Predicting Forced Responses of Probability Distributions via the Fluctuation–Dissipation Theorem and Generative Modeling

Ludovico T. Giorgini[1,*], Fabrizio Falasca[2], and Andre N. Souza[3]

[1]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA
[3]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[*]Corresponding author: `ludogio@mit.edu`

April 21, 2025

Abstract

We present a novel data-driven framework for estimating the response of higher-order moments of nonlinear stochastic systems to small external perturbations. The classical Generalized Fluctuation–Dissipation Theorem (GFDT) links the unperturbed steady-state distribution to the system's linear response. Standard implementations rely on Gaussian approximations, which can often accurately predict the mean response but usually introduce significant biases in higher-order moments—such as variance, skewness, and kurtosis. To address this limitation, we combine GFDT with recent advances in score-based generative modeling, which enable direct estimation of the score function from data without requiring full density reconstruction. Our method is validated on three reduced-order stochastic models relevant to climate dynamics: a scalar stochastic model for low-frequency climate variability, a slow–fast triad model mimicking key features of the El Niño-Southern Oscillation (ENSO), and a six-dimensional stochastic barotropic model capturing atmospheric regime transitions. In all cases, the approach captures strongly nonlinear and non-Gaussian features of the system's response, outperforming traditional Gaussian approximations.

Significance Statement

Predicting how complex stochastic systems respond to small external perturbations is central in physics, climate science, and engineering. We combine the Generalized Fluctuation–Dissipation Theorem with score-based generative modeling (KGMM) to accurately capture mean and higher-order (variance, skewness, kurtosis) responses—even in strongly non-Gaussian regimes.

1 Introduction

Understanding how a physical system responds to external perturbations is a central challenge in physics, climate science, and engineering [Palmer, 2001, Strogatz, 2018]. In many applications, one is not only interested in the shift of the system's *mean* state under a small forcing but also in changes in its *higher-order moments*—such as variance, skewness, and kurtosis. Capturing responses in higher-order moments is crucial for reconstructing the perturbed steady-state probability distribution and for building a principled, causal framework to analyze how perturbations shape the system's overall response, including distributional tails [Majda et al., 2010b]. For instance, in climate science, assessing the full redistribution of probability (e.g., changes in the frequency of heat waves or cold spells) is as crucial as estimating an average temperature increase [IPCC, 2013].

A key theoretical tool in this context is the *Generalized Fluctuation-Dissipation Theorem* (GFDT), which relates the steady-state probability density function (PDF) of a system to its linear response under a time-dependent *external* perturbation [Baldovin et al., 2020, Cooper and Haynes, 2011, Ghil and Lucarini, 2020, Giorgini et al., 2024b, Majda et al., 2005, 2008b]. In principle, GFDT offers a route to predict how the entire distribution of an observable changes without explicitly perturbing the system. However, practical application of the GFDT is hindered by the difficulty of accurately estimating the full, often high-dimensional and non-Gaussian, steady-state PDF. In many cases the *Gaussian* approximation—where the system's PDF is replaced by a Gaussian distribution sharing the same mean and covariance—is employed. It has been shown empirically that the Gaussian approximation has high-skill in predicting responses in the mean even in nonlinear systems (e.g. Baldovin et al. [2020], Gershgorin and Majda [2010], Gritsun and Branstator [2007]) but introduces systematic biases for higher-order moments [Majda and Qi, 2019, Majda et al., 2008c].

Recent advances in data-driven methodologies offer promising alternatives. In particular, *score-based generative modeling* has emerged as a powerful approach for sampling from complex distributions by learning the gradient of the log-PDF (the *score function*) directly from data, thereby circumventing the need for full density estimation [Song et al., 2021]. Moreover, modern clustering-based algorithms have shown that statistical properties of high-dimensional systems can be estimated more reliably from coarse-grained observational data than the systems' detailed dynamical trajectories [Falasca et al., 2025, 2024, Giorgini et al., 2024a, 2025b, Souza, 2024a,b, Souza and Silvestri, 2024].

These data-driven advances are particularly valuable in settings where directly solving the full partial differential equations (PDEs) governing a system's dynamics is computationally prohibitive or even unknown. This limitation has motivated the development of reduced-order models, which aim to reproduce the essential statistical behavior of complex systems at a fraction of the computational cost. In climate science, in particular, stochastic reduced-order models have played a central role in both understanding and forecasting system behavior [Hasselmann, 1976, Kondrashov et al., 2015, Lucarini and Chekroun, 2023, Majda et al., 1999, 2001, Penland, 1989, Penland and Sardeshmukh,

1995, Strounine et al., 2010]. For example, reduced Markovian models have been successfully employed to derive effective dynamics for slow variables in multiscale systems [Chen et al., 2019, Giorgini et al., 2022, Keyes et al., 2023, Kravtsov et al., 2005, Majda et al., 2008a]. These approaches facilitate our understanding of high-dimensional dynamical systems by focusing on core processes while also improving computational efficiency. Furthermore, reduced-order models also shed light on how non-Gaussian features and intermittent behavior emerge from interactions between resolved and unresolved scales.

Building on these ideas, recent studies have combined the GFDT framework with generative score-based techniques to improve predictions of the mean response in high-dimensional systems [Giorgini et al., 2024b]. In this paper, we build on the framework presented in [Giorgini et al., 2024b] by constructing *higher-order* response functions for nonlinear reduced-order stochastic models relevant to climate science. Specifically, we examine the response of the mean (first moment) and the second, third, and fourth central moments of an observable to small external perturbations. By incorporating information from these higher-order moments, our approach provides an accurate estimate of the perturbed steady-state PDF, which is crucial for understanding changes in variability and the occurrence of extreme events. We test the approach on the following reduced-order stochastic models of increasing complexity: a one-dimensional stochastic model for low-frequency climate variability, a slow–fast triad model designed to capture key features of the El Niño-Southern Oscillation (ENSO), and a six-dimensional stochastic barotropic model that reproduces atmospheric regime transitions [Charney and DeVore, 1979, Crommelin et al., 2004, De Swart, 1988]. Our results demonstrate that the new method can accurately reproduce the steady-state PDFs and outperforms traditional Gaussian approximations in capturing higher-order moment responses.

The remainder of the paper is organized as follows. In Section 2, we review the derivation and practical implementation of the GFDT and we detail the extraction of the score function from data through a recently proposed clustering-based algorithms. In Section 3, we present numerical experiments on the three reduced-order models, highlighting the improved skill of our approach in capturing higher-order responses compared to the standard Gaussian approximation. Section 4 focuses on limitations and practical considerations of our proposal in the presence of partial observations from high-dimensional dynamical systems. Finally, Section 5 summarizes our findings and discusses directions for future work.

2   Methods

In this section, we present the theoretical and computational foundations of our approach. We begin by deriving the Generalized Fluctuation-Dissipation Theorem (GFDT), which forms the basis for computing response functions in nonlinear stochastic systems. We describe the KGMM (K-means Gaussian Mixture Modeling) algorithm from [Giorgini et al., 2025a] for data-driven estimation of the score function, which plays a central role in implementing the GFDT for this work. Additional details on the maximum entropy framework used to reconstruct perturbed steady-state distributions from estimated moment responses are provided in the Appendix.

2.1 The Generalized Fluctuation-Dissipation Theorem

We consider a stochastic dynamical system governed by

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{F}(\boldsymbol{x}) + \boldsymbol{\sigma}\,\boldsymbol{\xi}(t), \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{F}(\boldsymbol{x})$ is a deterministic drift term, $\boldsymbol{\sigma}$ is the noise amplitude matrix, and $\boldsymbol{\xi}$ is a standard Gaussian white noise process. The system evolves towards a steady state described by a probability density $\rho_S(\boldsymbol{x})$, which satisfies the stationary Fokker–Planck equation:

$$\mathcal{L}_0\,\rho_S(\boldsymbol{x}) = 0, \tag{2}$$

where the Fokker–Planck operator $\mathcal{L}_0$ is defined as

$$\mathcal{L}_0\,\rho(\boldsymbol{x}) = -\nabla \cdot \big(\boldsymbol{F}(\boldsymbol{x})\,\rho(\boldsymbol{x})\big) + \frac{1}{2}\nabla\nabla^\top : \big(\boldsymbol{D}\,\rho(\boldsymbol{x})\big), \tag{3}$$

with $\boldsymbol{D} = \boldsymbol{\sigma}\boldsymbol{\sigma}^\top$ the diffusion matrix, and : denoting the double contraction of tensors.

We now introduce a small, time-dependent perturbation that factorizes into a small spatial component $\boldsymbol{u}(\boldsymbol{x})$ and a temporal modulation $f(t)$, so that the dynamics become

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{F}(\boldsymbol{x}) + \boldsymbol{u}(\boldsymbol{x})\,f(t) + \boldsymbol{\sigma}\,\boldsymbol{\xi}(t). \tag{4}$$

This perturbation induces a small variation in the probability density, denoted $\delta\rho(\boldsymbol{x}, t)$, which satisfies, to first order in $\boldsymbol{u}(\boldsymbol{x})$,

$$\frac{\partial\delta\rho(\boldsymbol{x}, t)}{\partial t} = \mathcal{L}_0\,\delta\rho(\boldsymbol{x}, t) + f(t)\,\mathcal{L}_1\,\rho_S(\boldsymbol{x}), \tag{5}$$

with the perturbation operator $\mathcal{L}_1$ defined by

$$\mathcal{L}_1\,\rho(\boldsymbol{x}) = -\nabla \cdot \big(\boldsymbol{u}(\boldsymbol{x})\,\rho(\boldsymbol{x})\big). \tag{6}$$

By formally integrating this linearized equation, the resulting first-order correction to the expectation value of any observable $A(\boldsymbol{x})$ is given by

$$\langle\delta A(t)\rangle = \int_0^t f(t')\,\big\langle A(\boldsymbol{x}(t))\,B(\boldsymbol{x}(t'))\big\rangle_0 \mathrm{d}t' = \int_0^t \boldsymbol{R}(t - t')\,f(t')\,\mathrm{d}t', \tag{7}$$

where the conjugate observable $B(\boldsymbol{x})$ is

$$B(\boldsymbol{x}) = \frac{\mathcal{L}_1\,\rho_S(\boldsymbol{x})}{\rho_S(\boldsymbol{x})} = -\frac{\nabla \cdot \big(\boldsymbol{u}(\boldsymbol{x})\,\rho_S(\boldsymbol{x})\big)}{\rho_S(\boldsymbol{x})}, \tag{8}$$

and $\boldsymbol{R}(t) = \langle A(\boldsymbol{x}(t))\,B(\boldsymbol{x}(0))\rangle_0$ is the linear response function, by definition equal to the response to an impulse (delta) perturbation $f(t) = \delta(t)$ Risken [1996]. $\langle\cdot\rangle_0$ denotes the expectation with respect to the unperturbed steady-state distribution $\rho_S(\boldsymbol{x})$.

Applying the product rule, we obtain

$$\nabla \cdot \big(\boldsymbol{u}(\boldsymbol{x})\,\rho_S(\boldsymbol{x})\big) = \rho_S(\boldsymbol{x})\,\nabla \cdot \boldsymbol{u}(\boldsymbol{x}) + \boldsymbol{u}(\boldsymbol{x}) \cdot \nabla\rho_S(\boldsymbol{x}), \tag{9}$$

and thus the conjugate observable takes the explicit form

$$B(\boldsymbol{x}) = -\nabla \cdot \boldsymbol{u}(\boldsymbol{x}) - \boldsymbol{u}(\boldsymbol{x}) \cdot \nabla\ln\rho_S(\boldsymbol{x}). \tag{10}$$

Note that climate studies traditionally focus on the specific case of changes in the external forcing as $\boldsymbol{u}(\boldsymbol{x})_i = \boldsymbol{e}_i$, further simplifying Eq. 10 to $B(\boldsymbol{x}) = -\nabla \ln \rho_S(\boldsymbol{x})$.

Often the unperturbed steady state is approximated with a multivariate Gaussian. In this case we have

$$\rho_S(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n \det \boldsymbol{\Sigma}}} \exp\left[-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right], \tag{11}$$

which yields

$$\ln \rho_S(\boldsymbol{x}) = -\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) + \text{const}, \tag{12}$$

and consequently,

$$\nabla \ln \rho_S(\boldsymbol{x}) = -\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}). \tag{13}$$

Substituting into the expression for $B(\boldsymbol{x})$, we obtain

$$B(\boldsymbol{x}) = -\nabla \cdot \boldsymbol{u}(\boldsymbol{x}) + \boldsymbol{u}(\boldsymbol{x})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}). \tag{14}$$

In this case, it is straightforward to verify that, under external forcings $\boldsymbol{u}(\boldsymbol{x})_i = \boldsymbol{e}_i$ the response function reduces to $\boldsymbol{R}(t) = \langle A(\boldsymbol{x}(t))\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})(0)\rangle_0$, as commonly adopted in climate studies. Note that while the analytical score is approximated using that of a Gaussian distribution, the ensemble averages $\langle \cdot \rangle_0$ are still obtained using the original data and therefore over the system's steady state distribution. This approach differs from linear inverse model strategies (e.g., Penland [1989]), which assume a time-invariant Gaussian measure from the outset. Such models would yield zero response to any quadratic functional, such as the variance (see Majda et al. [2010a]). Because of this distinction, the approximation discussed here is sometimes referred to as a quasi-Gaussian approximation in the literature. For simplicity, we will simply refer to it as the Gaussian approximation throughout this work.

2.2 Score Function Estimation via KGMM

The application of the Generalized Fluctuation-Dissipation Theorem (GFDT) requires knowledge of the score function $\nabla \ln \rho_S(x)$, i.e., the gradient of the logarithm of the system's steady-state distribution. For the vast majority of systems of interest, this quantity cannot be obtained analytically and must be inferred from data. To this end, a hybrid statistical-learning method called KGMM (K-means Gaussian Mixture Modeling) has been recently proposed Giorgini et al. [2025a] for accurate and efficient score function estimation in systems with low-dimensional *effective* dynamics.

The KGMM method is based on the observation that a probability density can be approximated as a Gaussian Mixture Model (GMM):

$$p(\boldsymbol{x}) = \frac{1}{N} \sum_{i=1}^{N} \mathcal{N}(\boldsymbol{x} \mid \boldsymbol{\mu}_i, \sigma^2 \boldsymbol{I}), \tag{15}$$

where the $\boldsymbol{\mu}_i$ are data samples drawn from the steady-state distribution $\rho_S(\boldsymbol{x})$, and σ^2 is the (isotropic) covariance amplitude of the Gaussian kernels. The corresponding score function reads

$$\nabla \ln p(\boldsymbol{x}) = -\frac{1}{\sigma^2} \sum_{i=1}^{N} \frac{\mathcal{N}(\boldsymbol{x} \mid \boldsymbol{\mu}_i, \sigma^2 \boldsymbol{I})(\boldsymbol{x} - \boldsymbol{\mu}_i)}{p(\boldsymbol{x})}. \tag{16}$$

While Eq. (16) provides a direct expression for the score, it becomes numerically unstable for small σ, as the density and its derivative become highly sensitive to local data fluctuations.

KGMM circumvents this issue by exploiting a probabilistic identity: define $\boldsymbol{x} = \boldsymbol{\mu} + \sigma\boldsymbol{z}$ with $\boldsymbol{z} \sim \mathcal{N}(0, \boldsymbol{I})$. Giorgini et al. [2025a] derived the following relation in the continuous-data limit:

$$\nabla \ln p(\boldsymbol{x}) = -\frac{1}{\sigma^2}\mathbb{E}[\boldsymbol{z} \mid \boldsymbol{x}], \tag{17}$$

i.e., the score function is the conditional expectation of the kernel displacements $\boldsymbol{z}$, rescaled by σ^2.

This identity is computed empirically using the following steps:

1. Generate perturbed samples $\boldsymbol{x}_i = \boldsymbol{\mu}_i + \sigma\boldsymbol{z}_i$, where $\boldsymbol{\mu}_i$ are the original data points and $\boldsymbol{z}_i \sim \mathcal{N}(0, \boldsymbol{I})$.

2. Partition the perturbed samples $\{\boldsymbol{x}_i\}$ into N_C clusters $\{\Omega_j\}$ using bisecting K-means clustering.

3. For each cluster Ω_j with centroid $\boldsymbol{C}_j$, compute the conditional expectation

$$\mathbb{E}[\boldsymbol{z} \mid \boldsymbol{x} \in \Omega_j] \approx \frac{1}{|\Omega_j|} \sum_{i:\boldsymbol{x}_i\in\Omega_j} \boldsymbol{z}_i. \tag{18}$$

4. Estimate the score function at the centroid $\boldsymbol{C}_j$ as

$$\nabla \ln \rho_S(\boldsymbol{C}_j) \approx -\frac{1}{\sigma^2}\mathbb{E}[\boldsymbol{z} \mid \boldsymbol{x} \in \Omega_j]. \tag{19}$$

5. Fit a neural network to interpolate the discrete estimates $\{(\boldsymbol{C}_j, \nabla \ln \rho_S(\boldsymbol{C}_j))\}$ over the full domain.

The number of clusters N_C must therefore be chosen carefully to balance the trade-off between resolution and noise. Empirically, a useful scaling relation is

$$N_C \propto \sigma^{-d}, \tag{20}$$

where d is the effective dimensionality of the dataset and σ is the kernel width. This scaling ensures that clusters remain small enough to capture local gradient structure while still containing enough points for robust averaging.

The choice of σ plays a central role in the KGMM algorithm. Small values of σ yield estimates of the score function that are closer to the one associated with the true steady-state distribution, as the perturbation introduced by the convolution kernel becomes negligible. However, this comes at the cost of increased statistical noise, since the displacements become more sensitive to sample variability. Conversely, larger values of σ smooth out the fluctuations, leading to more stable estimates but of a score function associated with a more strongly perturbed distribution. The optimal value of σ thus balances these competing effects—reducing bias while maintaining statistical reliability.

To interpolate the discrete score estimates $(\boldsymbol{C}_j, \nabla \ln \rho_S(\boldsymbol{C}_j))$, we train a fully connected feedforward neural network with two hidden layers, using the Swish activation function between layers and a linear activation on the output layer. The models and training parameters are:

- **Scalar model:** Hidden layers of 50 and 25 neurons; batch size 32; 2000 epochs.

- **Slow–fast triad model:** Hidden layers of 100 and 50 neurons; batch size 32; 200 epochs.

- **Barotropic model:** Hidden layers of 128 and 64 neurons; batch size 128; 300 epochs.

All networks are trained using the Adam optimizer with mean-squared-error loss on the predicted score.

3   Results

In what follows, we estimate the mean (first moment) and the second, third, and fourth central moments of the perturbed steady-state distribution for three representative non-linear stochastic systems of increasing complexity: a scalar stochastic model for low-frequency climate variability, a slow-fast triad model mimicking essential features of the El Niño-Southern Oscillation (ENSO), and a six-dimensional stochastic barotropic model capturing atmospheric regime transitions. These systems serve as ideal testbeds for evaluating the accuracy of different approximations to the Generalized Fluctuation-Dissipation Theorem (GFDT) in predicting higher-order responses under small external perturbations.

For each system, we construct the response functions as defined in Eq. (7). Specifically, we focus on impulse response functions for the mean (first moment) and the second, third, and fourth central moments of the invariant distribution by choosing as observables $\boldsymbol{x}$, $(\boldsymbol{x}-\boldsymbol{\mu})^2$, $(\boldsymbol{x}-\boldsymbol{\mu})^3$, and $(\boldsymbol{x}-\boldsymbol{\mu})^4$, where $\boldsymbol{x}$ denotes the system state and $\boldsymbol{\mu}$ its unperturbed mean. Specifically, we assess three distinct approaches: (i) a data-driven method that constructs the response using the score function estimated via the KGMM (K-means Gaussian Mixture Modeling) algorithm; (ii) a Gaussian approximation, which assumes the steady-state distribution to be multivariate Gaussian, fully characterized by its mean and covariance; and (iii) a reference or "ground truth" response. In the case of the scalar triad model, the ground truth has been derived analytically in Majda et al. [2009b] from the known stationary distribution. For the slow-fast triad and stochastic barotropic models, where an analytic form of the score is not available, the ground truth response is computed numerically by performing ensemble integrations of both the unperturbed and perturbed systems and evaluating the change in the moments, see for example the Appendix in Boffetta et al. [2003] or Giorgini et al. [2024b]. In all examples, we normalize the data such that the unperturbed time series in each dimension has zero mean and unit variance. This preprocessing step facilitates the comparison across variables and models, and ensures that the Gaussian approximation always corresponds to a standard normal distribution. For systems subject to a state-independent perturbation, as in the first and third examples, we normalize the response functions by the perturbation amplitude, rendering them independent of its magnitude.

In settings where direct computation of response functions is infeasible—either because the underlying equations are unknown or integration is prohibitively expensive—the only reliable way to validate the KGMM-estimated score (and, by extension, any response predictions) is to compare observed and KGMM-generated marginal PDFs. Concretely, we simulate the Langevin dynamics

$$\dot{\boldsymbol{x}}(t) \;=\; \boldsymbol{s}(\boldsymbol{x}(t)) \;+\; \sqrt{2}\,\boldsymbol{\xi}(t), \tag{21}$$

where $\boldsymbol{s}(\boldsymbol{x}) = \nabla \ln \rho_S(\boldsymbol{x})$ is the score learned via KGMM. The degree to which the simulated marginals reproduce the empirically observed steady-state PDFs thus provides the sole stringent diagnostic that the score has been learned accurately and that any inferred response functions can be trusted. For this reason, we have also compared the marginal steady-state PDFs obtained from these Langevin simulations directly against the empirical distributions, offering an independent check on the fidelity of the estimated score function.

This comparative analysis allows us to quantify the extent to which generative score modeling improves the accuracy of higher-order response predictions relative to traditional linear approximations, especially in regimes characterized by strong non-Gaussianity and nonlinear interactions.

3.1 Scalar Stochastic Model for Low-Frequency Climate Variability

We consider a one-dimensional stochastic model for low-frequency climate variability. The model was originally derived by Majda et al. [2009b] using stochastic reduction techniques developed in Majda et al. [1999, 2001]. For further details on the derivation and relevant literature, we refer the reader to Section 4b of Majda et al. [2010a]. This reduced-order stochastic model has been successfully used to fit nonlinear scalar dynamics from low-frequency data of a general circulation model and has also served as a testbed for fluctuation-dissipation theorem (FDT) analyses, see for example [Majda et al., 2010a, 2009a] in the context of the Gaussian approximation. The model is given by the scalar nonlinear stochastic differential equation:

$$\dot{x}(t) = F + ax(t) + bx^2(t) - cx^3(t) + \sigma_1\,\xi_1(t) + \sigma_2(x)\xi_2(t), \tag{22}$$

where $\xi_1(t)$ and $\xi_2(t)$ are independent standard white noise processes. In the present work, we consider a closely related scalar model introduced in [Chen et al., 2019], which builds upon similar principles but features no multiplicative noise, that is, $\sigma_2(x) = 0$. This leads to a simplified, purely additive stochastic model of the form:

$$\dot{x}(t) = F + ax(t) + bx^2(t) - cx^3(t) + \sigma\,\xi(t), \tag{23}$$

where $\xi(t)$ is standard white noise. This scalar nonlinear model retains the essential deterministic structure of the reduced triad system, while simplifying the stochastic component, thus making it an ideal testbed for analytical and numerical studies.

The coefficients of the model used in this work are listed in Table 1. The input parameters of the KGMM algorithm (see Section 2.2) are set to $\sigma = 0.05$ and $N_C = 348$.

Table 1: Parameters of the stochastic scalar model.

a	-0.0222
b	-0.2
c	0.0494
F	0.6
σ	0.7071

A key motivation for considering this model is the availability of an analytical expression for the score function, defined as the gradient of the logarithm of the stationary distribution:

$$s(x) = \frac{d}{dx} \log \rho_S(x) = \frac{2}{\sigma^2} \left(F + ax + bx^2 - cx^3 \right). \qquad (24)$$

Correspondingly, the stationary probability density function (PDF) of the system can be written in closed form as:

$$\rho_S(x) = \mathcal{N}^{-1} \exp \left[\frac{2}{\sigma^2} \left(Fx + \frac{a}{2}x^2 + \frac{b}{3}x^3 - \frac{c}{4}x^4 \right) \right], \qquad (25)$$

where $\mathcal{N}$ is a normalization constant ensuring that $\rho_S(x)$ integrates to 1.

We now construct impulse response functions through GFDT as $R(t) = \langle A(x(t)) B(x(0)) \rangle_0$, and remind the reader that $\langle \cdot \rangle_0$ denotes the expectation with respect to the unperturbed steady-state distribution $\rho_S(x)$. We consider the simple case where the conjugate observable is defined as $B(x) = -s(x) = -\frac{d}{dx} \log \rho_S(x)$. To estimate the score function $s(x)$ we consider three different strategies: (i) analytically from Eq. (24), (ii) using the KGMM method, and (iii) under the Gaussian approximation. In Figure 1, we compare the unperturbed steady-state PDF as well as the first four moment response functions obtained with these three approaches. The results show that the score function estimated via KGMM yields a nearly perfect match with the analytic response, both in terms of the steady-state PDF and the responses of the mean (first moment) and the second, third, and fourth central moments. This confirms the ability of KGMM to accurately reconstruct the score function. As expected from previous studies (e.g. [Baldovin et al., 2020]), the Gaussian approximation provides a very good estimation for changes in the mean. However, significant biases are present in the second, third and fourth centralized moments, reflecting its inability to capture the non-Gaussian features of the underlying dynamics.
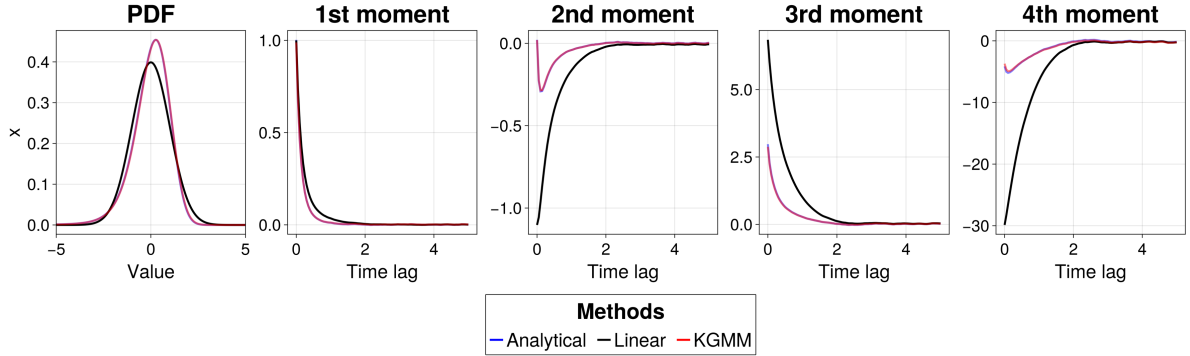

Figure 1: Left: True unperturbed PDF (blue) compared with a normal Gaussian (black) and the PDF obtained integrating Eq. (21) with the KGMM score function (red). Right: Response functions predicted via GFDT. The true response functions (blue) are obtained using the analytic score function and are compared with the ones obtained using the KGMM score function (red) and the linear approximation (black).

3.2 Slow-Fast Triad Model and Application to ENSO

The second test case considers the slow-fast triad model from [Thual et al., 2016]:

$$
\begin{aligned}
\dot{u}_1 &= -d_u u_1 - \omega u_2 + \tau + \sigma_{u1}\,\xi_1(t), \\
\dot{u}_2 &= -d_u u_2 + \omega u_1 + \sigma_{u2}\,\xi_2(t), \\
\dot{\tau} &= -d_\tau\,\tau + \sigma_\tau(u_1)\,\xi_3(t),
\end{aligned}
\tag{26}
$$

with parameters listed in Table 2. The input parameters of the KGMM algorithm (see Section 2.2) are set to $\sigma = 0.05$ and $N_C = 7353$.

d_u	0.2
ω	0.4
d_τ	2.0
$\sigma_{u1} = \sigma_{u2}$	0.3
$\sigma_\tau(u_1)$	$1.5(\tanh(u_1) + 1)$

Table 2: Parameters of the slow-fast triad model.

This model is designed to mimic multi-scale interactions characteristic of the El Niño-Southern Oscillation (ENSO), where the slow variables u_1 and u_2 represent the large-scale ocean-atmosphere state, and the fast variable τ represents rapid wind bursts. Differently from the previous Section, here we are interested in studying the system's response to perturbations in the damping coefficient of the fast variable τ. Specifically, the perturbation is applied as:

$$
\delta d_\tau = -0.2 d_\tau = -0.4.
\tag{27}
$$

This modification increases the memory time of τ (the wind bursts), thereby enhancing its strength. Physically, this perturbation is interpreted as an increase in Madden-Julian oscillation or monsoon activity in the Western Pacific, which in turn modifies the behavior of ENSO. A direct consequence of this perturbation is an increase in the variance of u_1 and u_2, leading to a higher occurrence of strong ENSO events.

In this case, the conjugate observable needs to be considered in its full formulation as $B(\boldsymbol{x}) = -\nabla \cdot \boldsymbol{u}(\boldsymbol{x}) - \boldsymbol{u}(\boldsymbol{x}) \cdot \nabla \ln \rho_S(\boldsymbol{x})$. Unlike the triad model, no analytic expression for the steady-state distribution or score function is available for this system. Therefore, the ground truth response was computed by directly simulating an ensemble of both unperturbed and perturbed systems and evaluating the change in moments (see e.g. Giorgini et al. [2024b]). These serve as a benchmark for validating the GFDT-based predictions. As in the previous example, we compared the response of the mean (first moment) and the second, third, and fourth central moments of each variable using three methods: GFDT with the KGMM-estimated score function, GFDT under the Gaussian linear approximation, and the numerical ground truth. We also compared the unperturbed marginal steady-state PDFs predicted by KGMM to the empirically observed distributions by integrating the Langevin equation in Eq. 21.

The results are reported in Figure 2. In all three dimensions, the score function estimated by KGMM allows the GFDT framework to accurately reproduce the response functions and the steady-state PDFs. The agreement with the numerically computed

ground truth is excellent, indicating that the KGMM method successfully reconstructs the non-Gaussian statistical structure of the system. In contrast, the linear approximation yields visible discrepancies in all moments, especially for the strongly skewed and heavy-tailed distribution of τ.
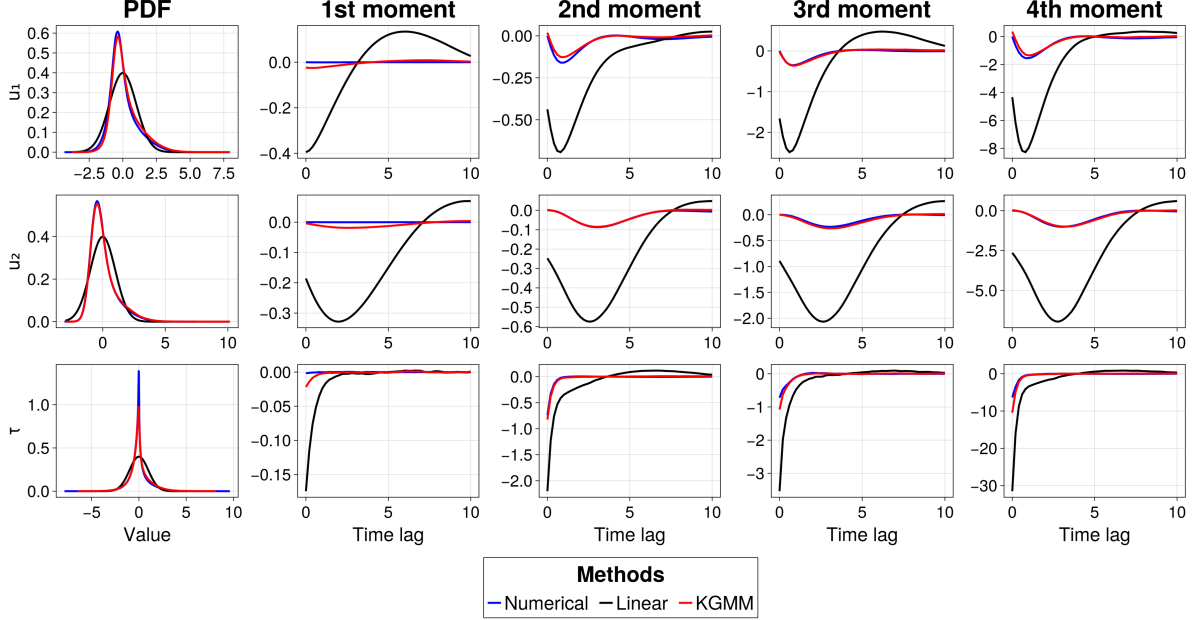

Figure 2: Comparison of unperturbed marginal steady-state PDFs (leftmost column) and responses of the first four moments (remaining columns) in the slow-fast triad model for variables u_1, u_2, and τ (rows). Response functions are shown as a function of time lag. Predictions obtained via GFDT using the KGMM-estimated score function (red), the Gaussian approximation (black), and numerical ground truth (blue) are reported.

3.3 Stochastic Barotropic Model for Atmospheric Regime Transitions

The large-scale circulation of the atmosphere often exhibits persistent patterns, such as zonal and blocked flow regimes, with abrupt transitions between them. These features can be captured through low-order models that isolate the essential mechanisms responsible for regime persistence and transitions [Charney and DeVore, 1979]. The six-dimensional model considered here is based on a Galerkin truncation of the barotropic vorticity equation on a β-plane channel with topography. The original formulation was introduced by Charney and DeVore [Charney and DeVore, 1979], and the specific version used here follows a slightly modified form proposed by De Swart [De Swart, 1988]. This model has been widely studied, particularly in the context of stochastic regime dynamics [e.g., Crommelin et al., 2004, Dorrington and Palmer, 2023]. It retains two zonal modes and two pairs of wave modes, enabling nonlinear interactions between the waves and the mean flow, as well as capturing the influence of topographic forcing. Physically, the model describes the evolution of the barotropic streamfunction field, subject to planetary rotation, Newtonian damping toward a prescribed zonal background state, and orographic forcing. The six prognostic variables represent amplitudes of selected Fourier modes, and the equations contain linear terms corresponding to damping and rotation,

11

as well as quadratic nonlinearities encoding advective interactions. The system supports multiple equilibria and intermittent transitions between weakly chaotic and more persistent quasi-stationary states. By adding stochastic forcing, one can probe how unresolved processes affect the stability and persistence of regimes [Dorrington and Palmer, 2023]. The stochastic version of the model takes the form:

$$
\begin{aligned}
\dot{x}_1 &= \tilde{\gamma}_1 x_3 - C(x_1 - x_1^*) + \sigma \xi_1(t), \\
\dot{x}_2 &= -(\alpha_1 x_1 - \beta_1)x_3 - Cx_2 - \delta_1 x_4 x_6 + \sigma \xi_2(t), \\
\dot{x}_3 &= (\alpha_1 x_1 - \beta_1)x_2 - \gamma_1 x_1 - Cx_3 + \delta_1 x_4 x_5 + \sigma \xi_3(t), \\
\dot{x}_4 &= \tilde{\gamma}_2 x_6 - C(x_4 - x_4^*) + \varepsilon(x_2 x_6 - x_3 x_5) + \sigma \xi_4(t), \\
\dot{x}_5 &= -(\alpha_2 x_1 - \beta_2)x_6 - Cx_5 - \delta_2 x_4 x_3 + \sigma \xi_5(t), \\
\dot{x}_6 &= (\alpha_2 x_1 - \beta_2)x_5 - \gamma_2 x_4 - Cx_6 + \delta_2 x_4 x_2 + \sigma \xi_6(t).
\end{aligned}
\tag{28}
$$

Here, x_1 and x_4 denote the amplitudes of zonal modes, while the remaining variables represent wave modes. The parameter C denotes the damping rate, and (x_1^*, x_4^*) are the prescribed zonal background states. Stochastic forcing enters additively through independent Gaussian white noise processes $\xi_i(t)$ with common amplitude σ.

The values of the coefficients used in this study are summarized in Table 3, corresponding to a system with channel aspect ratio $b = 1.6$, $\beta = 1.25$, and topographic forcing strength $\gamma = 0.2$. The forcing terms $x_1^* = 0.95$ and $x_4^* = -0.76095$ define the equilibrium zonal profile. The input parameters of the KGMM algorithm (see Section 2.2) are set to $\sigma = 0.05$ and $N_C = 30309$.

Table 3: Model coefficients for the six-dimensional stochastic barotropic system.

Parameter	Value	Description
α_1	0.86322463	Nonlinear advection (mode 1)
α_2	0.81394451	Nonlinear advection (mode 2)
β_1	0.89887640	Coriolis effect (mode 1)
β_2	0.48780488	Coriolis effect (mode 2)
δ_1	1.38115941	Triad interaction (mode 1)
δ_2	-0.12882575	Triad interaction (mode 2)
$\tilde{\gamma}_1$	0.19206748	Orographic forcing (mode 1)
$\tilde{\gamma}_2$	0.07682699	Orographic forcing (mode 2)
γ_1	0.05395154	Orographic damping (mode 1)
γ_2	0.04684573	Orographic damping (mode 2)
ε	1.44050611	Wave-wave interaction
C	0.1	Newtonian relaxation rate
x_1^*	0.95	Zonal background forcing (mode 1)
x_4^*	-0.76095	Zonal background forcing (mode 2)
σ	0.01	Noise amplitude

We now construct impulse response functions through GFDT as $\boldsymbol{R}(t) = \langle A(\boldsymbol{x}(t)) B(\boldsymbol{x}(0)) \rangle_0$, and remind the reader that $\langle \cdot \rangle_0$ denotes the expectation with respect to the unperturbed steady-state distribution $\rho_S(\boldsymbol{x})$. As in Section 3.1 we consider the case where $B(\boldsymbol{x}) = -\boldsymbol{s}(\boldsymbol{x}) = -\nabla \log \rho_S(\boldsymbol{x})$. To estimate the score function $\boldsymbol{s}(\boldsymbol{x})$ we consider three

different strategies: the Gaussian linear approximation, the KGMM-estimated score function, and direct numerical simulations of the perturbed system, which serve as the ground truth. Figure 3 shows the results for all six state variables of the model, comparing the unperturbed steady-state PDFs (leftmost column) and the temporal evolution of the response of the mean (first moment) and the second, third, and fourth central moments (remaining columns) for a state-independent, impulse perturbation applied to x_1.

The KGMM-based GFDT predictions exhibit remarkable agreement with the numerical results across all variables and moment orders. The Gaussian approximation shows good estimates for changes in the mean but significant deviations for higher order moments. This is especially true for responses in the perturbed variable x_1, where the effects of the perturbation are most pronounced. These discrepancies confirm the limitations of the Gaussian approximation in representing nonlinear and asymmetric responses, which are typical in regime-transition dynamics such as blocking or wave-mean flow interactions. Notably, the KGMM approach captures both sharp peaks and heavy tails in the marginal distributions, which are characteristic of systems with intermittent transitions.
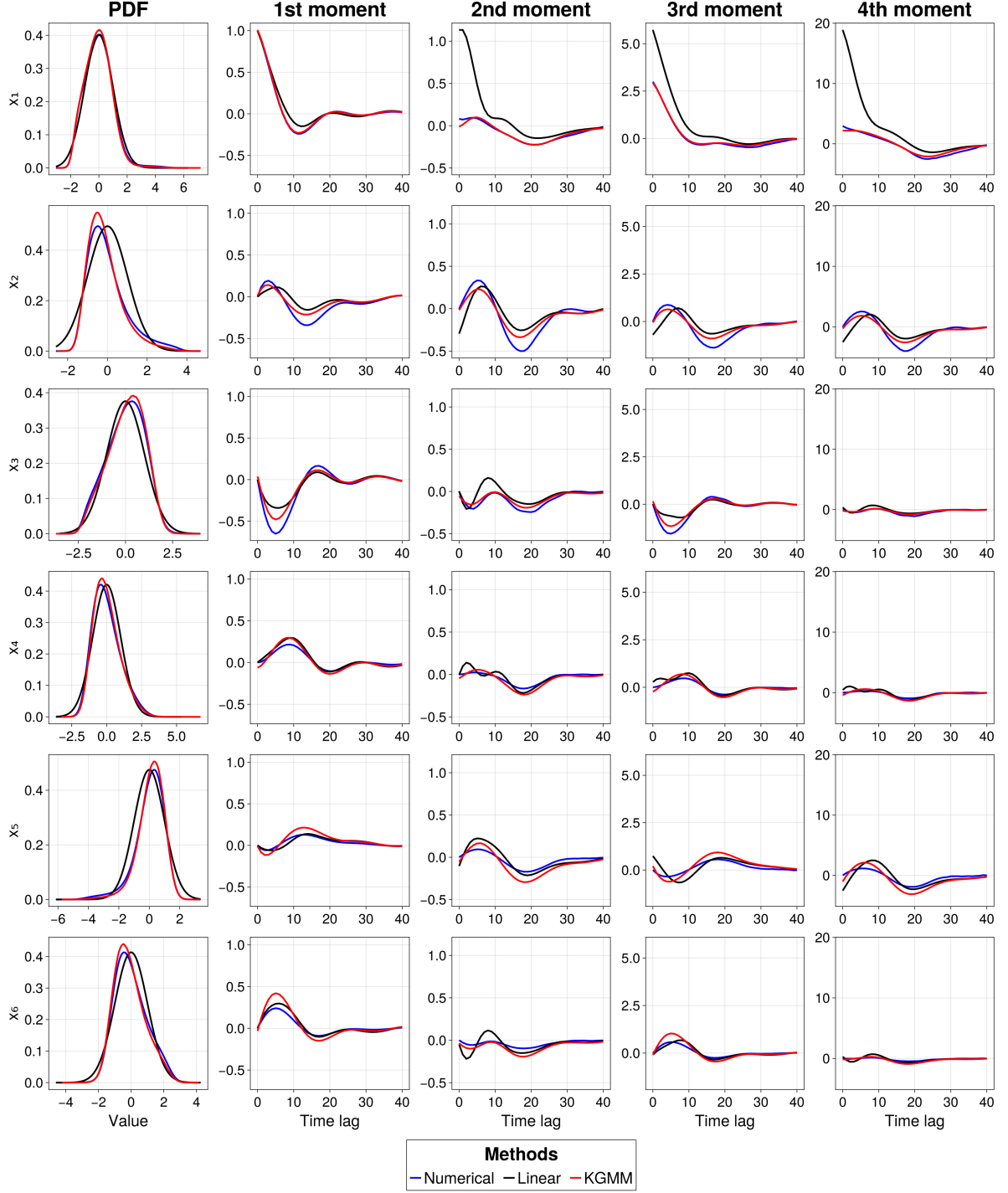
Figure 3: Comparison of the unperturbed marginal steady-state PDFs (left column) and response functions of the first four moments (remaining columns) for all six variables in the stochastic barotropic model. Results are shown for GFDT with score function estimated via KGMM (red), the linear Gaussian approximation (black), and numerical ensemble simulations (blue).

14

4 Discussion: Limitations and practical considerations

Our results show that when (i) a large amount of data is available and (ii) the full state vector $\boldsymbol{x}(t)$ is known, the system's probability density response to small external perturbations can indeed be recovered from the Generalized Fluctuation-Dissipation Theorem (GFDT), as expected from theoretical arguments [Castiglione et al., 2008]. Specifically, response functions can be estimated with high precision by leveraging recent advances in score-based generative modeling, such as the KGMM algorithm. The experiments presented here represent a valuable first step toward an equation-free framework for studying how the probability distribution of real-world dynamical systems responds to small external perturbations.

However, several limitations and caveats must be considered before applying these tools to high-dimensional complex systems. A straightforward practical limitation arises when only relatively short time series are available. As with many machine learning-based approaches, the performance of our method depends heavily on the sample size, which in turn is determined by the complexity and temporal scales of the phenomena under study. Even with access to infinitely long time series, a more fundamental issue remains in data-driven approaches: in most real-world scenarios, we rarely observe (or even know) the full state vector of the system. Instead, we typically have access to only a small subset of variables—a projection of the full dynamics. This longstanding problem, emphasized at least since Onsager and Machlup (1953) [Onsager and Machlup, 1953], remains a central concern in modern studies [Baldovin et al., 2018, 2020, Cecconi et al., 2012, Hosni and Vulpiani, 2018]. One might be tempted to address this challenge using the Takens embedding theorem [Takens, 1981], reconstructing the underlying dynamics from partial observations. However, this strategy is severely limited in practice: it is not applicable to stochastic systems and becomes rapidly infeasible as the system dimensionality increases; see Baldovin et al. [2018], Cecconi et al. [2012], Lucente et al. [2022] for an in-depth discussion. For these reasons, Takens-based approaches are not a reliable option for studying many real-world systems.

A more promising approach—aligned with the perspective taken in this work and with a long history in climate science—is to exploit the multiscale structure of real-world dynamical systems and focus on coarse-grained *effective* dynamics. In practice, in certain cases, it is possible to identify slow variables that exhibit low-dimensional effective dynamics. The cumulative effect of unresolved fast variables on the slow ones can then be modeled as stochastic forcing terms [Hasselmann, 1976, Lucarini and Chekroun, 2023, Majda et al., 2008c, Penland, 1989]. Response theory can then be applied by focusing on the coarse-grained dynamics [Lacorata and Vulpiani, 2007]. Accordingly, the success of the proposed method in real-world systems depends critically on the choice of the *proper* variables used to study the phenomena of interest; see also Appendix A in Falasca et al. [2025]. In practice, these choices are informed by physical intuition and prior knowledge, and they are shaped by the specific coarse-graining and data preprocessing procedures applied. Consequently, a purely "black box" approach is inadequate: the inference of responses functions in complex systems through data-driven methods requires deep physical insight and strong domain expertise.

5 Conclusions

In this work, we have introduced a framework for constructing higher-order response functions in nonlinear stochastic systems by combining the Generalized Fluctuation-Dissipation Theorem (GFDT) with score-based generative modeling. Our method circumvents the limitations of traditional Gaussian approximations by estimating the score function directly from data using the KGMM algorithm, allowing for accurate prediction of moment responses and perturbed probability density functions (PDFs) in non-Gaussian regimes.

We validated our approach on three prototypical systems of increasing complexity: a scalar stochastic model for low-frequency climate variability, a slow-fast triad model representative of El Niño-Southern Oscillation (ENSO) dynamics, and a six-dimensional stochastic barotropic model capturing atmospheric regime transitions. In all cases, the GFDT framework equipped with KGMM-estimated scores outperformed the linear Gaussian approximation, especially in capturing higher-order statistical responses and reconstructing the full perturbed PDF.

These results showcase the potential of incorporating data-driven score estimates in high-dimensional nonlinear systems, particularly when higher-order statistical features and extreme event predictions are of paramount importance. Future work will focus on extending this approach to larger-scale systems with partial observations and on exploring alternative generative models to further enhance the robustness and computational efficiency of the response predictions.

Appendix: Reconstruction of the Perturbed Steady-State Distribution via the Maximum Entropy Principle

In this appendix, we describe how the perturbed steady-state probability distributions can be reconstructed from the moment responses obtained through the GFDT framework. This methodology complements the main analysis presented in Section 3 by providing a way to visualize the full distributional response rather than just individual moments.

The GFDT framework allows us to estimate the perturbed steady-state probability density function $\rho_1(\boldsymbol{x})$ without directly simulating the perturbed system. Theoretically, this distribution satisfies the stationary Fokker–Planck equation associated with the perturbed dynamics:

$$\mathcal{L}_0 \rho_1(\boldsymbol{x}) + \mathcal{L}_1 \rho_S(\boldsymbol{x}) = 0, \tag{29}$$

where $\mathcal{L}_0$ is the unperturbed Fokker–Planck operator and $\mathcal{L}_1$ encodes the effect of the small perturbation, as defined in Section 2.1.

Rather than solving this partial differential equation directly, we infer the change in the distribution from its effect on a finite set of observables. In particular, we apply GFDT to observables corresponding to the first N powers of the system state $\boldsymbol{x}$, such as $\boldsymbol{x}$, $\boldsymbol{x}\boldsymbol{x}^\top$, $\boldsymbol{x}^{\otimes 3}$, and so on. This allows us to estimate the perturbation-induced variation in the first N moments of the steady-state distribution. For instance, the change in the

first two moments is given by:

$$\delta \boldsymbol{m}_1 = \mathbb{E}_1[\boldsymbol{x}] - \mathbb{E}_0[\boldsymbol{x}] = \mathbb{E}_1[\boldsymbol{x}] - \boldsymbol{\mu} = \langle \delta \boldsymbol{x} \rangle,$$

$$\delta \boldsymbol{m}_2 = \mathbb{E}_1[(\boldsymbol{x} - \mathbb{E}_1[\boldsymbol{x}])(\boldsymbol{x} - \mathbb{E}_1[\boldsymbol{x}])^\top] - \mathbb{E}_0[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^\top] \tag{30}$$

$$\approx \mathbb{E}_1[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^\top] - \mathbb{E}_0[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^\top] = \langle \delta[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^\top] \rangle.$$

These quantities provide indirect but crucial information about how the perturbed steady-state distribution $\rho_1(\boldsymbol{x})$ differs from the unperturbed one $\rho_S(\boldsymbol{x})$, even in the absence of direct samples from the perturbed dynamics.

To reconstruct the perturbed distribution $\rho_1(\boldsymbol{x})$ consistent with the estimated moments, we invoke the maximum entropy principle. Among all candidate distributions satisfying the known moment constraints, this principle selects the one with maximal Shannon entropy, ensuring the most unbiased estimate compatible with the available information [Mead and Papanicolaou, 1984].

Under this principle, the perturbed probability density function takes the exponential form:

$$\rho_1(\boldsymbol{x}) = \exp\left(\sum_{i=0}^{N} \boldsymbol{\lambda}_i \cdot \boldsymbol{\phi}_i(\boldsymbol{x}) \right), \tag{31}$$

where $\boldsymbol{\phi}_i(\boldsymbol{x})$ denotes the basis functions associated with the i-th moment (e.g., monomials, tensor products), and $\boldsymbol{\lambda}_i$ are Lagrange multipliers enforcing the moment constraints. These coefficients are obtained by solving the system:

$$\int_{\boldsymbol{\Omega}} \boldsymbol{\phi}_i(\boldsymbol{x}) \, \rho_1(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \boldsymbol{m}_i, \quad \text{for } i = 0, 1, \dots, N, \tag{32}$$

where $\boldsymbol{m}_0 = 1$ ensures normalization, and $\boldsymbol{m}_1, \dots, \boldsymbol{m}_N$ are the perturbed moments computed via GFDT. In this work, we restrict the basis functions $\boldsymbol{\phi}_i(\boldsymbol{x})$ to be multivariate polynomials, corresponding to the moment tensors of increasing order defined in Eq. (30).

This system of nonlinear integral equations generally lacks a closed-form solution and is solved numerically, for instance via iterative root-finding or variational methods. In our implementation, we first generate multiple candidate starting points based on heuristic estimates of key statistics (mean, variance, skewness, etc.) derived from the target moments. These candidates are then optimized in parallel using the Nelder–Mead method, complemented by preliminary LBFGS refinement to minimize the discrepancy between computed and desired moments. The best candidate undergoes a regularized Newton–Raphson refinement—where the Jacobian of the moment equations is computed using adaptive quadrature—to ensure rapid convergence toward the solution. Finally, a polished LBFGS optimization step is applied to further reduce the error. Once the multipliers $\boldsymbol{\lambda}_i$ are known, the resulting maximum entropy distribution provides an approximation of the perturbed steady-state PDF that incorporates the estimated changes in mean, variance, and higher-order moments.

In this study, the Maximum Entropy Principle is used primarily as a proof of concept, applied to the scalar stochastic model discussed in Section 3.1. This application allows us to validate the accuracy of our GFDT predictions by comparing the reconstructed perturbed distributions against ground truth distributions obtained analytically.

For this demonstration, we introduced a constant perturbation in the forcing term of the scalar model, shifting $F \to F + \epsilon$, and studied the response of the first four moments as

a function of the perturbation amplitude ϵ. Figure 4 shows both the moment responses and the reconstructed PDFs for different values of ϵ. The results confirm that up to $\epsilon = 0.06$, the responses predicted using both the analytic score function and the KGMM-estimated score via GFDT match the true moment variations very closely, validating that the linear-response regime holds within this range. For larger perturbations, increasing discrepancies appear, highlighting the breakdown of the linear response assumption.

The reconstructed perturbed PDFs shown in the bottom row of Figure 4 were estimated using the maximum entropy principle, based on the first two moments for the Gaussian approximation and the first four moments for the analytic and KGMM-based approaches (for $\epsilon = 0.06, 0.08$), and the first three moments for $\epsilon = 0.10, 0.12$. For these larger values of ϵ, including more moments in the maximum entropy reconstruction leads to convergence towards a multimodal distribution, which would be impossible to capture with a Gaussian approximation. Indeed, for $\epsilon = 0.12$, the algorithm used to reconstruct the perturbed PDF from the Gaussian approximation failed to converge.

While this approach worked well in the low-dimensional setting of the scalar model, its application to the higher-dimensional models presented in this paper would be computationally demanding and is beyond the scope of the current work. Nevertheless, the results from this simple case clearly demonstrate how the KGMM-based GFDT approach outperforms traditional linear approximations in predicting both the individual moments and the overall shape of perturbed probability distributions.
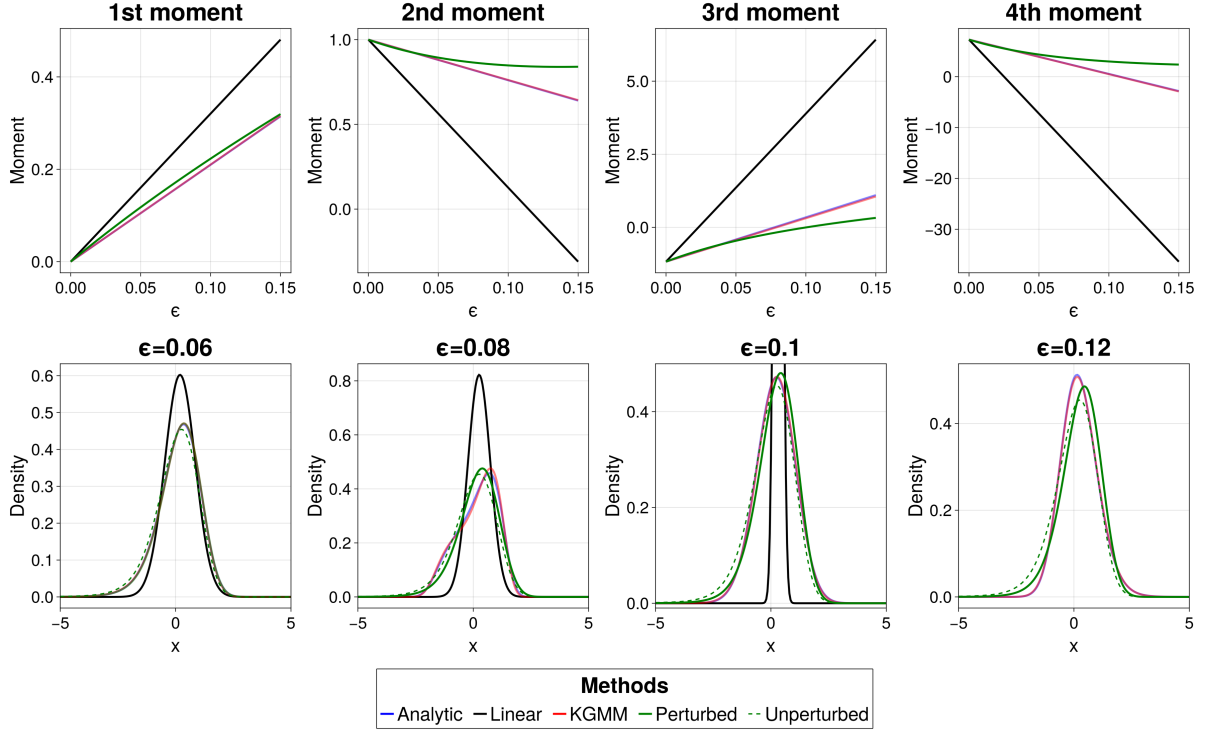
Figure 4: Top row: First four statistical moments—mean, variance, skewness, and kurtosis—as a function of the perturbation amplitude ϵ, obtained using the analytic expression of the PDF, and GFDT with the analytic (blue), linear (black) and KGMM (red) score function. Bottom row: Comparison between the true perturbed PDF (green) and the PDFs reconstructed via the maximum entropy method (see Section 2.2) and GFDT for $\epsilon = 0.06$, 0.08, 0.10, and 0.12. We compare three approaches: reconstruction using the analytic score (blue), the KGMM-estimated score (red), and a linear (Gaussian) approximation (black).

References

M. Baldovin, F. Cecconi, M. Cencini, A. Puglisi, and A. Vulpiani. The role of data in model building and prediction: A survey through examples. *Entropy*, 20, Oct 2018. doi: https://doi.org/10.3390/e20100807.

Marco Baldovin, Fabio Cecconi, and Angelo Vulpiani. Understanding causation via correlations and linear response theory. *Physical Review Research*, 2(4):043436, 2020.

G. Boffetta, G. Lacorata, S. Musacchio, and A. Vulpiani. Relaxation of finite perturbations: Beyond the fluctuation-response relation. *CHAOS*, 13(3):806–811, 2003.

P. Castiglione, M. Falcioni, A. Lesne, and A. Vulpiani. *Chaos and coarse-graining in statistical mechanics*. Cambridge University Press, 2008.

F. Cecconi, M. Cencini, M. Falcioni, and A. Vulpiani. Predicting the future from the past: An old problem from a modern perspective. *Am. J. Phys.*, 80:1001–1008, 2012. doi: https://doi.org/10.1119/1.4746070.

J.G. Charney and J.G. DeVore. Multiple flow equilibria in the atmosphere and blocking. *Journal of the atmospheric sciences*, 36:1205–1216, 1979.

Nan Chen, Xiao Hou, Qin Li, and Yingda Li. Understanding and predicting nonlinear turbulent dynamical systems with information theory. *Atmosphere*, 10(5):248, 2019.

Fenwick C Cooper and Peter H Haynes. Climate sensitivity via a nonparametric fluctuation–dissipation theorem. *Journal of the Atmospheric Sciences*, 68(5):937–953, 2011.

Daan T Crommelin, JD Opsteegh, and F Verhulst. A mechanism for atmospheric regime behavior. *Journal of the atmospheric sciences*, 61(12):1406–1419, 2004.

HE De Swart. Low-order spectral models of the atmospheric circulation: A survey. *Acta Applicandae Mathematica*, 11:49–96, 1988.

Joshua Dorrington and Tim Palmer. On the interaction of stochastic forcing and regime dynamics. *Nonlinear Processes in Geophysics*, 30(1):49–62, 2023.

F. Falasca, A. Basinski, L. Zanna, and M. Zhao. A fluctuation-dissipation theorem perspective on radiative responses to temperature perturbations. *Arxiv (Accepted, in Press in Journal of Climate)*, 2025. doi: https://doi.org/10.48550/arXiv.2408.12585.

Fabrizio Falasca, Pavel Perezhogin, and Laure Zanna. Data-driven dimensionality reduction and causal inference for spatiotemporal climate fields. *Physical Review E*, 109(4): 044202, 2024.

Boris Gershgorin and Andrew J. Majda. A test model for fluctuation-dissipation theorems with time-periodic statistics. *Physica D: Nonlinear Phenomena*, 239(17):1741–1757, 2010. ISSN 0167-2789. doi: https://doi.org/10.1016/j.physd.2010.05.009. URL `https://www.sciencedirect.com/science/article/pii/S0167278910001648`.

Michael Ghil and Valerio Lucarini. The physics of climate variability and climate change. *Reviews of Modern Physics*, 92(3):035002, 2020.

Ludovico T Giorgini, W Moon, N Chen, and J Wettlaufer. Non-gaussian stochastic dynamical model for the el niño southern oscillation. *Physical Review Research*, 4(2): L022065, 2022.

Ludovico T Giorgini, Andre N Souza, and Peter J Schmid. Reduced markovian models of dynamical systems. *Physica D: Non-linear Phenomena*, 470:134393, 2024a.

Ludovico T Giorgini, Tobias Bischoff, and Andre N Souza. Kgmm: A k-means clustering approach to gaussian mixture modeling for score function estimation. *arXiv preprint arXiv:2503.18054*, 2025a.

Ludovico T Giorgini, Andre N Souza, Domenico Lippolis, Predrag Cvitanović, and Peter Schmid. Learning dissipation and instability fields from chaotic dynamics. *arXiv preprint arXiv:2502.03456*, 2025b.

Ludovico Theo Giorgini, Katherine Deck, Tobias Bischoff, and Andre Souza. Response theory via generative score modeling. *Physical Review Letters*, 133(26):267302, 2024b.

A. Gritsun and G. Branstator. Climate response using a three-dimensional operator based on the fluctuation–dissipation theorem. *Journal of The Atmospheric Science*, page 2558–2575, 2007. doi: https://doi.org/10.1175/JAS3943.1.

K. Hasselmann. Stochastic climate models part i. theory. *Tellus*, 28:473–485, 1976. doi: https://doi.org/10.1111/j.2153-3490.1976.tb00696.x.

H. Hosni and A. Vulpiani. Forecasting in Light of Big Data. *Philos. Technol.*, 31:557–569, 2018. doi: https://doi.org/10.1007/s13347-017-0265-3.

IPCC. Climate change 2013: The physical science basis. `https://www.ipcc.ch/report/ar5/wg1/`, 2013. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

N D Keyes, L T Giorgini, and J S Wettlaufer. Stochastic paleoclimatology: Modeling the epica ice core climate records. *Chaos*, 33(9):093132, 2023.

D. Kondrashov, M.D. Chekroun, and M. Ghil. Data-driven non-markovian closure models. *Physica D: Nonlinear Phenomena*, 297:33–55, 2015. ISSN 0167-2789. doi: https://doi.org/10.1016/j.physd.2014.12.005. URL `https://www.sciencedirect.com/science/article/pii/S0167278914002413`.

S. Kravtsov, D. Kondrashov, and M. Ghil. Multilevel Regression Modeling of Nonlinear Processes: Derivation and Applications to Climatic Variability . *Journal of Climate*, 18:4404–4424, 2005.

G. Lacorata and A. Vulpiani. Fluctuation-response relation and modeling in systems with fast and slow dynamics. *Nonlin. Processes Geophys.*, 14:681–694, 2007. URL `https://doi.org/10.5194/npg-14-681-2007`.

V. Lucarini and M.D. Chekroun. Theoretical tools for understanding the climate crisis from Hasselmann's programme and beyond. *Nat Rev Phys (2023)*, 2023. URL `https://doi.org/10.1038/s42254-023-00650-8`.

D. Lucente, A. Baldassarri, A. Puglisi, A. Vulpiani, and M. Viale. Inference of time irreversibility from incomplete information: Linear systems and its pitfalls. *Phys. Rev. Res.*, 4:043103, Nov 2022. doi: 10.1103/PhysRevResearch.4.043103. URL `https://link.aps.org/doi/10.1103/PhysRevResearch.4.043103`.

A. J. Majda, I. Timofeyev, and Vanden-Eijnden. Models for stochastic climate prediction. *Proc. Natl. Acad. Sci. USA*, 96:14687–14691, 1999.

A. J. Majda, I. Timofeyev, and Vanden-Eijnden. A mathematical framework for stochastic climate models. *Proc. Natl. Acad. Sci. USA*, 54:891–974, 2001.

A. J. Majda, R. V. Abramov, and M. J. Grote. *Information Theory and Stochastics for Multiscale Nonlinear Systems*. CRM Monograph Series, American Mathematical Society, 2005.

A. J. Majda, B. Gershgorin, and Y. Yuan. Low-frequency climate response and fluctuation–dissipation theorems: Theory and practice. *Journal of the Atmospheric Sciences*, 67:1186–1201, 2010a.

A.J. Majda, R. Abramov, and B. Gershgorin. High skill in low-frequency climate response through fluctuation dissipation theorems despite structural instability. *Proc. Natl. Acad. Sci.*, 2(107):581–586, 2009a. doi: https://doi.org/10.1073/pnas.0912997107.

A.J. Majda, C. Franzke, and D. Crommelin. Normal forms for reduced stochastic climate models. *Proc. Natl. Acad. Sci.*, 10(106):3649–3653, 2009b. doi: https://doi.org/10.1073/pnas.0900173106.

Andrew J Majda and Di Qi. Linear and nonlinear statistical response theories with prototype applications to sensitivity analysis and statistical control of complex turbulent dynamical systems. *Chaos*, 29(103131), 2019. doi: https://doi.org/10.1063/1.5118690.

Andrew J Majda, Christian Franzke, and Boualem Khouider. An applied mathematics perspective on stochastic modelling for climate. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1875):2427–2453, 2008a.

Andrew J Majda, Christian Franzke, and Boualem Khouider. An applied mathematics perspective on stochastic modelling for climate. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1875):2427–2453, 2008b.

Andrew J Majda, Christian Franzke, and Boualem Khouider. An applied mathematics perspective on stochastic modelling for climate. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1875):2427–2453, 2008c.

Andrew J Majda, Rafail V Abramov, and Marcus J Grote. *Information Theory and Stochastics for Multiscale Nonlinear Systems*. American Mathematical Society, 2010b.

Lawrence R Mead and Nikos Papanicolaou. Maximum entropy in the problem of moments. *Journal of Mathematical Physics*, 25(8):2404–2417, 1984.

L. Onsager and S. Machlup. Fluctuations and irreversible processes. *Phys. Rev.*, 91:1505, 1953. URL https://journals.aps.org/pr/abstract/10.1103/PhysRev.91.1505.

T.N. Palmer. A nonlinear dynamical perspective on climate change. *Nature*, 410(6826):789–795, 2001.

C. Penland. Random Forcing and Forecasting Using Principal Oscillation Pattern Analysis. *Monthly Weather Review*, 117:2165–2185, 1989.

C. Penland and P.D. Sardeshmukh. The optimal growth of tropical sea surface temperature anomalies. *Journal of Climate*, 8(8):1999–2024, 1995.

H. Risken. *The Fokker-Planck Equation – Methods of Solution and Applications*. Springer-Verlag, 1996. doi: doi.org/10.1007/978-3-642-04898-2_150.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2021.

Andre N Souza. Representing turbulent statistics with partitions of state space. part 1. theory and methodology. *Journal of Fluid Mechanics*, 997:A1, 2024a.

Andre N Souza. Representing turbulent statistics with partitions of state space. part 2. the compressible euler equations. *Journal of Fluid Mechanics*, 997:A2, 2024b.

Andre N. Souza and Simone Silvestri. A modified bisecting k-means for approximating transfer operators: Application to the lorenz equations. *arXiv preprint arXiv:2412.03734*, 2024. doi: 10.48550/arXiv.2412.03734. URL `https://doi.org/10.48550/arXiv.2412.03734`. Submitted to arXiv on Dec 4, 2024.

Steven H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. CRC Press, 2018.

K. Strounine, S. Kravtsov, D. Kondrashov, and M. Ghil. Reduced models of atmospheric low-frequency variability: Parameter estimation and comparative performance. *Physica D: Nonlinear Phenomena*, 239(3):145–166, 2010. ISSN 0167-2789. doi: https://doi.org/10.1016/j.physd.2009.10.013. URL `https://www.sciencedirect.com/science/article/pii/S0167278909003285`.

F. Takens. *Detecting strange attractors in turbulence, in Dynamical Systems and Turbulence*, volume 898, page 21–48. Springer, Berlin,Heidelberg, 1981.

Sulian Thual, Andrew J Majda, Nan Chen, and Samuel N Stechmann. Simple stochastic model for el niño with westerly wind bursts. *Proceedings of the National Academy of Sciences*, 113(37):10245–10250, 2016.