Response Theory via Generative Score Modeling

Ludovico Theo Giorgini [1,2,*,†] Katherine Deck [3,*] Tobias Bischoff,[4,*] and Andre Souza [2,*]

[1]*Nordita, Royal Institute of Technology and Stockholm University, Stockholm 106 91, Sweden*
[2]*Massachusetts Institute of Technology, Cambridge, Massachusetts, USA*
[3]*Climate Modeling Alliance, California Institute of Technology, Pasadena, California, USA*
[4]*Independent Researcher, Pasadena, California, USA*

We introduce an approach for analyzing the responses of dynamical systems to external perturbations that combines score-based generative modeling with the generalized fluctuation-dissipation theorem. The methodology enables accurate estimation of system responses, including those with non-Gaussian statistics. We numerically validate our approach using time-series data from three different stochastic partial differential equations of increasing complexity: an Ornstein-Uhlenbeck process with spatially correlated noise, a modified stochastic Allen-Cahn equation, and the 2D Navier-Stokes equations. We demonstrate the improved accuracy of the methodology over conventional methods and discuss its potential as a versatile tool for predicting the statistical behavior of complex dynamical systems.

The study of high-dimensional dynamical systems is essential for advancing our understanding of various complex phenomena (e.g., [1–8]). These systems are characterized by numerous interacting degrees of freedom, manifesting feedback mechanisms across spatial and temporal scales. Notable examples of such complexity are observed in climate modeling (e.g., [9]), where feedback mechanisms lead to self-sustained spatiotemporal patterns, such as the El Niño Southern Oscillation in the tropical Pacific Ocean, the Asian Monsoon (particularly prominent in South and East Asia), the Indian Ocean Dipole, and the Madden-Julian Oscillation in the Indian and Pacific Oceans [10,11], among others. Similar complexities and the need to understand feedback mechanisms can be observed in other areas, such as financial markets (e.g., [12]) or neuroscience (e.g., [13,14]).

A central challenge is to characterize the causal relationships among these degrees of freedom without prior knowledge of the underlying evolution laws [15–18]. Causal inference seeks to unambiguously determine whether the behavior of one variable has been influenced by another based on observed time-series data. This estimation process is challenging due to the complexity and high dimensionality of these systems [19,20].

In recent decades, numerous methodologies have been developed to infer causal relationships directly from data (e.g., [21–25]), but calculating a system's response to small perturbations using linear response theory has become a dominant approach [26–29]. This method allows for evaluating responses without actually perturbing the system, leveraging instead the analysis of unperturbed dynamics [30,31]. It is underpinned by the generalized fluctuation-dissipation theorem (GFDT), an extension of the classical fluctuation-dissipation theorem [32]. A central obstacle in applying linear response theory is obtaining the system's unperturbed invariant distribution, i.e., the distribution of a system's attractor [33–35].

The invariant distribution is often approximated by assuming a multivariate Gaussian distribution ansatz to circumvent this issue [36]. This approximation is reasonable when examining coarse-grained (spatially or temporally) variables or particular physical observables [37,38]. The limits of this approximation are evident in atmospheric science, where observational data from wind and rainfall exhibit intermittency and other non-Gaussian features [39–42]. Another context for non-Gaussian statistics comes from neuroscience, where the brain's response to stimuli and the resulting patterns of neural activity display non-Gaussian properties [43].

Our Letter leverages recent advancements in score-based generative modeling (e.g., [44–49], and follow-up studies) to accurately represent a high-dimensional, nonlinear dynamical system's invariant distribution via a "score function," the gradient of the logarithm of a system's invariant distribution. Score models are used primarily as generative models, i.e., for sampling from the data distribution. Still, they have other applications, such as estimating the dimensionality of the data distribution; see Ref. [50]. This study demonstrates another use case beyond sample generation and in an area

---

where machine-learning approaches have not yet been fully utilized: A trained score model can be used to compute response functions using data from dynamical systems.

Using trajectories of the dynamical system, we train a model that approximates the score of the steady-state, or invariant, distribution. The model is a neural network based on spatial convolutions that can capture high-dimensional data distribution statistics (e.g., [51–53]). Like the Gaussian approximation approach, this method requires only time-series data. In this Letter, we apply the score-based method to three dynamical systems and compare its performance to other standard approaches for computing response functions. This approach outperforms the traditional Gaussian approximation when the underlying system is nonlinear.

Problem description—We focus on systems described by an evolution equation

$$\partial_t u = F(u) + \xi, \tag{1}$$

where u is the system state, which varies as a function of location in space and time and belongs to a space $\mathcal{X}$. The function $F \colon \mathcal{X} \to \mathcal{X}$ is a map, and ξ is a noise term, which can be taken to be zero to yield a deterministic system. However, we emphasize that the only requirements for applying the methodology here are time-series data, and one does not need to know an underlying evolution law such as Eq. (1), which represents systems of both ordinary or partial differential equations.

Introducing a finite-dimensional analog facilitates the connection with computations. When computations are performed on a grid, we use $u_i(t) \equiv u(x_i, t)$ as shorthand to denote the values of u at grid location x_i at time t, where $i \in [1, N^2]$ and N^2 is the number of grid points. Thinking of u_i as the components of a vector in $\mathbb{R}^{N^2}$, we also use $\vec{u}(t)$ to denote the vector of values u at time t and each grid location x_i. We use similar notation for other functions as well; e.g., if $\vec{f} \colon \mathbb{R}^{N^2} \to \mathbb{R}^{N^2}$, then $f_i[u(t)]$ denotes the value of $\vec{f}(\vec{u})$ at grid location x_i and time t.

We consider statistically stationary (steady-state) systems with a smooth invariant probability density function $\rho = \rho(u_1, u_2, \ldots, u_{N^2})$. We will use the GFDT, which is defined in Eq. (3), to understand how a small initial perturbation applied at time t alters the expected value of a system's state at a later time $t + \mathcal{T}$, compared to its unperturbed expected state. This relative change in the state after a time $\mathcal{T}$, $\delta \vec{u}(\mathcal{T})$, has the expectation

$$\langle \delta \vec{u}(\mathcal{T}) \rangle_p \equiv \langle \vec{u}(\mathcal{T}) \rangle_{\mathcal{J}'} - \langle \vec{u} \rangle_\rho, \tag{2}$$

where $\langle \cdot \rangle_\rho$ denotes an expectation over the steady-state attractor distribution ρ, $\langle \cdot \rangle_{\mathcal{J}'}$ denotes an expectation over the joint density $\mathcal{J}'[\vec{u}(t), \vec{v}(t + \mathcal{T})|\vec{v}(t) = \vec{u}(t) + \delta \vec{u}(t)]$, for a fixed perturbation $\delta \vec{u}(t)$ at time t. The ensemble average $\langle \cdot \rangle_p$ is defined by the right-hand side of Eq. (2).

Below, we will use $\langle \cdot \rangle_{\mathcal{J}}$ to denote an expectation over the joint density $\mathcal{J}[\vec{u}(t), \vec{v}(t + \mathcal{T})|\vec{v}(t) = \vec{u}(t)]$, which can be written as $\mathcal{J}[\vec{u}(t), \vec{u}(t + \mathcal{T})]$.

For a small perturbation to u at grid location j, denoted by δu_j, we wish to quantify the change in the average of u_i at a later time $\mathcal{T}$ [the discrete analog of Eq. (2)]. The GFDT states that this can be evaluated using a matrix-valued quantity called the response function R_{ij} (we also refer to it as the response matrix) [32] as

$$R_{ij}(\mathcal{T}) \equiv \frac{\langle \delta u_i(\mathcal{T}) \rangle_p}{\delta u_j(0)} = -\langle u_i(\mathcal{T}) s_j[\vec{u}(0)] \rangle_{\mathcal{J}}, \tag{3}$$

where the score function of the steady-state distribution has the usual definition

$$\vec{s} = \nabla \ln \rho. \tag{4}$$

Since we assume that we have access to time-series data, the problem of estimating the response function is thus reduced to estimating the score function $\vec{s}(\vec{u})$ of the steady-state distribution ρ. In deterministic dynamical systems, this invariant distribution is typically singular almost everywhere on the attractor, and it is necessary to introduce Gaussian noise into the system to make ρ sufficiently smooth to apply the GFDT [54] effectively.

If the steady-state distribution is a multivariate Gaussian distribution, the response matrix is given by

$$\boldsymbol{R}(\mathcal{T}) = \boldsymbol{C}(\mathcal{T}) \boldsymbol{C}^{-1}(0), \tag{5}$$

where $\boldsymbol{C}(\mathcal{T})$ is the correlation matrix, with elements $C_{ij}(\mathcal{T}) = \langle u_i(\mathcal{T}) u_j(0) \rangle_{\mathcal{J}}$. The property that $R_{ij}(0) = \delta_{ij}$, where δ_{ij} is the Kronecker delta, is a general property that holds for the discrete response matrix; see Supplemental Material [55] for more details, which include Refs. [56–62]. Constructing the correlation matrix is straightforward, allowing Eq. (5) to serve as a convenient method for estimating the response matrix when the underlying steady-state distribution is well approximated by a Gaussian distribution.

The response function can also be estimated using numerical simulations of the dynamical system. This estimation involves taking an ensemble of initial conditions, perturbing them, evolving the unperturbed and perturbed system forward in time, and then taking an ensemble average. However, this must be repeated for each perturbation of interest, which involves repeated potentially expensive simulations. It also requires knowledge of the equations of motion.

Here, we employ score-based generative modeling to derive the response function directly from data. The approach uses denoising score matching to facilitate the parametrization of the dynamical system's attractor score function directly, following methodologies suggested by

Song *et al.* [46,47]. Specifically, we use a U-Net to represent the score function [48,51], but other choices of architecture are necessary depending on the application. Our approach requires only data and does not require approximating the steady-state distribution as Gaussian or additional numerical simulations for different perturbations. Refer to Supplemental Material [55] for additional details on the architecture, training the model, obtaining the response function, and practical considerations.

We apply the proposed method to three different dynamical systems. We compare the response function computed in this way (generative) to a response computed by solving the perturbed dynamical system (dynamical), a response using an analytically derived score function (analytical) for the systems for which it is possible to derive one, and the response obtained through the Gaussian approximation (Gaussian).

Before describing the three systems under study, we note some commonalities. Each represents a stochastic partial differential equation (PDE) for a scalar variable. We take the spatial domain to be two-dimensional and periodic in each direction. The state field u maps a two-dimensional location into a scalar field, which we represent with a spatial discretization of N^2 degrees of freedom (or pixels, when representing the field as an image). In our Letter, we used $N = 32$ pixels. We use a coarse resolution so that comparing to the Gaussian approximation is directly tractable, as it requires inverting a $32^2 \times 32^2$ matrix. An advective term $A\partial_x u$, where A is a constant, was added because it introduces a preferred direction of information propagation and, therefore, a direction of causality. The systems differ in their complexity: The first system has Gaussian statistics, the second system is non-Gaussian but has an analytically tractable score function, and the last system is non-Gaussian, does not have an analytically tractable score function, and has "hidden dynamics" in the forcing.

An Ornstein-Uhlenbeck stochastic PDE—Our first system is an Ornstein-Uhlenbeck (OU) system [63] with spatially correlated noise [64–66]. The steady-state trajectories of the system exhibit Gaussian statistics, which makes the Gaussian approximation exact and, hence, serves as a baseline for testing our method.

Here, we consider a multidimensional version with an advection term and spatially correlated noise:

$$\partial_t u = -\lambda \Sigma u + \kappa \Delta u + A\partial_x u + \epsilon \Sigma^{1/2}\xi, \qquad (6)$$

where $u = u(x, y, t)$ is the unknown field which depends on space x, y and time t, $\Sigma \equiv (1 - \Delta)^{-1}$ is an inverse Helmholtz operator, $A = 2 \times 10^{-2}$, $\epsilon = 2 \times 10^{-3/2}$, $\lambda = 4 \times 10^{-2}$ are constant scalars, and ξ represents space-time noise with covariance

$$\mathbb{Cov}[\xi(x, y, t)] = \delta(x - x', y - y', t - t'), \qquad (7)$$

where $\delta(x, y, t)$ is the Dirac delta function. In this case, an analytically tractable score function is available:

$$s(u) = \frac{2}{\epsilon^2}\left(-\lambda u + \kappa \Delta u - \kappa \Delta^2 u\right), \qquad (8)$$

which does not depend on the advective term $A\partial_x u$.

A modified Allen-Cahn system—Our second system is the Allen-Cahn system, an example of a reaction-diffusion equation used for modeling phase separation in multicomponent systems (see Ref. [67], who studied applications to binary alloys). Reaction-diffusion equations [68] are partial differential equations combining a diffusive term with a (non)linear reaction term. They are commonly used to study pattern formation (e.g., [69–71]) and chemical processes.

We considered a modified version of the Allen-Cahn system to include an advection term and spatially correlated noise, as follows:

$$\partial_t u = \kappa \Delta u - A\partial_x u + \alpha \Sigma\big(u(1 - u^2)\big) + \epsilon \Sigma^{1/2}\xi, \quad (9)$$

where Σ, A, ϵ, u, and ξ are same as those used in the OU system. We choose parameters $\alpha = 16$ and $\kappa = 2.5 \times 10^{-4}$. The system has an analytically tractable score function:

$$s(u) = \frac{2}{\epsilon^2}[\alpha u(1 - u^2) + \kappa(\Delta u - \Delta^2 u)], \qquad (10)$$

which, similar to the OU example, does not depend on the advective term $A\partial_x u$. The steady-state trajectories of this system exhibit bimodal statistics and, hence, provide an example where the Gaussian approximation to the response function is expected to fail.

For both the OU and Allen-Cahn systems, the analytic score function and the machine-learned score will differ by a constant factor due to the difference between the Kronecker delta and the Dirac delta function. In both examples, the analytic score function is 32^2 larger. We divide the analytic score by this factor for comparison with the machine-learned score. Refer to Supplemental Material [55] for the derivation of the analytic score functions which contains Refs. [72,73].

Navier-Stokes equation—We next study a version of the 2D Navier-Stokes model as described in [74], where we take the state variable to be the vorticity u. This system was chosen as a stepping stone toward the more complex fluids observed in climate systems; however, the vorticity statistics are not strongly non-Gaussian due to the chosen resolution as would be expected from a coarse-grained system [37]. Hence, we expect good performance from the Gaussian approximation to the response function. The data-generating equation is

$$\partial_t u = \partial_x \psi \partial_y u - \partial_x u \partial_y \psi - A\partial_x u + \mathcal{D}u + \varsigma, \quad (11)$$
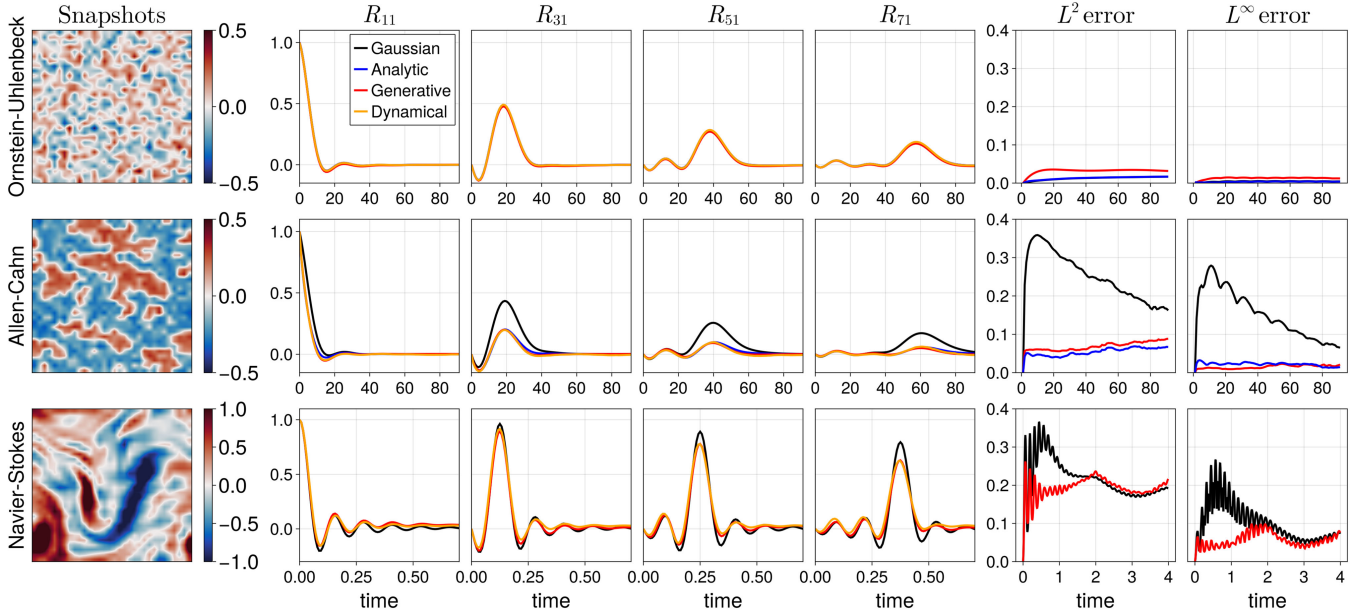
FIG. 1. Column 1: example snapshots of the three systems. In the top row, Ornstein-Uhlenbeck PDE; in the middle row, Allen-Cahn; in the bottom row, Navier-Stokes. Columns 2–5: response functions for three systems. The response to a perturbation at the pixel component $i = 1$ is evaluated at various pixel components with coordinates indicated at each panel. All evaluated components align with the advection direction. Responses are computed by integrating the dynamical system (orange lines, "dynamical"), via Gaussian approximation (black lines, "Gaussian"), using the score function derived from generative modeling (red lines, "generative"), and using the analytic score function (blue lines, "analytic"). Columns 6 and 7: response function errors for the three systems. Here, we show both the rms error and the infinity norm error of the response function as a function of time. Here, we use the dynamic response function as the "ground truth." We see that the generative response (red) generally outperforms the Gaussian response (blue) for the nonlinear cases (Allen-Cahn and Navier-Stokes), especially in the infinity norm, and performs similarly for the linear system (Ornstein-Uhlenbeck). The response function computed with the exact score does not have zero error relative to the dynamical response function because of errors introduced in numerical discretization and empirical averaging.

$$\psi = \Delta^{-1} u, \qquad (12)$$

$$\mathcal{D}u = -\nu_h \Delta^{-2} u - \nu \Delta^2 u - \nu_0 \int u \mathrm{d}x \mathrm{d}y, \qquad (13)$$

where $\mathcal{D}$ is a dissipation operator with constants [$\nu_h = 10^{-2}$, $\nu = 10^{-5}$, $\nu_0 = (2\pi)^{-2}$], $A = \pi$ is a mean advection term, and ς is a random wave forcing from [74]. The random wave forcing introduces hidden dynamics into the system; thus, the score function is estimated using only partial state information. The inverse Laplacian is defined to preserve the mean of the original variable.

Results—We first generate time-series data corresponding to Eqs. (6), (9), and (11). These data are used to compute response functions and the training sets for the score model. To create the training data, we determine autocorrelation times for each and subsample the trajectories to create statistically independent snapshots. Figure 1 column 1 shows training data samples of the numerically simulated systems. All details regarding the numerical simulations (time-stepping, spatial discretization), the datasets, and the training of the score models are provided in Supplemental Material [55], along with an in-depth

comparison between data samples generated via numerical simulation of the Allen-Cahn system and those produced using the learned score function of this system through reverse diffusion. Training the neural network took a few hours to a few days for each score model on a single Nvidia A100.

Owing to the periodic boundary conditions and symmetries of the systems under study, the response matrix elements exhibit translational invariance in the x and y directions. However, due to the advection term, the response function in the x direction depends on both the distance and the direction (sign) between pixels, while the y direction depends on only the distance. This invariance implies that the index j in the response matrix R_{ij} can be arbitrarily fixed. We chose $j = 1$ and calculated the response matrix elements along the direction of advection (denoted by i), as this is the direction information propagation, where we observe a more significant response to perturbations at pixel component $i = 1$. The resulting responses of nearby pixels as a function of time for the three systems are presented in Fig. 1, columns 2–5. The generative model is consistent with the dynamical response function (our "ground truth") in all cases. For the OU system, the Gaussian response function also matches the

ground truth; this is expected as this system exhibits Gaussian statistics. For the Allen-Cahn system, it is clear that the Gaussian approximation is a poor approximation to the response function. And lastly, for the Navier-Stokes system, we see improvement in using the generative score over the Gaussian response, despite the coarse Navier-Stokes system being quasi-Gaussian.

In Fig. 1 columns 6 and 7, we compared the error over the entire state field between the generative, Gaussian, and analytic (for the systems where it is available) response functions with the dynamical response function considered as ground truth for the three systems as a function of time. We used two different metrics to determine the errors: the L^2 and the L^∞ norm. The L^2 norm measures the average error across all pixels in the domain, providing a sense of the overall error distribution. It is helpful in understanding the general accuracy of the response function over the entire field. On the other hand, the L^∞ norm measures the maximum error at any single pixel, highlighting the worst-case error. This norm is particularly relevant to ensure the minimization of the largest deviations. From the comparison in Fig. 1, it is evident that the generative response function generally outperforms the Gaussian response in the nonlinear cases (Allen-Cahn and Navier-Stokes), especially in terms of the L^∞ norm at early to intermediate times. This observation indicates that the generative model is better at capturing the extreme responses at individual pixels. The Gaussian approximation is exact for the Ornstein-Uhlenbeck system, which has a Gaussian steady-state distribution, and both the generative and Gaussian response functions perform similarly.

Drawbacks and practical considerations—While score-based generative models offer a powerful tool for estimating response functions, there are practical challenges. One issue is that training these models requires substantial data or a sufficiently good inductive bias in the parametrized model for the score function. Additionally, setting up and training neural networks can often require significant resources, engineering, and time. Expertise in training neural networks is essential. For example, suppose the data or training time is insufficient, and one is using a U-Net to represent the score function. In that case, the generative response function tends to match the Gaussian approximation's performance due to the U-Net's linear bypass connections. A further drawback with any machine-learning method is the lack of performance "out of sample"; indeed, in Supplemental Material [55], we demonstrate that the score model for the Allen-Cahn system performs poorly when evaluated on data dissimilar from the training data.

For many applications, especially those mentioned in the introduction, several additional steps may need to be taken to apply the methods from this Letter effectively. One crucial step is the incorporation of domain-specific knowledge into the design of the neural network architecture. This will significantly improve the model's inductive bias and alleviate the data burden. Additionally, domain-specific challenges such as nonstationarity or incomplete state information must be incorporated [75]. This is particularly salient for climate data, which exhibits diurnal, seasonal, and longer timescale variations [76,77]. As always, it is crucial to validate the trained model with known benchmarks or analytical solutions.

Conclusions—In this Letter, we use score-based diffusion models to compute response functions, offering a purely data-driven method to calculate accurate response functions. We validated this approach on a system with Gaussian statistics (Ornstein-Uhlenbeck) and two dynamical systems with more complex distributions: a reaction-diffusion equation (Allen-Cahn) and the Navier-Stokes equations. The former is a class of SPDEs that have applications across biology, chemistry, and physics, and the latter is a system exhibiting some of the characteristics of more complex fluids like Earth's atmosphere or ocean. We compared the response function of these systems computed in multiple ways: using an *analytic* expression for the score, using the *generative* score function, using a *dynamical* approach by introducing perturbations to initial conditions and solving the equations of motion, and using a *Gaussian* approximation. For the nonlinear cases, we found that the Gaussian response function overpredicted the response to neighboring perturbations and that the generative response function agreed well with both the dynamical and analytical response functions. Compared with the Gaussian response, the RMSE of the generative response function had significantly reduced error. This suggests that score-based models are a useful tool for studying system response functions, especially when the dynamics are strongly nonlinear.

Data availability—The code utilized in this work is publicly available for further study and replication [78]. We used the publicly available code base [79] as the backbone for training the models.

[1] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, Science **328,** 876 (2010).

[2] A. Halu, R. J. Mondragón, P. Panzarasa, and G. Bianconi, PLoS One **8,** e78293 (2013).

[3] J. Jumper *et al.*, Nature (London) **596,** 583 (2021).

[4] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, Annu. Rev. Fluid Mech. **52,** 477 (2019).

[5] J. Barbier, arXiv:2010.14863.

[6] M. F. Singh, C. Wang, M. W. Cole, and S. Ching, arXiv:2104.02827.

[7] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Rev. Mod. Phys. **80,** 1275 (2008).

[8] A.-L. Barabási and R. Albert, Science **286,** 509 (1999).

[9] M. Ghil and V. Lucarini, Rev. Mod. Phys. **92,** 035002 (2020).

[10] A. Timmermann, S.-I. An, J.-S. Kug, F.-F. Jin, W. Cai, A. Capotondi, K. M. Cobb, M. Lengaigne, M. J. McPhaden, M. F. Stuecker *et al.*, Nature (London) **559,** 535 (2018).

[11] Z. Martin, S.-W. Son, A. Butler, H. Hendon, H. Kim, A. Sobel, S. Yoden, and C. Zhang, Nat. Rev. Earth Environ. **2,** 477 (2021).

[12] N. Badwan, Macroeconomic Analysis for Economic Growth, 145 (2022).

[13] B. A. Hunt, P. K. Tewarie, A. G. G. Smith, C. Porcaro, K. D. Singh, and P. R. Murphy, PLoS Biol., **20,** e3001686 (2021).

[14] O. Sporns and R. F. Betzel, Network Neuroscience **5,** 405 (2021).

[15] A. E. BozorgMagham, S. Motesharrei, S. G. Penny, and E. Kalnay, PLoS One **10,** e0131226 (2015).

[16] K. Lagemann, C. Lagemann, B. Taschler, and S. Mukherjee, Nat. Mach. Intell. **5,** 1306 (2023).

[17] A. Wismüller, A. M. Dsouza, M. A. Vosoughi, and A. Abidin, Sci. Rep. **11,** 7817 (2021).

[18] N. D. B. Keyes, L. T. Giorgini, and J. S. Wettlaufer, Chaos **33,** 093132 (2023).

[19] E. Aurell and G. Del Ferraro, J. Phys. Conf. Ser. **699,** 012002 (2016).

[20] R. Friedrich, J. Peinke, M. Sahimi, and M. R. R. Tabar, Phys. Rep. **506,** 87 (2011).

[21] C. W. J. Granger, Econometrica **39,** 199 (1969).

[22] T. Schreiber, Phys. Rev. Lett. **85,** 461 (2000).

[23] J. Pearl, Stat. Surv. **3,** 96 (2009).

[24] G. Camps-Valls, A. Gerhardus, U. Ninad, G. Varando, G. Martius, E. Balaguer-Ballester, R. Vinuesa, E. Diaz, L. Zanna, and J. Runge, arXiv:2305.13341.

[25] J. Kaddour, A. Lynch, Q. Liu, M. J. Kusner, and R. Silva, arXiv:2206.15475.

[26] M. Baldovin, F. Cecconi, A. Provenzale, and A. Vulpiani, Sci. Rep. **12,** 15320 (2022).

[27] F. Cecconi, G. Costantini, C. Guardiani, M. Baldovin, and A. Vulpiani, Phys. Biol. **20,** 056002 (2023).

[28] F. Falasca, P. Perezhogin, and L. Zanna, Phys. Rev. E **109,** 044202 (2024).

[29] F. Falasca, A. Basinski-Ferris, L. Zanna, and M. Zhao, arXiv:2408.12585v2.

[30] M. Baldovin, F. Cecconi, and A. Vulpiani, Phys. Rev. Res. **2,** 043436 (2020).

[31] V. Lucarini, J. Stat. Phys. **173,** 1698 (2018).

[32] U. M. B. Marconi, A. Puglisi, L. Rondoni, and A. Vulpiani, Phys. Rep. **461,** 111 (2008).

[33] F. Daum, IEEE Aerospace and Electronics Systems Magazine **20,** 57 (2005).

[34] P. Sjöberg, P. Lötstedt, and J. Elf, Comput. Visualization Sci. **12,** 37 (2009).

[35] F. C. Cooper and P. H. Haynes, J. Atmos. Sci. **68,** 937 (2011).

[36] C. E. Leith, J. Atmos. Sci. **32,** 2022 (1975).

[37] P. D. Sardeshmukh and P. Sura, J. Clim. **22,** 1193 (2009).

[38] A. Gritsun and V. Lucarini, Physica (Amsterdam) **349D,** 62 (2017).

[39] C. Proistosescu, A. Rhines, and P. Huybers, Geophys. Res. Lett. **43,** 5425 (2016).

[40] P. C. Loikith and J. D. Neelin, Geophys. Res. Lett. **42,** 8577 (2015).

[41] M. Branicki and A. J. Majda, Nonlinearity **25,** 2543 (2012).

[42] C. Martinez-Villalobos and J. D. Neelin, J. Atmos. Sci. **76,** 3611 (2019).

[43] N. Kriegeskorte and X.-X. Wei, Nat. Rev. Neurosci. **22,** 703 (2021).

[44] P. Vincent, Neural Comput. **23,** 1661 (2011).

[45] J. Ho, A. Jain, and P. Abbeel, Adv. Neural Inf. Process. Syst. **33,** 6840 (2020), https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html.

[46] Y. Song and S. Ermon, Adv. Neural Inf. Process. Syst. **32** (2019), https://proceedings.neurips.cc/paper/2019/hash/3001ef257407d5a371a96dcd947c7d93-Abstract.html?ref=https://githubhelp.com.

[47] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, arXiv:2011.13456.

[48] T. Bischoff and K. Deck, Artif. Intell. Earth Syst. **3,** e230039 (2024).

[49] Z. Liu, D. Luo, Y. Xu, T. Jaakkola, and M. Tegmark, arXiv:2304.02637.

[50] J. Stanczuk, G. Batzolis, T. Deveney, and C.-B. Schönlieb, arXiv:2212.12611.

[51] O. Ronneberger, P. Fischer, and T. Brox, in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 2015, Proceedings, Part III 18* (Springer, New York, 2015), pp. 234–241.

[52] B. Adcock, S. Brugiapaglia, N. Dexter, and S. Moraga, arXiv:2012.06081.

[53] P. Beneventano, P. Cheridito, R. Graeber, A. Jentzen, and B. Kuckuck, arXiv:2112.14523.

[54] A. Gritsun and G. Branstator, J. Atmos. Sci. **64,** 2558 (2007).

[55] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.133.267302 for provide detailed theoretical derivations and computational methods to support the main findings. They include the derivation of the Generalized Fluctuation-Dissipation Theorem, analysis of score functions in stochastic systems, and practical considerations for numerical implementations. Additionally, they describe the use of generative modeling techniques, such as denoising score matching, and provide numerical methods and validation results for systems like Navier-Stokes and Allen-Cahn. Visualizations of response functions, errors, and synthetic data further illustrate the robustness of the approach.

[56] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli, in *Proceedings of the International Conference on Machine Learning* (PMLR, 2015), pp. 2256–2265.

[57] Y. Song and S. Ermon, Adv. Neural Inf. Process. Syst. **33**, 12438 (2020), https://proceedings.neurips.cc/paper/2020/hash/92c3b916311a5517d9290576e3ea37ad-Abstract.html.

[58] B. D. Anderson, Stochastic Processes Appl. **12**, 313 (1982).

[59] J. Choi, J. Lee, C. Shin, S. Kim, H. Kim, and S. Yoon, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, New York, 2022), pp. 11472–11481.

[60] M. Tancik, P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. Barron, and R. Ng, Adv. Neural Inf. Process. Syst. **33**, 7537 (2020), https://proceedings.neurips.cc/paper_files/paper/2020/hash/55053683268957697aa39fba6f231c68-Abstract.html.

[61] D. P. Kingma and J. Ba, arXiv:1412.6980.

[62] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, 2nd ed., Dover Books on Mathematics (Dover Publications, Mineola, NY, 2001).

[63] G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36**, 823 (1930).

[64] R. A. Maller, G. Müller, and A. Szimayer, *Handbook of Financial Time Series* (Springer, 2009), p. 421.

[65] T. D. Frank, A. Daffertshofer, and P. J. Beek, Phys. Rev. E **63**, 011905 (2000).

[66] P. Vatiwutipong and N. Phewchean, Adv. Differ. Equat. **2019**, 1 (2019).

[67] S. M. Allen and J. W. Cahn, Acta Metall. **20**, 423 (1972).

[68] A. N. Kolmogorov, I. G. Petrovskii, and N. S. Piskunov, Bulletin. Moscow Univ. Mech. Bull. **1**, 1 (1937).

[69] A. M. Turing, Bull. Math. Biol. **52**, 153 (1990).

[70] T. Callahan and E. Knobloch, Physica (Amsterdam) **132D**, 339 (1999).

[71] S. Kondo and T. Miura, Science **329**, 1616 (2010).

[72] A. N. Souza and M. Tao, Eur. J. Appl. Math. **30**, 830 (2019).

[73] I. Corwin and H. Shen, Bull. Am. Math. Soc. **57**, 409 (2020).

[74] G. R. Flierl and A. N. Souza, J. Fluid Mech. **986**, A8 (2024).

[75] A. J. Majda, B. Gershgorin, and Y. Yuan, J. Atmos. Sci. **67**, 1186 (2010).

[76] B. Gershgorin and A. J. Majda, Physica (Amsterdam) **239D**, 1741 (2010).

[77] A. J. Majda and D. Qi, Chaos **29**, 103131 (2019).

[78] https://anonymous.4open.science/r/PrivateGenLRT-052E/generate/allen_cahn_model.jl

[79] T. Bischoff and K. Deck https://github.com/CliMA/CliMAgen.jl/.