Response Theory via Generative Score Modeling

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

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

(Dated: June 23, 2024)

We introduce an approach for analyzing the responses of dynamical systems to external perturbations that combines score-based generative modeling with the Fluctuation-Dissipation Theorem (FDT). The methodology enables accurate estimation of system responses, notably for systems with non-Gaussian statistics, which can present challenges for conventional approximate methods. We numerically validate our approach using time-series data from a stochastic partial differential equation where the score and response functions are available analytically. Furthermore, we demonstrate the improved accuracy of our methodology over conventional methods and its potential as a versatile tool for understanding complex dynamical systems. Applications span disciplines from climate science and finance to neuroscience.

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 spatio-temporal patterns, such as the El Niño Southern Oscillation (ENSO) 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 is a challenging task due to the inherent 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., 21–26], but calculating a system's response to small perturbations using linear response theory has become a dominant approach. This method, underpinned by the Fluctuation-Dissipation Theorem (FDT), allows for the evaluation of responses without actually perturbing the system, leveraging instead the analysis of unperturbed dynamics [27, 28]. A central obstacle in applying linear response theory is obtaining the system's unperturbed invariant distribution, i.e., the distribution of the system state on the attractor, which is particularly challenging in high-dimensional systems [29, 30].

To circumvent this issue, the invariant distribution is often approximated using a multivariate Gaussian distribution, which introduces potential biases. This is particularly evident in climate science, where observational data such as temperature, wind, and rainfall exhibit characteristics of intermittency and other non-Gaussian features [31–33]. Another context for non-Gaussian statistics comes from neuroscience, where the brain's response to stimuli and the resulting patterns of neural activity frequently display non-Gaussian properties [34].

Our work leverages recent advancements in score-based generative modeling [e.g., 35–38, and follow-on 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, but they have other applications, such as estimating the dimensionality of the data distribution, see [39]. 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 for time-series 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., Adcock *et al.* [40], Beneventano *et al.* [41]). We find that this approach outperforms the traditional Gaussian approximation. At the same time, it avoids the significant computational cost of direct numerical simulation of the response function, as the same score model can be used repeatedly for any perturbation of interest.

Problem description. – In this Letter, 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 $F : \mathcal{X} \to \mathcal{X}$ 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 Equation (1). The equation 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 $u_i(t) \equiv u(\vec{x}_i, t)$ as shorthand to denote the values of u and grid location $\vec{x}_i$ at time t. Thinking of u_i as the components of a vector in $\mathbb{R}^N$, we also use $\vec{u}(t)$ to denote the vector of values u at time t and each grid location $\vec{x}_i$. We use similar notation for other functions as well, e.g. if $\vec{f} : \mathbb{R}^N \to \mathbb{R}^N$, then $f_i(u(t))$ denotes the value of $\vec{f}(\vec{u})$ at grid location x_i and time t.

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

$$\langle \delta\vec{u}(\tau) \rangle_p \equiv \langle \vec{u}(\tau) \rangle_{\mathcal{J}'} - \langle \vec{u} \rangle_\rho, \tag{2}$$

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 $\mathcal{J}'(\vec{u}(t), \vec{v}(t+\tau)|\vec{v}(t) = \vec{u}(t)+\delta\vec{u}(t))$, 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 Equation (2). Below, we will use $\langle\cdot\rangle_{\mathcal{J}}$ to denote an expectation over the joint density $\mathcal{J}(\vec{u}(t), \vec{v}(t+\tau)|\vec{v}(x,t) = \vec{u}(t))$.

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 τ (the discrete analog of Equation (2). The FDT 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) [42] as

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

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 effectively apply the FDT [43].

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

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

where $\boldsymbol{C}(\tau)$ is the correlation matrix, with elements $C_{ij}(\tau) = \langle u_i(\tau) u_j(0) \rangle_{\mathcal{J}}$. The property that $C_{ij}(0) = \delta_{ij}$, where δ_{ij} is the Kronecker delta, is a general property that holds for the discrete response matrix, e.g., $R_{ij}(0) = \delta_{ij}$. Constructing the correlation matrix from available datasets is relatively straightforward, allowing Eq. (5) to serve as a convenient method for estimating the response matrix; however, this approach can lead to inaccuracies when the underlying steady-state distribution is not Gaussian.

The response function can also be estimated using numerical simulations of the dynamical system. This involves taking an ensemble of initial conditions, perturbing them, and evolving the unperturbed and perturbed system forward in time. The response can then be estimated using Equation (2) and Equation (3). However, this must be repeated for each perturbation of interest and hence involves repeated simulation of the system at a possibly large computational cost.

Here, we employ score-based generative modeling, a technique from the machine-learning community, to derive the response function directly from data. Our approach uses denoising score matching and facilitates the parameterization of the dynamical system's attractor score function directly, following methodologies suggested by Song et al. [37, 38]. This approach does not require approximating the steady-state distribution as Gaussian or require additional numerical simulations for different perturbations.

A modified Allen-Cahn system. – We apply the proposed method to estimate the response function of the Allen-Cahn system, which is an example of a reaction-diffusion equation used for modeling phase separation in multi-component systems [see 44, who studied applications to binary alloys].

Reaction-diffusion equations [45] are partial differential equations combining a diffusive term with a (non)linear reaction term. They are commonly used to study pattern formation (e.g., Turing [46], Callahan and Knobloch [47], Kondo and Miura [48]) and chemical processes. We compare the results to a response computed by solving the perturbed dynamical system, a response using an analytically derived score function, and the response obtained through the Gaussian approximation frequently used in the literature. Both the response function determined from dynamics and that based on the analytic estimate of the score can be interpreted as "truth", but note that in many cases, neither of these baseline "truths" will be attainable due to analytic intractability, the significant computational cost require to compute the response function numerically, or inability to attain the underlying system that generated the time-series.

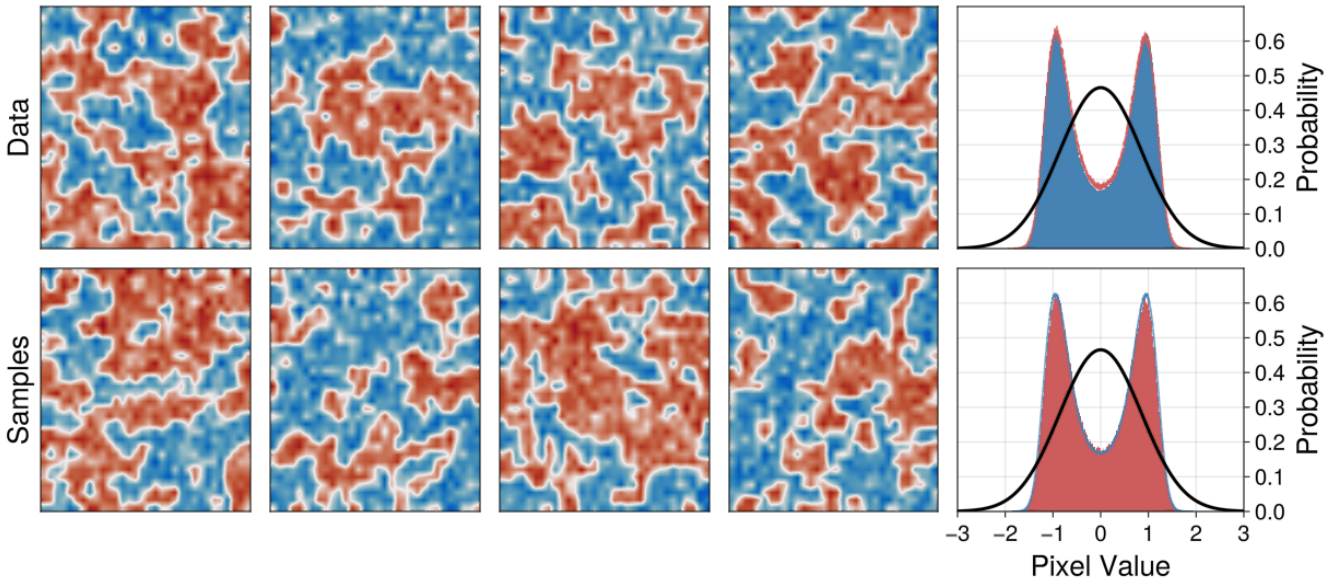

FIG. 1. **Example snapshots and steady-state pixel distribution of the solution to the Allen-Cahn equation**. In the **top row**, we show the results of a numerical simulation of Equation (8) on a domain of size $2\pi \times 2\pi$ using 32×32 pixels; in the **bottom row**, we show samples generated via the learned score function using reverse diffusion [38]. The rightmost column shows the corresponding pixel distributions, where the blue and red lines indicate the equivalent distributions from the top and bottom, respectively. The black line shows the Gaussian approximation fitted to the bimodal pixel distribution.

The Allen-Cahn system is defined by the following stochastic partial differential equation

$$\partial_t u = \Delta u + u(1-u^2) + \epsilon\xi, \tag{6}$$

where $u = u(x, y, t)$ is the unknown field which depends on space x, y and time t. Furthermore, ϵ is a constant scalar setting the amplitude of the stochastic forcing, and ξ represents space-time noise with covariance

$$\mathbb{C}\mathrm{ov}\left[\xi(x,y,t)\right] = \delta(x-x', y-y', t-t'), \tag{7}$$

where $\delta(x, y, t)$ is the Dirac delta function. Here, we adjust the system to include an advection term and spatially correlated noise, as follows,

$$\partial_t u = \kappa\Delta u - A\partial_x u + \alpha\Sigma\left(u(1-u^2)\right) + \epsilon\Sigma^{1/2}\xi, \tag{8}$$

where $\Sigma \equiv (1-\Delta)^{-1}$ is an inverse Helmholtz operator, and A, α, κ are constants. We take the spatial domain to be periodic. In this work, the 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 study, we used $N = 32$ pixels and we choose parameters $\alpha = 8{\times}10^{3/2}$, $\epsilon = 2{\times}10^{-3/2}$, $\kappa = 2.5{\times}10^{-4}$, $A = 2{\times}10^{-2}$. The domain is 2π periodic in each direction.

The advection term, $A\partial_x u$, was added because it introduces a preferred direction of information propagation and, therefore, a direction of causality. This equation generates steady-state trajectories exhibiting non-Gaussian and bimodal statistics. The bimodal nature arises because the potential functional of the system,

$$\mathcal{V}[u] = \iint \left[\frac{\alpha}{2\epsilon^2}(1-u^2)^2 + \frac{\kappa}{\epsilon^2}\left(|\nabla u|^2 + |\Delta u|^2\right)\right] \mathrm{d}x\,\mathrm{d}y, \tag{9}$$

has two global minima at $u = \pm 1$. The correlated noise is necessary for spatially local transitions between ± 1. The system has an analytically tractable score function,

$$s(u) = -\frac{\delta\mathcal{V}}{\delta u} = \frac{2}{\epsilon^2}\left[\alpha u(1-u^2) + \kappa(\Delta u - \Delta^2 u)\right], \tag{10}$$

and the advective term $A\partial_x u$ does not change the invariant distribution.

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 functions. In our example, the analytic score function is 32^2 larger. We divide the analytic score by this factor for comparison with the machine-learned score. The availability of the analytic score makes the system a good candidate for exploring the comparative performance between different response function estimates (analytic, numerically estimated with the dynamical system, Gaussian-approximated, and machine-learned).

To train the score model, we generate time series data corresponding to Equation (8), consisting of 128 ensemble members and 2,000 temporally coherent data separated by one time unit within each ensemble member. We then subsample this data by a factor of 10 for each ensemble

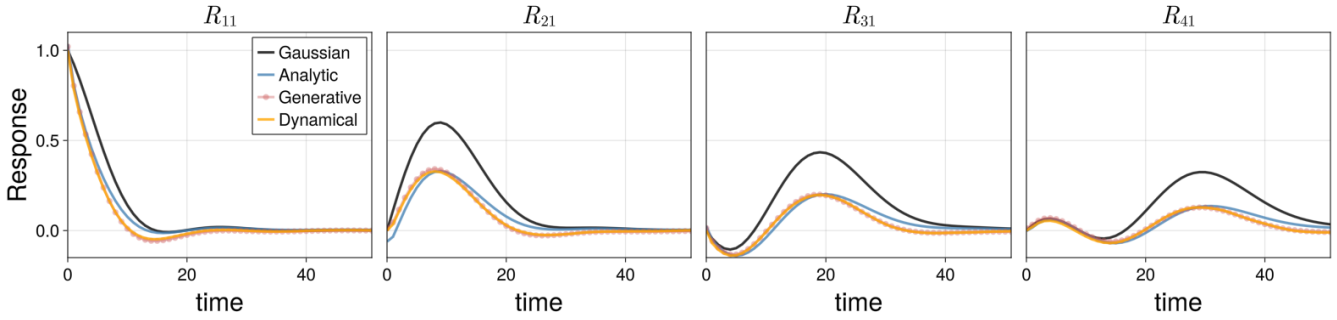

FIG. 2. **Response functions for the Allen-Cahn system**. The response to a perturbation at the pixel component $i = 1$ is evaluated at various pixel components with coordinates indicated at each panel. All evaluated components align with the advection direction. Responses are computed by integrating the dynamical system (orange lines, "Dynamical") via Gaussian approximation (black lines, "Gaussian"), using the score function derived from generative modeling (red lines and dots, "Generative"), and using the analytic score function (blue lines, "Analytic"). Statistical uncertainty in the curves is much smaller than the difference between curves and is not shown.

member to generate 25,600 states for training the score model. For the spatial discretization, we use a spectral method [49]. For the temporal discretization, we used a Runge-Kutta 4 scheme for the deterministic dynamics, with a timestep of size $\Delta t = 1/32$. We accounted for the noise term after the final Runge-Kutta stage.

Example training data samples and the steady-state pixel distribution of the numerically simulated system are shown in the upper row of Figure (1). The machine-learned method is used to generate new samples in the bottom row of Figure (1) and the corresponding steady-state pixel distribution. The machine-learned score function can generate samples similar to the original distribution. In comparing the steady-state pixel distributions of both the generated fields and data samples to the Gaussian distribution defined by the data's mean and variance (as used in the Gaussian approximation), a notable limitation emerges, which is visible in the rightmost panels of Figure (1).

We employed a time series of 2000 points (spaced in time by $t = 1$ time units) and 128 ensemble members for the computation of response functions using both the Gaussian approximation and generative modeling techniques. Owing to the periodic boundary conditions, the response matrix elements exhibit translational invariance: it depends only on the distance between pixels. This invariance implies that the index i 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 larger response to perturbations at pixel component $i = 1$.

The resulting responses of nearby pixels as a function of time are presented in Figure (2). For this system, the distance between adjacent pixels divided by the advection velocity $\mathcal{A}$ gives a travel time of approximately 10 time units, which is clearly observed and captured by all response functions. However, the figure illustrates that the response matrix elements computed using the score function from generative modeling more accurately reproduce the true response (either numerically computed or analytically derived) than those obtained via the Gaussian approximation for the full nonlinear system. To quantify the performance, we report root-mean-square-errors (RMSE) between the exact analytic estimate and the generative score model estimate and between the analytic estimate and the Gaussian estimate in Table I.

RMSE Table

Method	R_{11}	R_{21}	R_{31}	R_{41}
Generative	0.05	0.05	0.03	0.06
Analytic	0.21	0.23	0.19	0.15
Gaussian	0.62	0.99	0.93	0.80

TABLE I. **Root-mean-squared-errors (RMSE)** of the response functions, obtained using the generative score-model, analytic score-model, and the Gaussian approximation, respectively and computed using the dynamical response function as baseline over the first 50 time units. See Figure (2). The columns denote the response of a pixel (0,1,2,3) units away from a perturbed pixel. The rows indicate the three methods.

Since we have access to the analytic score function, we have an alternative method of assessing the generative model's fidelity. We evaluate Equation (10) (scaled by $1/32^2$) and the generative score for several choices of fields u in Figure (3). We test both on data similar to what the neural network has seen before (but not in the training set) and fields far from the original attractor distribution. When applying the score function to a constant field (left panel) for a choice of amplitude (such as $u = 0, u \pm 1, u = 0.5$, etc.), we obtain a cubic polynomial for the analytic score and a similar shape for the generative score, albeit with a significantly smaller amplitude. When evaluated on a smooth field (top three panels on the right), the generative score predicts the "rings". Still, it

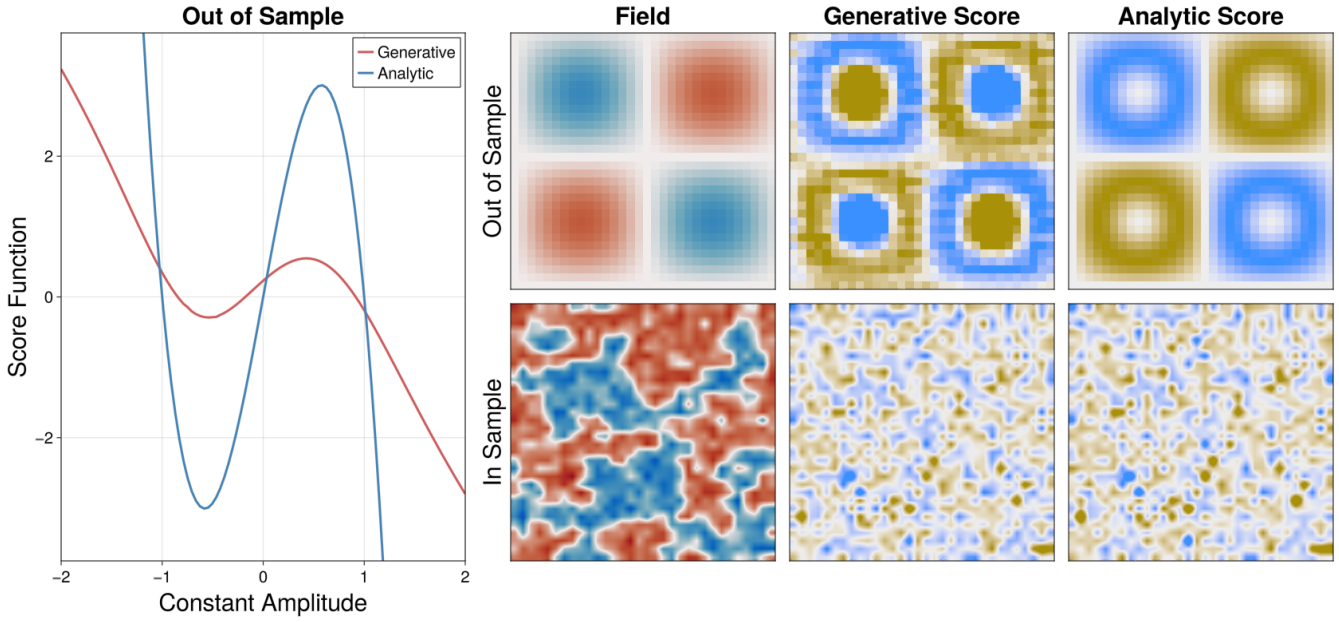

FIG. 3. **Score function evaluation.** Here, we evaluate the analytic score function and the generative score function for several cases. In the leftmost plot, we show the evaluation of a constant function with amplitude given by the horizontal axis. On the right, we show the evaluation of the score function in two separate cases: The first is on the function $\sin(x)\sin(y)$, and the second is a data sample in the test set. Color ranges are the same for fields in which comparisons can be made.

cannot obtain the center and mistakenly assigns different values to that region. And lastly, when evaluated on data similar to the steady-state distribution of the Allen-Cahn equation (bottom three panels on the right), we see excellent agreement between the analytic and predictive score. We reemphasize that the neural network was only exposed to images of the steady-state distribution. In summary, the neural network reproduces the score function when evaluated on data similar to what it has been exposed to. However, for significantly out-of-sample functions, the neural network produces quantitatively incorrect answers with similar qualitative features.

Conclusions. – In this study, 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 version of the stochastic Allen-Cahn equation, an example of a reaction-diffusion equation, a class of SPDEs that have applications across biology, chemistry, and physics. We compared the response function of the system computed in four ways: using an analytic expression for the score, using the learned score function, using perturbations to initial conditions and solving the equations of motion, and using a Gaussian approximation. In this case, we found that the Gaussian response function over-predicted the response of pixels to neighboring perturbations and that the score-based response function agreed well with both the numerically determined and analytic response functions. Compared with the Gaussian response, the RMSE of the score-model-based response function reduced error by an order of magnitude. Furthermore, we found that "in-sample" functions yielded similar score function evaluations between the analytic and generative score.

The codes utilized in this research are publicly available for further study and replication in the following GitHub repository: https://anonymous.4open.science/r/PrivateGenLRT-052E/generate/allen_cahn_model.jl.

LTG gratefully acknowledges support from the Swedish Research Council (Vetenskapsradet) Grant No. 638-2013-9243. KD acknowledges support by Eric and Wendy Schmidt (by recommendations of the Schmidt Futures) and by the Cisco Foundation. AS acknowledges support by Schmidt Sciences through the Bringing Computation to the Climate Challenge, an MIT Climate Grand Challenge Project and conversations with Fabrizio Falasca on linear response theory. TB acknowledges the support of the community at Livery Studio and stimulating conversations with Bryan Riel on generative modeling.

* All the authors contributed equally
† ludovico.giorgini@su.se

[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 **596**, 583 (2021).
[4] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, Annual Review of Fluid Mechanics **52**, 477 (2019).
[5] J. Barbier, arXiv preprint arXiv:2010.14863 (2020).
[6] M. F. Singh, C. Wang, M. W. Cole, and S. Ching, arXiv preprint arXiv:2104.02827 (2021).
[7] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Reviews of Modern Physics **80**, 1275 (2008).
[8] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[9] M. Ghil and V. Lucarini, Reviews of Modern Physics **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 **559**, 535 (2018).
[11] Z. Martin, S.-W. Son, A. Butler, H. Hendon, H. Kim, A. Sobel, S. Yoden, and C. Zhang, Nature Reviews Earth & Environment **2**, 477 (2021).
[12] N. Badwan, Perspective chapter: International financial markets and financial capital flows - forms, factors and assessment tools, in *Macroeconomic Analysis for Economic Growth*, edited by M. J. Ibrahim (IntechOpen, 2022).
[13] B. A. Hunt, P. K. Tewarie, A. G. G. Smith, C. Porcaro, K. D. Singh, and P. R. Murphy, PLOS Biology 10.1371/journal.pbio.3001686 (2021).
[14] O. Sporns and R. F. Betzel, Nature Neuroscience 10.1038/s41593-021-00858-0 (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, Deep learning of causal structures in high dimensions under data limitations (2023).
[17] A. Wismüller, A. M. Dsouza, M. A. Vosoughi, and A. Abidin, Scientific reports **11**, 7817 (2021).
[18] N. D. B. Keyes, L. T. Giorgini, and J. S. Wettlaufer, Chaos: An Interdisciplinary Journal of Nonlinear Science **33**, 093132 (2023).
[19] E. Aurell and G. Del Ferraro, in *Journal of Physics: Conference Series*, Vol. 699 (IOP Publishing, 2016) p. 012002.
[20] R. Friedrich, J. Peinke, M. Sahimi, and M. R. R. Tabar, Physics Reports **506**, 87 (2011).
[21] C. W. Granger, Econometrica: journal of the Econometric Society , 424 (1969).
[22] T. Schreiber, Physical review letters **85**, 461 (2000).
[23] J. Pearl, Statistics Surveys **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 preprint arXiv:2305.13341 (2023).
[25] J. Kaddour, A. Lynch, Q. Liu, M. J. Kusner, and R. Silva, arXiv preprint arXiv:2206.15475 (2022).
[26] F. Falasca, P. Perezhogin, and L. Zanna, Phys. Rev. E **109**, 044202 (2024).
[27] M. Baldovin, F. Cecconi, and A. Vulpiani, Physical Review Research **2**, 043436 (2020).
[28] V. Lucarini, Journal of Statistical Physics **173**, 1698 (2018).
[29] F. Daum, IEEE Aerospace and Electronic Systems Magazine **20**, 57 (2005).
[30] P. Sjöberg, P. Lötstedt, and J. Elf, Computing and Visualization in Science **12**, 37 (2009).
[31] C. Proistosescu, A. Rhines, and P. Huybers, Geophysical Research Letters **43**, 5425 (2016).
[32] P. C. Loikith and J. D. Neelin, Geophysical Research Letters **42**, 8577 (2015).
[33] M. Branicki and A. J. Majda, Nonlinearity **25**, 2543 (2012).
[34] N. Kriegeskorte and X.-X. Wei, Nature Reviews Neuroscience **22**, 703 (2021).
[35] P. Vincent, Neural computation **23**, 1661 (2011).
[36] J. Ho, A. Jain, and P. Abbeel, Advances in neural information processing systems **33**, 6840 (2020).
[37] Y. Song and S. Ermon, Advances in neural information processing systems **32** (2019).
[38] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, arXiv preprint arXiv:2011.13456 (2020).
[39] J. Stanczuk, G. Batzolis, T. Deveney, and C.-B. Schönlieb, Your diffusion model secretly knows the dimension of the data manifold (2023), arXiv:2212.12611 [cs.LG].
[40] B. Adcock, S. Brugiapaglia, N. Dexter, and S. Moraga, arXiv preprint arXiv:2012.06081 (2020).
[41] P. Beneventano, P. Cheridito, R. Graeber, A. Jentzen, and B. Kuckuck, arXiv preprint arXiv:2112.14523 (2021).
[42] U. M. B. Marconi, A. Puglisi, L. Rondoni, and A. Vulpiani, Physics reports **461**, 111 (2008).
[43] A. Gritsun and G. Branstator, Journal of the atmospheric sciences **64**, 2558 (2007).
[44] S. M. Allen and J. W. Cahn, Acta Metallurgica **20**, 423 (1972).
[45] A. Kolmogorov, I. Petrovskii, and N. Piskunov, Selected Works of AN Kolmogorov **1** (1937).
[46] A. M. Turing, Bulletin of mathematical biology **52**, 153 (1990).
[47] T. Callahan and E. Knobloch, Physica D: Nonlinear Phenomena **132**, 339 (1999).
[48] S. Kondo and T. Miura, science **329**, 1616 (2010).
[49] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, 2nd ed., Dover Books on Mathematics (Dover Publications, Mineola, NY, 2001).