Inferring Pattern Formation from Early-stage Snapshots via Long-Short-Range Neural Network

Qian Serena Hou[1,*] and Zecheng Gan[1,2,†]

[1] *Thrust of Advanced Materials, and Guangzhou Municipal Key Laboratory of Materials Informatics, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China*

[2] *Department of Mathematics, The Hong Kong University of Science and Technology, Hong Kong SAR, China*

(Dated: January 15, 2026)

Inferring long-term pattern evolution from *early-stage* data is a fundamental challenge across physical sciences, as traditional models rely on known governing equations, while data-driven methods typically demand extensive training datasets. We propose the Long-Short-Range Neural Network (LSR-Net), a model-free operator learning framework that overcomes this by learning from only two sets of early-stage snapshots. The key novelty of LSR-Net lies in decomposing the evolution operator into learnable long-/short-range integral kernels, effectively separating nonlocal interactions from local dynamics. Further introducing a *sum-of-exponentials* (SOE) ansatz for the long-range kernel reduces the computational complexity from $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log n)$. Tested on the Allen–Cahn, Cahn–Hilliard and Flory–Huggins models, LSR-Net achieves a prediction error $< 6\%$ for horizons $7\times$ beyond the training interval, outperforming advanced neural operator baselines and demonstrating robustness to low-quality measurements. Crucially, its successful application to infer the continental-scale Antarctic temperature field from real atmospheric data validates a general framework for predicting spatiotemporal dynamics in complex physical systems.

Introduction.–Pattern formation and interfacial dynamics are observed in a wide variety of physical, chemical, and material systems [1, 2]. Notable examples include the assembly of block copolymer blends [3], the emergence of patterns in colloidal electroconvection under 2D confinement [4–6], as well as in solidification fronts [7], oscillatory chemical reactions [8, 9] and excitable biological media [10]. Understanding these processes is of fundamental interest and has significant potential for technological applications, such as nanofabrication [11], heterogeneous catalysis [12], and the design of advanced materials [13].

A prototypical category of pattern formation phenomena is described by phase-field models, where the interface between two phases is captured by a steep yet continuous state variable $\phi(\boldsymbol{x}, t)$, whose evolution is driven by gradients of a nonconvex energy landscape [1, 14]. This category includes the well-known Allen–Cahn and Cahn–Hilliard equations [15, 16], and their variants such as the Swift–Hohenberg and Flory–Huggins models [17–19]. These models have found broad applications in phase-field-crystal (PFC) modeling [20], chemical kinetics [21], and polymer solutions [22].

Despite extensive numerical and theoretical investigations using phase-field models, several challenges remain. First, many intriguing natural phenomena occur across multiple spatial and temporal scales. Constitutive relations must effectively account for interactions between these scales, let alone the inherent nonlinearity and coupling among different physical fields (e.g., thermal, mechanical, and chemical). Although tremendous efforts have been made in this direction [23–29], obtaining adequate constitutive relations remains a challenge. Furthermore, ensuring the constructed models accurately predict real-world behavior requires thorough validation against experimental data, which can conflict with economic and time constraints. For instance, the polymer aging process can span decades. Therefore, predicting long-term pattern evolution from early-stage measurements would be highly beneficial [30]. However, this area remains largely overlooked at present.

In this Letter, we propose a novel Long-Short-Range Neural Network (LSR-Net) designed to accurately predict the pattern evolution from early-stage snapshots. By representing the forward evolution operator as a composition of long-/short-range learnable integral kernels with nonlinear activations, LSR-Net effectively captures global pattern conformation changes and resolves sharp interfaces locally. LSR-Net is model-free, i.e., it does not require constructing sophisticated constitutive relations, thus enabling its application to infer long-term pattern evolutions across diverse physical scenarios. In contrast to existing methods such as the Fourier Neural Operator (FNO) [31], which uses truncated Fourier series for long-range convolutions, we introduce the sum-of-exponentials (SOE) kernel approximation ansatz [32–34] with *learnable* weights and exponents. The proposed approach inherently captures the law of free-energy dissipation and exhibits exceptional robustness against noisy, low-resolution, and partially observed data. Applied to the ERA5 reanalysis dataset [35]–a comprehensive global climate record–for the challenging task of predicting the Antarctic 850 hPa temperature field evolution, LSR-Net successfully *extrapolates* to long-term variations despite being trained only on early-stage (12-hour) data. In this real-world test, it demonstrates superior performance compared to established neural operator baselines, in-

* qhou637@connect.hkust-gz.edu.cn

† Corresponding author. zechenggan@ust.hk

cluding FNO and the Spherical Fourier Neural Operator (SFNO) [36].

Learning evolution operators.–The goal of LSR-Net is to predict the system state $\phi_t = \phi(\boldsymbol{x}, t) \in \boldsymbol{\mathcal{X}}$ after a long-time evolution, given an initial condition $\phi_0 \in \boldsymbol{\mathcal{X}}$, where $\boldsymbol{\mathcal{X}}$ is a function space of states. Specifically, with two sets of *early-stage* observations $\{\boldsymbol{\phi}_0^{(i)}\}_{i=1}^N$ and $\{\boldsymbol{\phi}_T^{(i)}\}_{i=1}^N$ at $t = 0$ and T, we aim to build an approximation for the evolution operator as a learnable mapping $G_{\boldsymbol{\theta}}^T : \mathcal{X} \to \mathcal{X}$, $\boldsymbol{\theta} \in \Theta$ for some finite-dimensional parameter space Θ. This involves identifying $\boldsymbol{\theta}^\star$ such that $G_{\boldsymbol{\theta}^\star}^T[\phi_t] \approx \phi_{t+T}$ holds for all $t \geq 0$. In the operator learning framework, this problem can be approached in a model-free, data-driven manner by minimizing the mean squared error (MSE) between the predicted and reference system states at T, i.e.,

$$\boldsymbol{\theta}^\star = \arg\min_{\boldsymbol{\theta}} \frac{1}{N} \sum_{i=1}^N \left\| G_{\boldsymbol{\theta}}^T[\phi_0^{(i)}] - \phi_T^{(i)} \right\|^2 . \quad (1)$$

Once the evolution operator is accurately learned, long-term predictions can be achieved by iteratively applying $G_{\boldsymbol{\theta}^\star}^T$ to the current states, progressing as $\phi_0 \mapsto \phi_T \mapsto \phi_{2T} \mapsto \cdots$. It is important to note that learning operators is typically more challenging compared to function regression. Existing machine learning approaches, such as physics-informed neural networks (PINN) [37], essentially solve a regression problem, whose performance is enhanced by integrating known (or partially known) physical models. However, they are limited to single instances of system parameters. Meanwhile, operator learning frameworks like FNO [31] and DeepONet [38] still struggle to capture long-range dependencies, often resulting in inadequate long-term predictions [39, 40].

LSR-Net architecture.–As depicted in Fig. 1, the architecture of LSR-Net is composed of M learnable LSR-Net blocks. Each block integrates both long-range (LR) and short-range (SR) learnable integral kernels, followed by nonlinear activation functions σ. To prevent gradient exploding/vanishing during backpropagation, a residual connection is added into each LSR-Net block. This component is found to be essential for robust training with real-world atmospheric datasets. Denote the integral kernel transformation within each block as $\{g_{\boldsymbol{\theta}_i}\}_{i=1}^M$, the full evolution operator in LSR-Net is expressed as:

$$G_{\boldsymbol{\theta}}^T[\phi] = (\sigma \circ g_{\boldsymbol{\theta}_M}) \circ (\sigma \circ g_{\boldsymbol{\theta}_{M-1}}) \circ \cdots \circ (\sigma \circ g_{\boldsymbol{\theta}_1})[\phi] , \quad (2)$$

where each kernel integral operator $g_{\boldsymbol{\theta}_i}$ is defined by:

$$g_{\boldsymbol{\theta}_i}[\phi^i(\boldsymbol{x})] = \int_\Omega \mathcal{K}(\boldsymbol{x} - \boldsymbol{y}; \boldsymbol{\theta}_i)\, \phi^i(\boldsymbol{y})\, d\boldsymbol{y} . \quad (3)$$

Here $\mathcal{K}(\boldsymbol{x} - \boldsymbol{y}; \boldsymbol{\theta}_i)$ is a learnable, non-local kernel parameterized by $\boldsymbol{\theta}_i$, and $\phi^i(\boldsymbol{x})$ denotes the intermediate output from the previous block if $i > 1$. Consequently, $\boldsymbol{\theta} = \{\boldsymbol{\theta}_i\}_{i=1}^M$ constitutes the set of all trainable parameters, with σ chosen as `Tanh` in LSR-Net, which has proven effective in modeling particle/fluid interfaces [41]. Note that when physical models are known, tailored activation functions can be derived analytically under certain circumstances [42].

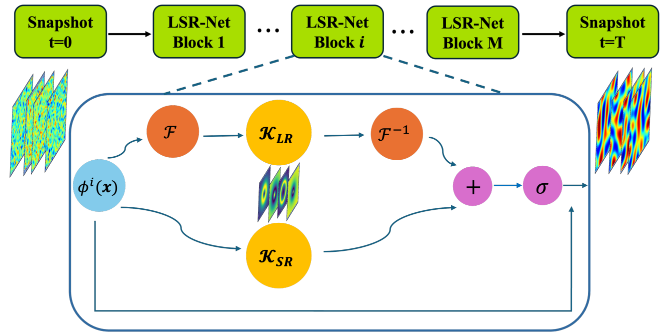


FIG. 1: The architecture of LSR-Net. Start from input snapshot at $t = 0$, the network applies M-layers of LSR-Net blocks, producing the predicted output snapshot at $t = T$. Each LSR-Net block integrates both long-range (LR) and short-range (SR) learnable integral kernels, followed by activation functions. The LR convolution is efficiently computed in Fourier space using the proposed sum-of-exponentials (SOE) kernel ansatz. A residual connection is added for each LSR-Net block to improve the training stability.

To effectively capture both non-local interactions and local fine-scale features in pattern formation, we further decompose $\mathcal{K}$ into LR and SR components:

$$\mathcal{K} = \mathcal{K}_{\rm LR} + \mathcal{K}_{\rm SR} . \quad (4)$$

In line with standard practices in convolutional neural networks (CNNs), the SR component can be learned using a convolution layer with a small kernel size (typically $5 \sim 7$). However, learning the LR component is challenging using standard convolution layers, which typically require a large kernel size or an increased number of layers. To overcome this issue, we propose the following sum-of-exponentials (SOE) ansatz for $\mathcal{K}_{LR}$:

$$\mathcal{K}_{LR}(\boldsymbol{x} - \boldsymbol{y}; \boldsymbol{\theta}_{LR}) \approx \sum_{l=1}^P \beta_l e^{-\alpha_l |\boldsymbol{x} - \boldsymbol{y}|} , \quad (5)$$

where $\boldsymbol{\theta}_{LR} = \{\alpha_l, \beta_l\}_{l=1}^P$ are the learnable weights and exponents for $\mathcal{K}_{LR}$, and P a hyperparameter representing the number of LR convolution channels. The SOE ansatz is grounded in approximation theory [43], which asserts that any smooth, and possibly long-range kernel can be approximated to arbitrary precision by a weighted sum of exponential functions. Note that Eq. (5) assumes translational symmetry of the evolution operator. In cases where this symmetry is broken, it can be addressed by lifting the SOE ansatz to higher dimensions [44]. Finally, we note that existing studies typically focus on identifying the optimal set of parameters $\{\alpha_l, \beta_l\}_{l=1}^P$ for *given* kernel function forms [34, 45, 46]. In contrast, here we do not assume a known kernel; rather, all parameters are learned during the training process, enabling the automatic discovery of *hidden* long-range interaction structures.

Efficient long-range convolutions.–Leveraging the SOE form in Eq. (5), we propose an efficient method to compute long-range convolutions. Based on the simple fact that $\left\{e^{-\alpha|x|},\ 2\alpha/(\alpha^2+k^2)\right\}$ forms a Fourier transform pair, the LR convolution can be efficiently handled in Fourier space via the convolution theorem:

$$\int_{\Omega} \mathcal{K}_{LR}(\boldsymbol{x}-\boldsymbol{y};\boldsymbol{\theta}_{LR})\,\phi(\boldsymbol{y})\,d\boldsymbol{y} = \mathcal{F}^{-1}\left[\widehat{\mathcal{K}}_{LR}(\boldsymbol{k})\cdot\widehat{\phi}(\boldsymbol{k})\right]\ , \tag{6}$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform, and $\widehat{\mathcal{K}}_{LR}(\boldsymbol{k})$ becomes a trainable *Fourier multiplier* in $\boldsymbol{k}$-space:

$$\widehat{\mathcal{K}}_{LR}(\boldsymbol{k}) = \sum_{i=1}^{P}\beta_i\left(\frac{\alpha_i}{k^2+\alpha_i^2}\right)\ . \tag{7}$$

In many real-world scenarios, the LR interaction kernel can be *oscillatory.* This can be accommodated by extending the exponents into the complex plane [45, 46]. Analogously, by using the fact that $\left\{e^{-\alpha|x|}\cos(sx),\ \alpha/[\alpha^2+(k+s)^2]+\alpha/[\alpha^2+(k-s)^2]\right\}$ is a Fourier transform pair, the Fourier multiplier can be modified as follows:

$$\widehat{\mathcal{K}}^{os}_{LR}(\boldsymbol{k}) = \sum_{i=1}^{P}\beta_i\left[\frac{\alpha_i}{(k+s_i)^2+\alpha_i^2}+\frac{\alpha_i}{(k-s_i)^2+\alpha_i^2}\right]\ ,$$

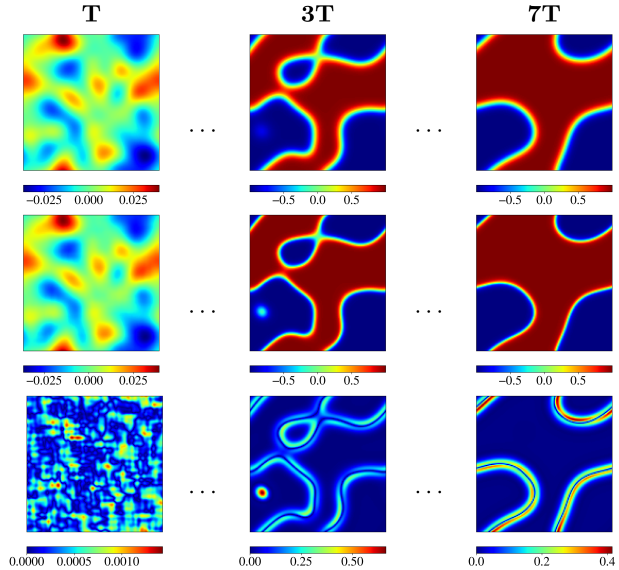
which simplifies to non-oscillatory case when $s_i = 0$. Note that even for the oscillatory case, there are only 3 trainable parameters per LR convolution channel. This not only yields a more compact model, but also leads to superior training accuracy compared to FNO and DeepOnet, as detailed in Supplementary Materials (SM), Sec. I. E.

Although the two sides of Eq. (6) are mathematically equivalent, their computational costs differ significantly. Direct computation of the LR convolution requires $\mathcal{O}(n^2)$ operations, where n is the total number of pixels in a snapshot image. In contrast, performing the computation in $\boldsymbol{k}$-space reduces the operations to $\mathcal{O}(n\log n)$, namely (1) take fast Fourier transform (FFT) of the input snapshot; (2) point-wise multiplication with the Fourier multiplier; and (3) inverse fast Fourier transform (iFFT) back. Finally, note that this approach assumes the early-stage measurements are standard images defined on a regular grid, so that FFT can be applied. If measurements are taken on a non-equispaced grid (or more generally on point clouds), the same process can be achieved using non-uniform FFT techniques [47–50]. We also implement a NUFFT-based version of LSR-Net. This version enables the model to handle measured data in the form of point clouds in $\mathbb{R}^3$, allowing it to predict temperature field evolutions over Antarctica.

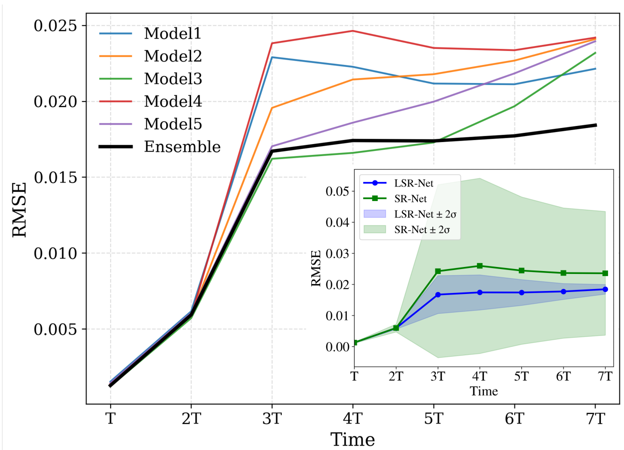
Long-term prediction via ensembling.–To test our approach, we generated simulated data for the evolution of the Allen–Cahn, Cahn–Hilliard, and Flory–Huggins models, where the true long-term dynamics are known. For each model, we collected $N = 500$ early-stage snapshots from the full simulation by varying initial conditions sampled from Gaussian random fields. These were split into 320 samples for training, 80 for validation, and 100 for testing. Importantly, to evaluate the long-term predictive capability of LSR-Net, we selected T such that the pattern formation process had only just begun. The trained model was then applied recursively to generate long-term predictions until the patterns nearly reached equilibrium. The inherent randomness in network initialization and training introduces significant variance across individual runs, which accumulates and degrades long-term prediction performance. To address this, we adopt an *ensembling* strategy to reduce prediction uncertainty [51]. Consistent behavior is observed across all three benchmark models. For clarity, the main text focuses on results from the Allen-Cahn dataset, while results for the Cahn–Hilliard and Flory–Huggins datasets, along with data preparation, model training and ensembling details, are provided in SM Sec. I. A-D.

The long-term prediction results of LSR-Net on the Allen–Cahn model are presented in Fig. 2a, where the first row shows the ground truth snapshots at representative times T, $3T$, and $7T$, while the second and third rows show the corresponding LSR-Net predictions and point-wise absolute errors, respectively. Despite some accumulated errors during the inference stage (see Fig. 2b), LSR-Net effectively captures the long-term pattern evolution features for the Allen–Cahn model, significantly extending beyond the training time horizon T. These findings demonstrate the model's capability to robustly learn and forecast long-term pattern evolution in complex systems. Notably, as shown in Fig. 2b, the ensemble predictions (black line) not only outperform the *best* individual model but also, crucially, exhibit significantly reduced long-term prediction uncertainty. Furthermore, a comparison with SR-Net (which omits the LR module; see inset) clearly highlights the advantage of LSR-Net: it simultaneously boosts long-term accuracy and reduces uncertainty, thereby validating the LSR-Net design.

It is well known that pattern formation described by models such as the Allen–Cahn and Cahn–Hilliard equations take the form of gradient flows to minimize the corresponding Ginzburg–Landau free-energy functionals. Therefore, it is crucial for the LSR-Net predictions to adhere to the fundamental physical law of energy dissipation. To validate this, we calculated the LSR-Net predicted free energies for the Allen–Cahn, Cahn–Hilliard, and Flory–Huggins models (SM, Fig. 1, Fig. 4(a-b)). In all cases, LSR-Net consistently preserves the energy dissipation law. Notably, LSR-Net does not require prior knowledge of free energy functionals; the physical law of energy dissipation is learned purely from early-stage snapshots. This distinguishes our method from recent works where the explicit energy functional forms are incorporated into the network design to preserve energy



(a) Long-term inference for the Allen–Cahn dynamics



(b) Improved prediction accuracy & reduced uncertainty via ensembling

FIG. 2: (a) Long-term inference of the Allen-Cahn dynamics by LSR-Net, where T the training time horizon. Top row: ground truth at times T , $3T$, and $7T$. Middle row: corresponding LSR-Net predictions. Bottom row: point-wise absolute errors. (b) Prediction RMSE *vs.* Inference time, with ensembling applied over 5 independent runs. Inset: ensemble-averaged RMSE (solid curve) and prediction uncertainty (shaded region) *vs.* time. Results from SR-Net (without LR modules) are included for comparison.

dissipation [52, 53].

Robustness against low-quality measurements.–To evaluate the robustness of LSR-Net, we conducted three types of tests with low-quality measurements: (a) noisy snapshots; (b) low-resolution data; and (c) partial observations. In Fig. 3, we present test results on the evolution of a star-shaped interface, a curvature-driven interfacial dynamics governed by the Allen–Cahn equation [53]. Details regarding the interfacial dynamics model setup, and extra robustness tests on the Cahn–Hilliard and Flory–Huggins datasets are provided in SM, Sec.I. F. The left panel of Fig. 3 presents the results with added noise to the original snapshots from full simulations, where the top row shows the ground truth; the second and third rows depict LSR-Net predictions with 5% and 10% added Gaussian noise, respectively. Despite the added noise, LSR-Net predictions reliably preserve the phase morphology and interfacial dynamics. Another interesting observation is that, though the training snapshots (at $t = 0$ and T) are noisy, the long-term dynamics inferred by LSR-Net automatically filters those and obtains smooth predictions, which can be understood from the *spectral bias* mechanism of neural networks [54, 55]. The right panel of Fig. 3 illustrates test results using low-resolution data and partial observations. The top two rows correspond to LSR-Net predictions trained on low-resolution snapshots with down-sampling factors of 2 and 4, respectively. The bottom row shows predictions based on partial observations, where only a subregion (the top-right quadrant) of the full snapshots is retained for LSR-Net training. Remarkably, LSR-Net successfully captures interfacial dynamics in both scenarios, underscoring its generalization capability towards real experimental applications.

Inferring Antarctic t_{850} pattern evolution.–We finally apply LSR-Net to predict the evolution of the Antarctic 850 hPa temperature field (t_{850})–a challenging real-world system with *unknown* governing equations. The model is trained on ERA5 reanalysis data from 1950–2024 [35], using the first two days of each month at 00:00 and 12:00 UTC to learn the 12-hour forward evolution operator. As before, predictions are enhanced via ensembling. Detailed data preparation, pre-processing, and training procedures are described in SM Sec. I. G. A comparison of 36-hour Antarctic t_{850} predictions from FNO, SFNO, and LSR-Net (under identical training and ensembling protocols) is presented in Fig. 4. While FNO over-smooths important variations and SFNO exhibits polar artifacts, LSR-Net successfully captures the low-temperature pattern, especially along the coastline, demonstrating LSR-Net's effectiveness in forecasting complex real-world spatiotemporal processes.

Conclusions & Discussions.–In this work, we propose the long-short-range neural network (LSR-Net), tailored for long-term inference in pattern formation scenarios. This model-free approach requires only two sets of early-stage snapshots. The key novelty lies in the integration of a long-range convolution module, which efficiently captures global interactions through a trainable Fourier multiplier derived from the sum-of-exponentials (SOE) ansatz in kernel approximation theory. Notably, this requires only $3 \sim 4$ trainable parameters per long-range convolution channel. Our design significantly reduces the model size by $1 \sim 3$ orders of magnitude compared to DeepOnet and FNO, while also achieving higher accuracy, especially for long-term predictions.

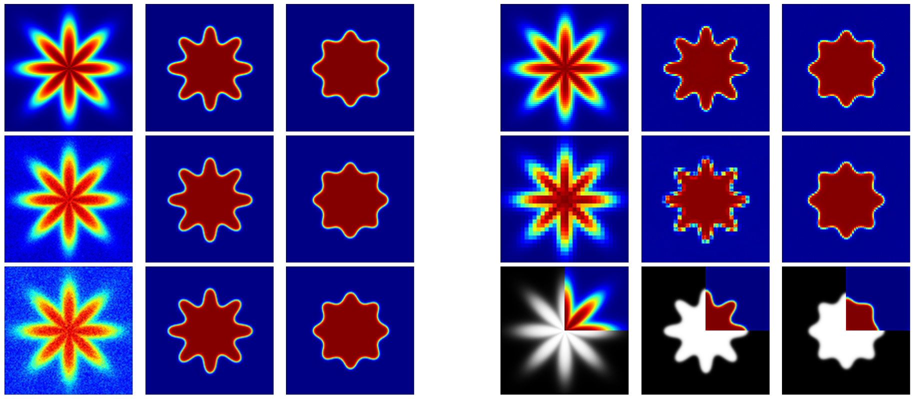
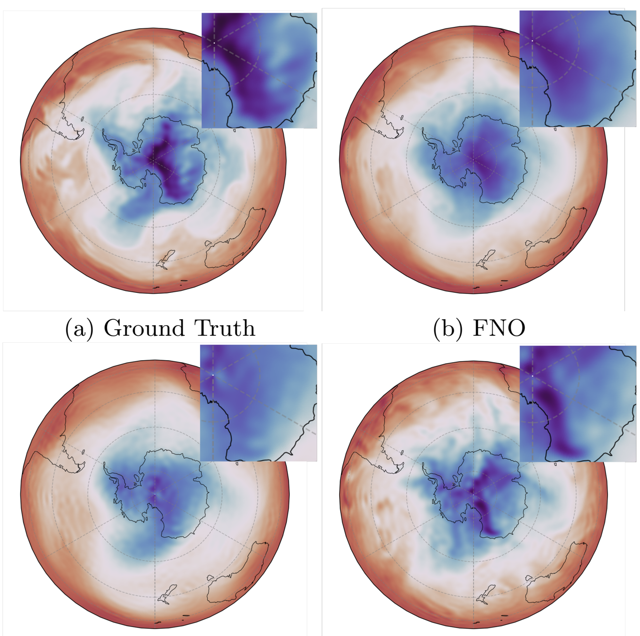


FIG. 3: Robustness of LSR-Net inference results tested on the evolution of a star-shaped interface. Left panel: tests against added noise, where the top row shows the ground truth; the second and third rows depict LSR-Net predictions with 5% and 10% added Gaussian noise, respectively. Right panel: the upper two rows correspond to LSR-Net predictions trained on low-resolution snapshots with down-sampling factors of 2 and 4, respectively; and the bottom row shows predictions based on partial observations, where only the top-right quadrant of the full snapshots is retained for LSR-Net training. In each panel, the three columns show inference results at T, $3T$, and $7T$, respectively.



(a) Ground Truth

(b) FNO

(c) SFNO

(d) LSR-Net

FIG. 4: Comparison of 36-hour forecasts for the Antarctic 850 hPa temperature field (t_{850}). (a) Ground truth. (b)-(d) Predictions using FNO, SFNO, and LSR-Net, respectively.

The performance of LSR-Net has been validated across various datasets collected from full simulations of the Allen–Cahn, Cahn–Hilliard, and Flory–Huggins models. A key result is that LSR-Net, trained directly on atmospheric datasets, successfully infers continental-scale Antarctic temperature field evolution. This demonstrates its capability for reliable prediction of large-scale, complex spatiotemporal dynamics, even in the absence of known governing equations. Interestingly, LSR-Net also exhibits exceptional robustness against low-quality measurements, including noisy snapshots, low-resolution images, and partial observations, highlighting its transformative potential for addressing challenging real-world applications.

In the future, we plan to use LSR-Net to infer thermodynamics/kinetic properties from real images of phase-separating materials [56], as well as from *in situ* X-ray scattering measurements of crystallization in reciprocal space [57, 58]. Additionally, we aim to test LSR-Net on datasets from advanced coherent diffractive imaging methods for noncrystalline materials [59]. We also aim to improve the performance of LSR-Net for large-scale systems via learning operators in reduced latent spaces [60], as well as optimize the algorithm implementation on GPUs [61], thereby expanding its applicability to more complex real-world scenarios.

We finally note that, as a model-free approach, LSR-Net is inherently less suited to directly providing new physical insights. In contrast, while model-based methods can enhance theoretical understanding, they often require substantially more–or more diverse–data to accurately learn the underlying constitutive relations [28, 29]. Ideally, these approaches can be *complementary*: a trained LSR-Net, learned from early-stage observations, can be used as a *generative model* to produce longer-term synthetic data, which may then support the discovery of underlying physical laws through model-based techniques.

Acknowledgments.–This work was supported by the Natural Science Foundation of China (Grant No. 12201146). The LSR-Net code and data used in this work is available on GitHub at *https://github.com/zcgan/LSR-Net-DataRepo.*

[1] M. C. Cross and P. C. Hohenberg, Rev. Mod. Phys. **65**, 851 (1993).
[2] M. Seul and D. Andelman, Science **267**, 476 (1995).
[3] M. P. Stoykovich, M. Muller, S. O. Kim, H. H. Solak, E. W. Edwards, J. J. De Pablo, and P. F. Nealey, Science **308**, 1442 (2005).
[4] Y. Han and D. G. Grier, J. Chem. Phys. **122** (2005).
[5] Y. Han and D. G. Grier, J. Chem. Phys. **125** (2006).
[6] Y. Han and D. G. Grier, J. Chem. Phys. **137** (2012).
[7] J. S. Langer, Rev. Mod. Phys. **52**, 1 (1980).
[8] A. Zaikin and A. Zhabotinsky, Nature **225**, 535 (1970).
[9] I. R. Epstein and K. Showalter, J. Phys. Chem. **100**, 13132 (1996).
[10] S. Sinha and S. Sridhar, *Patterns in excitable media: Genesis, dynamics, and control* (CRC Press, 2014).
[11] Y. A. Vlasov, X.-Z. Bo, J. C. Sturm, and D. J. Norris, Nature **414**, 289 (2001).
[12] W. Gao, Z. D. Hood, and M. Chi, Acc. Chem. Res. **50**, 787 (2017).
[13] H. Qiu, Y. Gao, C. E. Boott, O. E. Gould, R. L. Harniman, M. J. Miles, S. E. Webb, M. A. Winnik, and I. Manners, Science **352**, 697 (2016).
[14] I. Steinbach, Model. Simul. Mater. Sc. **17**, 073001 (2009).
[15] S. M. Allen and J. W. Cahn, Acta Metall. **27**, 1085 (1979).
[16] J. W. Cahn and J. E. Hilliard, J. Chem. Phys. **28**, 258 (1958).
[17] J. Swift and P. C. Hohenberg, Phys. Rev. A **15**, 319 (1977).
[18] P. J. Flory, J. Chem. Phys. **10**, 51 (1942).
[19] M. L. Huggins, J. Chem. Phys. **9**, 440 (1941).
[20] H. Emmerich, H. Löwen, R. Wittkowski, T. Gruhn, G. I. Tóth, G. Tegze, and L. Gránásy, Adv. Phys. **61**, 665 (2012).
[21] M. Z. Bazant, Acc. Chem. Res. **46**, 1144 (2013).
[22] M. Doi, *Introduction to polymer physics* (Oxford university press, 1996).
[23] F. Zhou, T. Maxisch, and G. Ceder, Phys. Rev. Lett. **97**, 155704 (2006).
[24] R. Malik, F. Zhou, and G. Ceder, Nat. Mater. **10**, 587 (2011).
[25] D. A. Cogswell and M. Z. Bazant, ACS nano **6**, 2215 (2012).
[26] L. Zhang, J. Han, H. Wang, R. Car, and W. E, Phys. Rev. Lett. **120**, 143001 (2018).
[27] G. H. Teichert, A. R. Natarajan, A. Van der Ven, and K. Garikipati, Comput. Methods in Appl. Mech. Eng. **353**, 201 (2019).
[28] H. Zhao, B. D. Storey, R. D. Braatz, and M. Z. Bazant, Phys. Rev. Lett. **124**, 060201 (2020).
[29] H. Zhao, H. D. Deng, A. E. Cohen, J. Lim, Y. Li, D. Fraggedakis, B. Jiang, B. D. Storey, W. C. Chueh, R. D. Braatz, *et al.*, Nature **621**, 289 (2023).
[30] Y. Tian, K. Li, R. Tian, and C. Lu, Sens. Actuator B-Chem. **393**, 134214 (2023).
[31] Z. Li, N. B. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. M. Stuart, and A. Anandkumar, in *9th International Conference on Learning Representations, ICLR* (2021).
[32] G. Beylkin and L. Monzón, Appl. Comput. Harmon. Anal. **19**, 17 (2005).
[33] G. Beylkin and L. Monzón, Appl. Comput. Harmon. Anal. **28**, 131 (2010).
[34] S. Jiang and L. Greengard, Commun. Pure Appl. Math. **61**, 261 (2008).
[35] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J.-R. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J.-N. Thépaut, Quarterly Journal of the Royal Meteorological Society **146**, 1999 (2020).
[36] B. Bonev, T. Kurth, C. Hundt, B. Lusch, M. Mardani, R. Anirudh, M. Chertkov, J. Kossaifi, Y. Mroueh, M. Farajtabar, *et al.*, in *Proceedings of the 40th International Conference on Machine Learning (ICML)* (PMLR, 2023) pp. 2806–2823.
[37] M. Raissi, P. Perdikaris, and G. E. Karniadakis, J. Comput. Phys. **378**, 686 (2019).
[38] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis, Nat. Mach. Intell. **3**, 218 (2021).
[39] Z. Li, W. Peng, Z. Yuan, and J. Wang, Phys. Fluids **35** (2023).
[40] D. Nayak and S. Goswami, arXiv:2505.17341 (2025).
[41] H. Tanaka and T. Araki, Phy. Rev. Lett. **85**, 1338 (2000).
[42] Y. Lan, Z. Li, J. Sun, and Y. Xiang, J. Comput. Phys. **491**, 112343 (2023).
[43] R. A. DeVore and A. Kunoth, *Multiscale, nonlinear and adaptive approximation* (Springer, 2009).
[44] S. Jiang, L. Greengard, and S. Wang, Adv. Comput. Math. **41**, 529 (2015).
[45] Z. Gao, J. Liang, and Z. Xu, J. Sci. Comput. **93**, 40 (2022).
[46] Y. Lin, Z. Xu, Y. Zhang, and Q. Zhou, arXiv:2503.03183 (2025).
[47] A. Dutt and V. Rokhlin, SIAM J. Sci. Comput. **14**, 1368 (1993).
[48] L. Greengard and J.-Y. Lee, SIAM Rev. **46**, 443 (2004).
[49] A. H. Barnett, J. Magland, and L. af Klinteberg, SIAM J. Sci. Comput. **41**, C479 (2019).
[50] Y. Peng, L. Lin, L. Ying, and L. Zepeda-Núñez, J. Comput. Phys. **473**, 111692 (2023).
[51] I. Khurjekar, I. Saha, L. Graham-Brady, and S. Goswami, arXiv preprint arXiv:2507.03863 (2025).
[52] G. Fan, T. Jin, Y. Lan, Y. Xiang, and L. Zhang, arXiv:2309.10002 (2023).
[53] R. Mattey and S. Ghosh, Comput. Methods in Appl. Mech. Eng. **439**, 117897 (2025).
[54] N. Rahaman, A. Baratin, D. Arpit, F. Draxler, M. Lin, F. Hamprecht, Y. Bengio, and A. Courville, in *International conference on machine learning* (PMLR, 2019) pp. 5301–5310.
[55] Z.-Q. J. Xu, Y. Zhang, and T. Luo, Commun. Appl. Math. Comput. , 1 (2024).
[56] J. Lim, Y. Li, D. H. Alsem, H. So, S. C. Lee, P. Bai, D. A. Cogswell, X. Liu, N. Jin, Y.-s. Yu, *et al.*, Science **353**, 566 (2016).

[57] M. C. Weidman, D.-M. Smilgies, and W. A. Tisdale, Nat. Mater. **15**, 775 (2016).

[58] J. J. Geuchies, C. Van Overbeek, W. H. Evers, B. Goris, A. De Backer, A. P. Gantapara, F. T. Rabouw, J. Hilhorst, J. L. Peters, O. Konovalov, *et al.*, Nat. Mater. **15**, 1248 (2016).

[59] J. Miao, T. Ishikawa, I. K. Robinson, and M. M. Murnane, Science **348**, 530 (2015).

[60] K. Kontolati, S. Goswami, G. Em Karniadakis, and M. D. Shields, Nat. Commun. **15**, 5101 (2024).

[61] Y. Kim, G. Ryu, and Y. Choi, Math. Probl. Eng. **2021**, 5263989 (2021).