100$ ($|\varepsilon| \geq 0.2$). PINN results from Son and Lee [24].

Case	ε	Method	ε error	α error	d error
1	-0.30	Quadratic	21.7%	2.7%	13.1%
		PINN	5%	58%	6%
2	-0.20	Quadratic	10.0%	1.6%	8.3%
		PINN	6%	46%	7%
3	$+0.20$	Quadratic	8.9%	6.1%	1.1%
		PINN	5%	48%	5%
4	$+0.30$	Quadratic	10.9%	6.3%	1.2%
		PINN	7%	52%	6%
Mean		Quadratic	12.9%	4.2%	5.9%
		PINN	5.7%	51.2%	6.0%

The comparison reveals complementary performance. PINNs achieve 2.3 times better estimation of ε (5.7% vs 12.9%) but our method achieves 12 times better estimation of α (4.2% vs 51.2%). The estimation of d is comparable (5.9% vs 6.0%). The per-case d errors show pronounced asymmetry between subcritical cases (-0.30, -0.20) and supercritical

cases (+0.20, +0.30): in the supercritical regime the limit cycle gives a strong signal in T_1, which combined with τ_{corr} tightly constrains d; in the subcritical regime the trajectory fluctuates near the origin and d is more weakly identified.

The PINN advantage for ε reflects the physics-informed loss function: the Fokker–Planck residual provides strong local constraints on ε through the drift term. Our advantage for α reflects the information content of τ_{corr}: the relaxation time scale is sensitive to α through its effect on the shape of the potential well, information that is not captured by the stationary PDF alone.

We note that the relatively higher ε error in our method may also reflect the slow convergence of higher-order moment estimates: $T_2 = \langle a^4 \rangle$ has substantially larger sampling variance than T_1, and since ε estimation relies on a near-cancelling combination of moment-dependent terms (see Appendix A), this sampling noise propagates directly to the ε estimate. PINNs, by contrast, extract ε information from the full PDF shape near $a = 0$, which may be less variance-sensitive.

C. Effect of trajectory length

We examine how estimation accuracy varies with observation duration T. As discussed in Sec. II, T is measured in units of ω_0^{-1}; $T = 100$ corresponds to approximately 16 oscillation periods or 8–30 correlation times. To isolate the effect of trajectory length, we retrain a separate polynomial model for each T value (same $N = 3000$ training samples, each simulated at that T). Each (T, case) combination is evaluated on 30 independent test seeds.

Figure 3 shows the mean relative error per parameter as a function of T, with error bars indicating the standard error of the mean over the 30 seeds. The behavior is compared with a $1/\sqrt{T}$ sampling-variance scaling, which is consistent with the central-limit-theorem expectation for moment-based estimators on stationary data. The dashed reference line shows a $1/\sqrt{T}$ scaling. The estimation of ε tracks this scaling somewhat, limited by a residual bias floor of $\sim$5–8 % from the polynomial's finite-order approximation. In contrast, the α and d errors flatten at a sampling-variance floor of $\sim$4–5 % across the tested range.

For practical thermoacoustic monitoring, the implication is that longer observation windows reduce ε error in the expected $1/\sqrt{T}$ fashion; $T = 100$–200 provides a reasonable compromise between accuracy and latency in real-time settings.

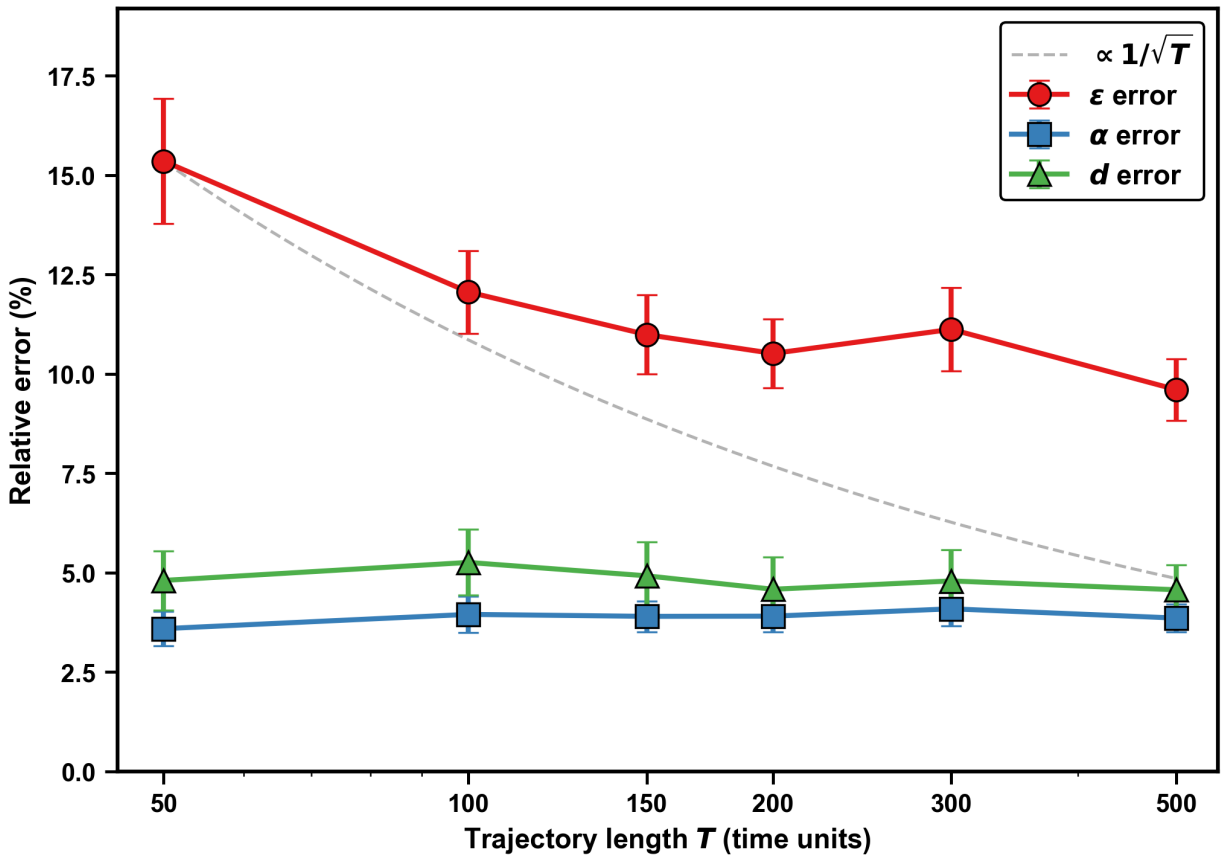

FIG. 3. Parameter estimation errors versus trajectory length T (main regime $|\varepsilon| \geq 0.2$, 30 seeds per case, per-T training). Error bars show the standard error of the mean; dashed line shows the $1/\sqrt{T}$ reference anchored at $T = 50$.

D. Method comparison

Figure 4 presents comprehensive method comparison at two trajectory lengths. Panel (a) compares our quadratic regression against SBNN and the Son and Lee [24] PINN at $T = 100$, showing parameter-specific errors. Panel (b) shows per-case ε errors at $T = 500$ for all methods including Volterra.

At $T = 500$ on the four main-regime cases, the methods rank as follows (mean errors across cases with 5 seeds per case): Quadratic 9.4 %/4.3 %/4.2 % ($\varepsilon/\alpha/d$); k-NN 11.3 %/9.9 %/13.5 %; SBNN 11.5 %/6.2 %/1.0 %. The quadratic regression with 30 parameters matches or outperforms the SBNN with 25 000 parameters on ε and α, while the SBNN retains the lowest d error — its expressive headroom gives it an advantage on the parameter

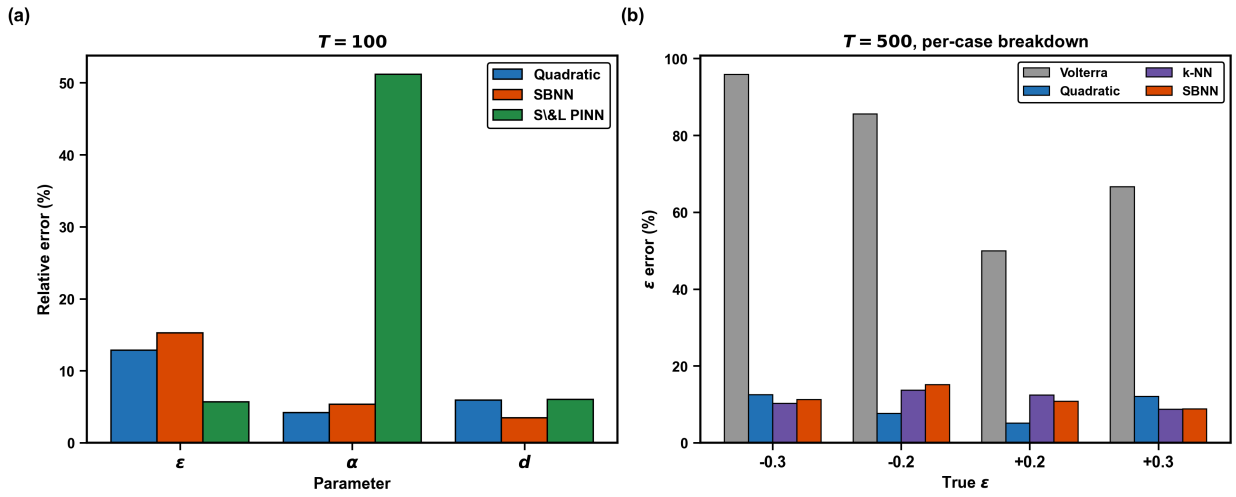

FIG. 4. Method comparison. (a) Parameter errors at $T = 100$ for Quadratic, SBNN, and PINN methods. (b) Per-case ε errors at $T = 500$ showing Volterra failure.

most sensitive to fine PDF shape. The k-NN approach shows consistently higher errors, likely due to the curse of dimensionality in the three-dimensional feature space combined with the complex, nonlinear inverse mapping.

The Volterra approach exhibits systematic failure, with mean ε error in the main regime of 74.6 % compared to 9.4 % for quadratic regression. Examination of per-case predictions reveals that Volterra consistently underestimates $|\varepsilon|$, predicting values around 0.02–0.10 regardless of the true value. The spectral width method that Volterra employs fails because it assumes a linear response regime where perturbative expansion converges. Near the bifurcation, this assumption breaks down.

SINDy. The SINDy baseline provides instructive contrast. With a polynomial library through cubic order in $(x, \dot{x})$ and sequential thresholded least squares ($\lambda = 0.005$), SINDy correctly identifies the linear restoring force ($\hat{c}_x \approx -\omega_0^2$) and selects the cubic nonlinearity term $x^2\dot{x}$ from the library, recovering the structural form of Eq. (1). However, on the Son and Lee benchmark cases at $T = 500$, parameter estimates are systematically biased. The growth-rate coefficient ε is thresholded to zero in four of the six benchmark cases (relative error 100 %) and is partially recovered only for the most supercritical case $\varepsilon = +0.3$ (relative error 20.7 %); the asymmetry reflects that the $\varepsilon\dot{x}$ term is a strong, persistent signal only when the system exhibits a sustained limit cycle. The cubic coefficient α is estimated within 11 % for $\varepsilon = +0.3$ but with errors of 50–2000 % for the smaller-amplitude and subcritical

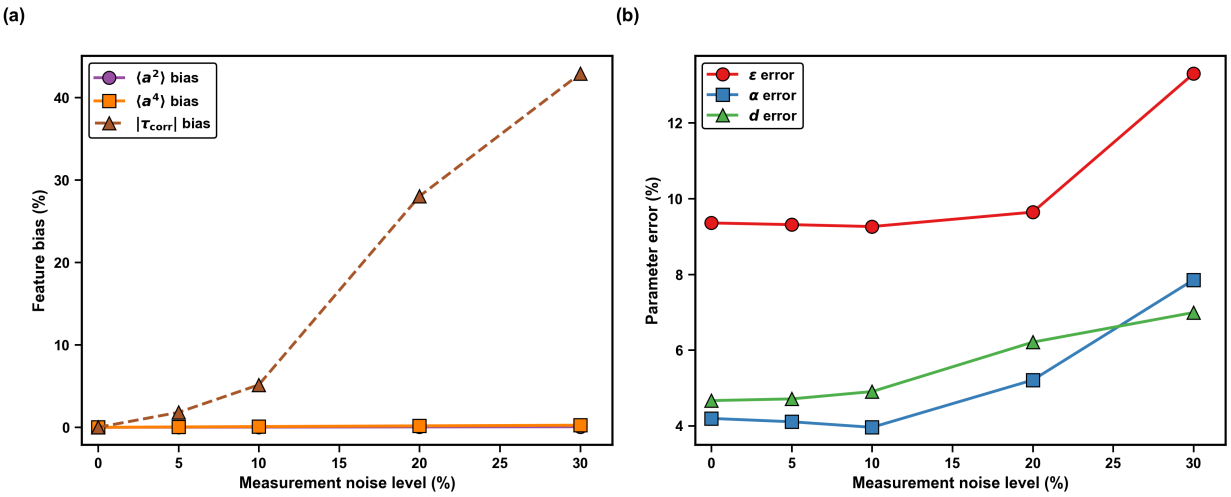

FIG. 5. Noise sensitivity analysis at $T = 500$. (a) Residual feature biases after noise correction: T_1 and T_2 stay essentially unbiased ($\leq 1.5\,\%$ and $\leq 0.5\,\%$), while τ_{corr} exhibits a measurable downward bias of $-43\,\%$ at $30\,\%$ noise. (b) Parameter estimation errors: ε, α, d remain near their clean-data values up to $\sim 20\,\%$ noise.

cases (main-regime mean $752\,\%$). The noise intensity d, estimated from the regression-residual variance, is biased low by 48–$63\,\%$ across all cases (main-regime mean $56.6\,\%$). These shortcomings are not specific to multiplicative noise: even with additive forcing as in Eq. (1), the combination of finite-difference differentiation and pointwise OLS amplifies the stochastic component and inflates the regression target in a way that ordinary sparse regression cannot disentangle from the deterministic terms. Stochastic extensions [37–39] that work in the integral or likelihood domain are required for principled identification in this regime. The comparison highlights the value of working in the stationary-statistics domain, where one effectively integrates over the trajectory rather than differentiating it.

E. Noise sensitivity

We assess sensitivity to measurement noise by adding Gaussian noise with standard deviation $\sigma_\eta = r \times \mathrm{std}(x)$ to the position trajectory. We apply the periodogram-based noise correction described in Sec. III E to the empirical moments before passing them to the regression. Figure 5 presents the results.

Three observations are notable (Table IV). First, the noise correction essentially elimi-

TABLE IV. Noise sensitivity at $T = 500$, mean over the four main-regime cases ($\varepsilon \in \{\pm 0.2, \pm 0.3\}$). Moment biases are residual values after the periodogram-based noise correction; τ_{corr} bias is reported as a percentage of the clean-data $\tau_{\text{corr}} \approx 3.9$.

σ_η/σ_x	T_1 bias	T_2 bias	τ_{corr} bias	ε err	α err	d err
0%	0.0%	0.0%	0.0%	9.4%	4.2%	4.7%
5%	0.3%	0.1%	−1.8%	9.3%	4.1%	4.7%
10%	0.5%	0.2%	−5.1%	9.3%	4.0%	4.9%
20%	1.0%	0.3%	−28.0%	9.6%	5.2%	6.2%
30%	1.4%	0.5%	−42.9%	13.3%	7.9%	7.0%

nates moment-bias: residual T_1 and T_2 biases stay below 1.5 % and 0.5 % respectively at all tested noise levels, confirming the leading-order analysis of Sec. III E. Second, all three parameter errors remain within 1–2 % (absolute) of their clean-data values up to $\sigma_\eta/\sigma_x = 0.2$: ε stays within 0.2 %, α within 1 %, d within 1.5 %. The qualitative picture is graceful degradation of all three parameters under realistic noise levels. Third, the residual high-noise error is dominated by the bias in τ_{corr}: at 30 % noise the empirical autocorrelation is suppressed by the noise-induced amplitude fluctuations and the exponential fit underestimates τ_{corr} by −43 %. The ε error is most affected by this because its regression coefficients have the largest sensitivity to τ_{corr} via cross-terms; α and d stay below 8 %.

For applications, two implications follow: (i) the periodogram-based moment correction is essential for retaining moment-based accuracy under noise and should be applied whenever the noise floor is non-negligible; (ii) a robust τ_{corr} estimator (e.g., based on the integral $\int_0^{\tau^*} C(\tau)\,\mathrm{d}\tau$ over a truncated lag range, or a more flexible parametric form such as a stretched exponential) would improve the high-noise regime further; we leave this for future work.

F. Critical regime ($|\varepsilon| < 0.1$)

This regime presents a unique challenge in that the dynamics is dominated by noise rather than deterministic structure. Figure 6 presents the critical regime performance.

The method achieves MAE below 0.030 across the critical regime, with the cross-regime mean at 0.025 and the per-case maximum at 0.030 (at $\varepsilon = +0.03$). The bias is small and

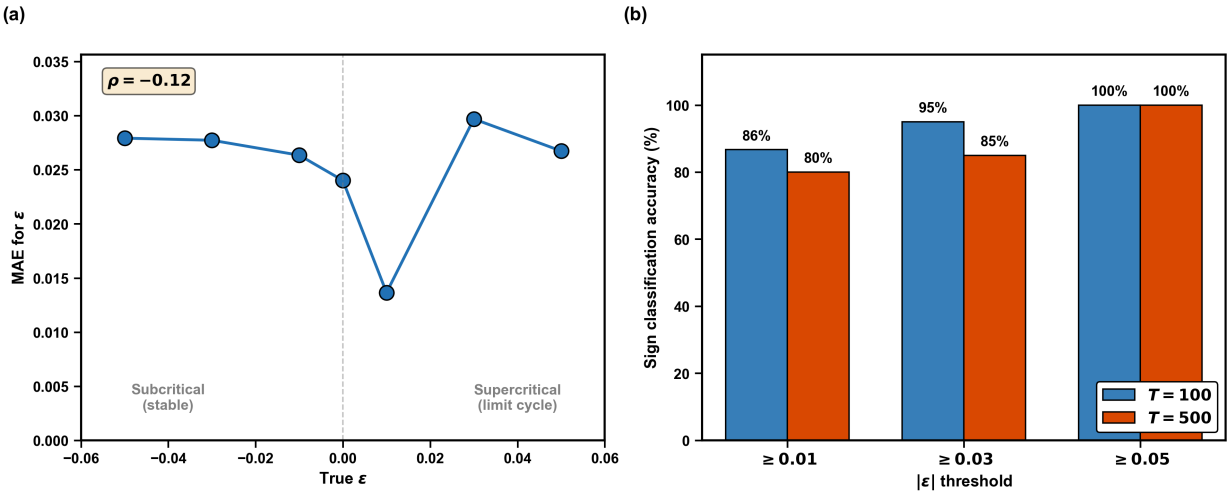

FIG. 6. Critical regime performance at $T = 500$. (a) MAE versus true ε across the critical regime $|\varepsilon| \leq 0.05$. (b) Sign classification accuracy across thresholds.

well-behaved.

Correctly determining the sign of ε is crucial for applications: it distinguishes subcritical (stable fixed point, $\varepsilon < 0$) from supercritical (limit cycle, $\varepsilon > 0$) regimes.

At $T = 500$, the method achieves 100 % sign classification accuracy for $|\varepsilon| \geq 0.05$ and 85 % for $|\varepsilon| \geq 0.03$; sign determination is therefore reliable when the operating point is at least $|\varepsilon| = 0.05$ away from the bifurcation. For typical thermoacoustic systems this corresponds to growth rates resolvable within standard observation windows. Closer to the bifurcation point, the linear discriminant $\text{sign}(\hat{\varepsilon})$ is dominated by the residual estimator bias rather than by the true sign of ε; a Bayesian decision rule incorporating the empirical estimator variance would be more appropriate in that regime.

G. Generalization across parameter space

The baseline benchmark cases presented above use fixed $\alpha = -0.1$ and $d = 0.1$. We next evaluate generalization when these parameters also vary.

Performance degrades at parameter boundaries, particularly for α approaching zero (75% error at $\alpha = -0.05$) where the nonlinear saturation becomes weak and difficult to detect, and for d approaching the edges of the training range (61% error at $d = 0.05$, 24% at $d = 0.15$) where the polynomial extrapolates beyond well-covered training data. The random test

TABLE V. Generalisation performance at $T = 500$.

Test Condition	ε Error	α Error	d Error
Standard ($\alpha = -0.10$, $d = 0.10$)	9.6%	3.4%	4.1%
$\alpha = -0.15$, $d = 0.10$	10.0%	26.0%	5.0%
$\alpha = -0.05$, $d = 0.10$	25.1%	75.2%	4.2%
$\alpha = -0.10$, $d = 0.05$	20.7%	4.4%	61.3%
$\alpha = -0.10$, $d = 0.15$	7.9%	2.4%	23.6%
Random (30 cases)	24.3%	28.3%	17.7%

cases, spanning the full parameter ranges, show moderate degradation (24% for ε, 28% for α, 18% for d).

These generalization results also demonstrate τ_{corr}'s importance: in the fixed-d benchmark (Table III), the degeneracy is artificially removed. When d varies (Table V), τ_{corr}'s role in breaking the degeneracy becomes essential, as shown by the R^2 analysis in Sec. III.

VI. DISCUSSION

A. Complementary strengths of PINN and sufficient statistics approaches

The comparison between our quadratic regression and the previously reported PINN reveals fundamentally different inductive biases, each suited to particular aspects of the problem. The PINN excels at ε estimation because the Fokker–Planck residual loss directly constrains the drift coefficient $\mu(a) = (\varepsilon/2)a + (\alpha/8)a^3 + d/(2\omega_0^2 a)$. The linear term $(\varepsilon/2)a$ dominates for small amplitudes, and the PDF shape near $a = 0$ is highly sensitive to ε. This physics-based regularization effectively extracts information from the full PDF shape, not just moments.

Our approach excels at the estimation of α because τ_{corr} encodes the potential well curvature, which depends strongly on α through the a^3 term in the drift. The stationary PDF is relatively insensitive to α when $|\alpha|$ is small, but the relaxation dynamics, that is, how quickly the system returns to equilibrium after perturbation, depends on the potential landscape.

For practical applications, the choice between methods should be guided in part by which

parameter is most critical. Thermoacoustic stability margin assessment prioritizes ε, favouring PINNs. Limit-cycle amplitude prediction prioritizes α, favouring our approach. Control system design may require accurate estimates of all parameters, suggesting an ensemble approach.

B. Why quadratic regression suffices

The surprising effectiveness of quadratic regression with only 30 coefficients, compared to 25 000 for the SBNN, requires further explanation. Several factors contribute. First, the sufficient statistics $(T_1, T_2, \tau_{\mathrm{corr}})$ are theoretically optimal features; adding more features does not improve asymptotic performance. Second, the inverse mapping from features to parameters is smooth and approximately quadratic in the standardised coordinates used for the regression. Third, the three-dimensional feature space avoids the curse of dimensionality that plagues high-dimensional approaches.

The limiting factors are not model capacity but rather fundamental identifiability limits, moment variance at short trajectories, and the fact that three sufficient statistics contain limited information about the full transition density. The two-stage moment-inversion baseline (Sec. IV D) is informative here: it makes the geometric content of the inverse problem explicit, and its substantially worse performance compared with the polynomial regression highlights that the polynomial implicitly averages over the ill-conditioning of the moment-inversion step.

This has practical implications: no neural network training is required since coefficients are determined by least squares; the formulas are interpretable and enable analysis and debugging; inference is a single matrix-vector multiply taking microseconds; and the results are reproducible without stochastic optimization or hyperparameter tuning.

C. Regime classification in the critical region

The sign classification results (100 % accuracy for $|\varepsilon| \geq 0.05$ at $T = 500$, 85 % at $|\varepsilon| \geq 0.03$; see Fig. 6b) have practical significance that may exceed the importance of magnitude estimation for some applications. Knowing whether a thermoacoustic system is subcritical (damped, $\varepsilon < 0$) or supercritical (self-excited, $\varepsilon > 0$) determines whether it will

exhibit sustained oscillations or decay to quiescence. This binary classification is often the practically relevant question.

The practical threshold $|\hat{\varepsilon}| \geq 0.05$ for fully reliable regime determination provides actionable guidance. If the estimated ε falls in the range $[-0.05, +0.05]$, the system is near the bifurcation point and regime classification is uncertain at the $\sim 15\,\%$ misclassification rate; additional data or caution is warranted. Outside this range, the predicted regime is reliable for operational decisions.

D. Limitations

Several limitations of our sufficiency-statistics approach should be noted. Regarding the ε estimation gap: at $T = 100$, our ε error (12.9%) is about 2.3× worse than PINNs (5.7%). The Fokker–Planck regularization in PINNs provides stronger constraints on ε from the PDF shape near $a = 0$. The white-noise assumption means that colored noise with significant correlation time requires modified treatment, as shown by Bonciolini *et al.* [20]. The noise sensitivity of estimated parameters at very high noise levels (above $\sigma_\eta/\sigma_x = 0.2$) is dominated by bias in τ_{corr}, which arises because the empirical autocorrelation is suppressed by noise-induced amplitude fluctuations; a more robust τ_{corr} estimator would mitigate this. And the fundamental identifiability limit for α from moments alone means the estimation of α degrades most severely when the full parameter space is considered.

The method assumes the system has reached statistical stationarity, requiring observation times T substantially longer than τ_{corr}. In systems with drifting parameters (e.g., changing thermal boundary conditions), if parameters evolve on timescales comparable to or faster than τ_{corr}, the moment and correlation estimates become biased. This quasi-steady assumption is inherent to any statistics-based identification method and contrasts with instantaneous filtering approaches (e.g., Kalman filters) that can track time-varying parameters. For tracking time-varying parameters, the data assimilation approaches of Nóvoa and Magri [28] provide complementary capabilities through sequential Bayesian updating, though at higher computational cost and other complications.

Near subcritical Hopf bifurcations, the bistability documented by Subramanian *et al.* [40] creates additional challenges: the system may occupy different attractors depending on initial conditions and perturbation history, leading to multimodal amplitude distributions that

violate the single-peaked exponential family assumption underlying our sufficient statistics. Similarly, the intermittency phenomena reported by Nair *et al.* [41] near the onset of instability in turbulent combustors produce non-stationary dynamics where the amplitude statistics mix contributions from qualitatively different dynamical regimes. In such transitional states, our method would report parameters corresponding to some effective average behavior rather than the instantaneous system state.

The polynomial coefficients provided in Appendix A are valid only within the training parameter ranges: $\varepsilon \in [-0.35, +0.35]$, $\alpha \in [-0.15, -0.05]$, $d \in [0.05, 0.15]$. Unlike PINNs or MLE methods that may degrade gracefully outside their training domain, polynomial regression exhibits Runge-type divergence when extrapolated. For $|\varepsilon| > 0.35$ (strong instability), the quadratic approximation breaks down as the system enters strongly nonlinear regimes where stochastic averaging itself becomes questionable. For $d < 0.05$ (weak noise), the system approaches the deterministic limit where the $\tau_{\text{corr}} \times d$ scaling no longer holds. For $d > 0.15$ (strong noise), noise-dominated dynamics may exhibit non-exponential autocorrelation decay. In practice, users should verify that extracted features fall within training bounds before trusting predictions. If features lie outside these ranges, the polynomial should not be applied; instead, retraining on an expanded parameter set or switching to a more flexible method is recommended.

VII. CONCLUSIONS

We have developed a sufficient statistics approach for parameter inference in the stochastic Van der Pol oscillator that complements physics-informed neural networks. The key contributions are as follows.

On the theoretical side, we identified τ_{corr} as the third sufficient statistic that breaks the parameter degeneracy noted by Boujo and Noiray [17]. Through the Fokker–Planck relaxation dynamics, τ_{corr} provides information about d that is independent of the stationary distribution.

On the methodological side, we demonstrated that quadratic polynomial regression with 30 coefficients provides good performance, matching neural networks with 25 000 parameters. This 850-fold reduction in model complexity comes with comparable accuracy and substantial gains in interpretability and computational efficiency (Fig. 4).

In practical comparison at equal data ($T = 100$), our method achieves 12 times better α estimation than the previously reported PINNs (4.2% vs 51.2%) while the previously reported PINNs achieve 2.3 times better ε estimation (5.7% vs 12.9%). The methods exploit fundamentally different information: PINNs leverage the local structure of the Fokker–Planck equation, while our method leverages the global statistics including temporal correlations. The choice between methods should be guided in part by which parameter is most critical for the application (Table III).

In the critical regime near bifurcation ($|\varepsilon| \leq 0.05$), our method achieves MAE below 0.03 for ε and enables reliable regime classification: sign determination (subcritical vs supercritical) achieves 100 % accuracy for $|\varepsilon| \geq 0.05$ at $T = 500$ and 85 % at $|\varepsilon| \geq 0.03$ (Fig. 6).

With the periodogram-based moment correction (Sec. III E), all three parameter estimates remain within 1–2 % of their clean-data values up to $\sigma_\eta/\sigma_x = 0.2$ measurement noise; the residual high-noise degradation is dominated by bias in τ_{corr} rather than in the moments (Fig. 5). The graceful degradation across all three parameters under realistic noise levels makes the methodology suitable for both laboratory and field applications.

The sufficient-statistics approach provides compact and interpretable expressions suitable for thermoacoustic stability monitoring, particularly when the nonlinear saturation coefficient α, which determines limit cycle amplitude, is the quantity of primary interest. Future work will extend the methodology to coloured noise forcing, multi-mode systems, and noise-robust estimation strategies that exploit the differential parameter sensitivity identified here. Integration with data assimilation frameworks [28] could enable hybrid approaches that combine the interpretability of sufficient statistics with the time-tracking capabilities of sequential Bayesian methods. Extension to azimuthal modes in annular combustors, building on the Langevin regression framework of Indlekofer *et al.* [23], would address industrially relevant configurations. Connections to adjoint sensitivity analysis [42] may enable gradient-based optimization of combustor parameters, while network-based approaches to thermoacoustic suppression [43] could leverage our identified parameters for targeted control strategies. The forced synchronization phenomena documented by Kashinath *et al.* [44] suggest opportunities for open-loop control informed by the sufficient statistics framework.

ACKNOWLEDGMENTS

This research received no grant from any funding agency, commercial or not-for-profit sectors. The author reports no conflict of interest. The author used Claude AI to debug and refactor the code used in the study.

DATA AVAILABILITY STATEMENTS

All data presented are obtained from numerical simulation. The code that supports the findings of this study is available from the corresponding author upon reasonable request.

Appendix A: Complete polynomial formulas

Note: These coefficients are valid only within the training parameter range: $\varepsilon \in [-0.35, +0.35]$, $\alpha \in [-0.15, -0.05]$, $d \in [0.05, 0.15]$, chosen to bracket typical laboratory thermoacoustic values (see Sec. IV). Extrapolation outside this range may produce unreliable results. Users should verify that the extracted features $(T_1, T_2, \tau_{\text{corr}})$ lie within the ranges observed in our training. If $\tau_{\text{corr}} < 0.1$ or $\tau_{\text{corr}} > 150$, the polynomial predictions are unreliable and retraining on an expanded parameter set is recommended.

Feature standardisation (computed from training set with $N = 3000$, $T = 500$, $\omega_0 = 2\pi$):

$$\widetilde{\log T_1} = \frac{\log T_1 - (-0.800)}{2.686}, \tag{A1}$$

$$\widetilde{\log T_2} = \frac{\log T_2 - (-1.266)}{5.065}, \tag{A2}$$

$$\widetilde{\tau} = \frac{\tau_{\text{corr}} - 9.018}{10.366}. \tag{A3}$$

For each physical parameter $\theta \in \{\varepsilon, \alpha, d\}$, $\hat{\theta}$ is a quadratic polynomial in $(\widetilde{\log T_1}, \widetilde{\log T_2}, \widetilde{\tau})$ with coefficients listed in Table VI.

Coefficient interpretation. For $\hat{\alpha}$, the near-equal-magnitude opposite-sign coefficients on $\widetilde{\log T_1}$ (-0.87) and $\widetilde{\log T_2}$ $(+0.88)$ reflect the relation $\alpha \propto T_2/T_1^2$ implied by the limit-cycle scaling. For $\hat{d}$ the same pattern is amplified $(-1.45, +1.46)$ and supplemented by second-order terms with cancellation, accommodating the additional information from τ_{corr}. For

TABLE VI. Polynomial coefficients.

term	ε	α	d
intercept	$+0.019$	-0.104	$+0.163$
$\widetilde{\log T_1}$	-0.049	-0.866	-1.445
$\widetilde{\log T_2}$	$+0.203$	$+0.876$	$+1.461$
$\widetilde{\tau}$	-0.011	$+0.011$	-0.032
$\widetilde{\log T_1}^2$	$+5.599$	-1.430	-6.083
$\widetilde{\log T_2}^2$	$+6.385$	-0.677	-4.940
$\widetilde{\tau}^2$	$+0.005$	-0.002	$+0.001$
$\widetilde{\log T_1} \cdot \widetilde{\log T_2}$	-12.002	$+2.112$	$+10.958$
$\widetilde{\log T_1} \cdot \widetilde{\tau}$	$+0.383$	$+0.038$	$+0.054$
$\widetilde{\log T_2} \cdot \widetilde{\tau}$	-0.483	-0.013	-0.013

$\hat{\varepsilon}$ the linear terms are small and the second-order terms dominate, consistent with ε being controlled by the difference $T_2 - T_1^2$.

[1] T. C. Lieuwen and V. Yang, *Combustion Instabilities in Gas Turbine Engines: Operational Experience, Fundamental Mechanisms, and Modeling*, Progress in Aeronautics and Astronautics, Vol. 210 (AIAA, 2005).

[2] T. Poinsot, Prediction and control of combustion instabilities in real engines, Proc. Combust. Inst. **36**, 1 (2017).

[3] M. P. Juniper and R. I. Sujith, Sensitivity and nonlinearity of thermoacoustic oscillations, Annu. Rev. Fluid Mech. **50**, 661 (2018).

[4] N. Noiray and B. Schuermans, Deterministic quantities characterizing noise driven Hopf bifurcations in gas turbine combustors, Int. J. Non-Linear Mech. **50**, 152 (2013).

[5] N. Noiray, Linear growth rate estimation from dynamics and statistics of acoustic signal envelope in turbulent combustors, J. Eng. Gas Turbines Power **139**, 041503 (2017).

[6] B. van der Pol and J. van der Mark, The heartbeat considered as a relaxation oscillation, and an electrical model of the heart, Phil. Mag. **7(S6)**, 763 (1928).

[7] K. Grudziński and J. J. Żebrowski, Modeling cardiac pacemakers with relaxation oscillators, Physica A **336**, 153 (2004).

[8] R. FitzHugh, Impulses and physiological states in theoretical models of nerve membrane, Biophys. J. **1**, 445 (1961).

[9] J. Nagumo, S. Arimoto, and S. Yoshizawa, An active pulse transmission line simulating nerve axon, Proc. IRE **50**, 2061 (1962).

[10] R. Stoop, T. Jasa, Y. Uwate, and S. Martignoli, From hearing to listening: Design and properties of an actively tunable electronic hearing sensor, Sensors **7**, 3287 (2006).

[11] R. Belousov, F. Berger, and A. J. Hudspeth, Volterra-series approach to stochastic nonlinear dynamics: Linear response of the Van der Pol oscillator driven by white noise, Phys. Rev. E **102**, 032209 (2020).

[12] J. Venkatramani, S. Sarkar, and S. Gupta, Intermittency in pitch-plunge aeroelastic systems explained through stochastic bifurcations, Nonlinear Dyn. **92**, 1225 (2018).

[13] A. P. Kuznetsov, S. P. Kuznetsov, and N. V. Stankevich, A simple autonomous quasiperiodic self-oscillator, Commun. Nonlinear Sci. Numer. Simul. **15**, 1676 (2010).

[14] N. Noiray and A. Denisov, A method to identify thermoacoustic growth rates in combustion chambers from dynamic pressure time series, Proc. Combust. Inst. **36**, 3843 (2017).

[15] I. V. Sysoev and B. P. Bezruchko, Noise robust approach to reconstruction of van der Pol-like oscillators and its application to Granger causality, Chaos **31**, 083118 (2021).

[16] J. Timmer, Parameter estimation in nonlinear stochastic differential equations, Chaos Solitons Fractals **11**, 2571 (2000).

[17] E. Boujo and N. Noiray, Robust identification of harmonic oscillator parameters using the adjoint Fokker–Planck equation, Proc. R. Soc. A **473**, 20160894 (2017).

[18] S. J. Lade, Finite sampling interval effects in Kramers–Moyal analysis, Phys. Lett. A **373**, 3705 (2009).

[19] C. Honisch and R. Friedrich, Estimation of Kramers–Moyal coefficients at low sampling rates, Phys. Rev. E **83**, 066701 (2011).

[20] G. Bonciolini, E. Boujo, and N. Noiray, Output-only parameter identification of a colored-noise-driven Van-der-Pol oscillator: Thermoacoustic instabilities as an example, Phys. Rev. E **95**, 062217 (2017).

[21] M. Lee, V. Gupta, and L. K. B. Li, System identification of a low-density jet via its noise-

induced dynamics, J. Fluid Mech. **862**, 200 (2019).

[22] M. Lee, Y. Guan, V. Gupta, and L. K. B. Li, Input–output system identification of a thermoacoustic oscillator near a Hopf bifurcation using only fixed-point data, Phys. Rev. E **101**, 013102 (2020).

[23] T. Indlekofer, A. Faure-Beaulieu, J. R. Dawson, and N. Noiray, Dynamics of azimuthal thermoacoustic modes in imperfectly symmetric annular combustors, J. Fluid Mech. **944**, A15 (2022).

[24] H. Son and M. Lee, A PINN approach for identifying governing parameters of noisy thermoacoustic systems, J. Fluid Mech. **984**, A21 (2024).

[25] M. Xie, X. Zhao, D. Zhao, J. Fu, C. Shelton, and B. Semlitsch, Predicting bifurcation and amplitude death characteristics of thermoacoustic instabilities from PINNs-derived van der Pol oscillators, J. Fluid Mech. **998**, A46 (2024).

[26] S. Mahesh, A. Mukhopadhyay, H. Son, and M. Lee, Learning thermoacoustic interactions in combustors using a physics-informed neural network, Eng. Appl. Artif. Intell. **138**, 109388 (2024).

[27] T. Sugandi, B. Dharmaputra, and N. Noiray, A deep operator network for Bayesian parameter identification of self-oscillators, Data-Centric Eng. **5**, e35 (2024).

[28] A. Nóvoa and L. Magri, Real-time thermoacoustic data assimilation, J. Fluid Mech. **948**, A35 (2022).

[29] A. Nóvoa, N. Noiray, J. R. Dawson, and L. Magri, A real-time digital twin of azimuthal thermoacoustic instabilities, J. Fluid Mech. **1001**, A49 (2024).

[30] R. L. Stratonovich, *Topics in the Theory of Random Noise*, Vol. 1 (Gordon and Breach, 1963).

[31] R. Z. Khasminskii, On the principle of averaging the Itô stochastic differential equations, Kybernetika **4**, 260 (1968).

[32] C. W. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*, 4th ed. (Springer, 2009).

[33] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (Elsevier (North-Holland), Amsterdam, 2007).

[34] H. Risken, *The Fokker–Planck Equation: Methods of Solution and Applications*, 2nd ed. (Springer, 1996).

[35] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed. (Springer, 1998).

[36] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. U.S.A. **113**, 3932 (2016).

[37] L. Boninsegna, F. Nüske, and C. Clementi, Sparse learning of stochastic dynamical equations, J. Chem. Phys. **148**, 241723 (2018).

[38] J. L. Callaham, J.-C. Loiseau, G. Rigas, and S. L. Brunton, Nonlinear stochastic modelling with Langevin regression, Proc. R. Soc. A **477**, 20210092 (2021).

[39] D. A. Messenger and D. M. Bortz, Weak SINDy: Galerkin-based data-driven model selection, Multiscale Model. Simul. **19**, 1474 (2021).

[40] P. Subramanian, R. I. Sujith, and P. Wahi, Subcritical bifurcation and bistability in thermoacoustic systems, J. Fluid Mech. **715**, 210 (2013).

[41] V. Nair, G. Thampi, and R. I. Sujith, Intermittency route to thermoacoustic instability in turbulent combustors, J. Fluid Mech. **756**, 470 (2014).

[42] L. Magri, M. P. Juniper, and J. P. Moeck, Sensitivity of the Rayleigh criterion in thermoacoustics, J. Fluid Mech. **882**, R1 (2020).

[43] A. Krishnan, R. Manikandan, P. Miber, R. I. Sujith, N. Marwan, and J. Kurths, Suppression of thermoacoustic instability by targeting the hubs of the turbulent networks in a bluff-body stabilized combustor, J. Fluid Mech. **916**, A20 (2021).

[44] K. Kashinath, L. K. B. Li, and M. P. Juniper, Forced synchronization of periodic and aperiodic thermoacoustic oscillations: lock-in, bifurcations and open-loop control, J. Fluid Mech. **838**, 690 (2018).