Deep Ritz Neural Operator for Dendritic Growth via Energy Splitting

Chih-Kang Huang [1 2 3] **Miha Založnik** [1] **Ludovick Gagnon** [2] **Benoît Appolaire** [1]

Abstract

We propose a neural operator approach bridging classical convex-concave splitting with physics-informed learning to accelerate phase-field simulations. By training via an energy-splitting variational formulation, we enforce the energy dissipation property of the underlying models. We further introduce a Reaction-Diffusion Neural Operator architecture, specifically designed to incorporate the operator-splitting of diffusion and reaction terms in the model equation. For the anisotropic dendritic growth simulation, we show that our approach provides better generalization and higher accuracy than the data-driven training approach, while achieving faster inference than traditional Fourier spectral methods.

1. Introduction

Phase transformations play an important role in materials processing as they determine the structure and resulting properties of the material. For instance, during additive manufacturing (3D printing) of metal parts, the molten metal first transforms into a polycrystalline solid structure, followed by further phase transformations at solid state. The final properties depend on the nature, size, and shape of the resulting solid phases. Modern modeling of such phase transformations relies on phase-field models (Fix, 1983; Tourret et al., 2022), which represent the transforming phases through an order parameter (a phase field) that is typically -1 in one phase and $+1$ in the other. Phase-field evolutions follows a gradient flow of a thermodynamic functional (Plapp, 2011), and can be coupled to temperature, chemical composition, stress, or fluid flow.

[1]Université de Lorraine, CNRS, IJL, F-54000 Nancy, France [2]Université de Lorraine, Inria, CNRS, IECL, F-54000 Nancy, France [3]Université Côte d'Azur, Inria, CNRS, LJAD, Nice, France. Correspondence to: Chih-Kang Huang <chih-kang.huang@inria.fr>, Miha Založnik <miha.zaloznik@univ-lorraine.fr>, Ludovick Gagnon <ludovick.gagnon@inria.fr>, Benoît Appolaire <benoit.appolaire@univ-lorraine.fr>.

Proceedings of the AI4Physics Workshop at the 43rd International Conference on Machine Learning (AI4Physics@ICML 2026), Seoul, South Korea. 2026. Copyright 2026 by the author(s).

Such phase-field models are formulated as reaction-diffusion partial differential equations (PDEs), that are generally solved with classical discretization (finite-difference, finite-element, spectral methods, etc.). However, their multi-scale, nonlinear nature requires fine spatial and temporal discretizations, resulting in high computational costs. This makes large-scale simulations and extensive parametric studies challenging (Takaki, 2023). Deep learning–based surrogate models offer a promising alternative, with the potential to significantly accelerate the simulation of phase transformations (Choi et al., 2024).

State of the Art Neural operators approximate parametric PDE solutions by learning mappings between function spaces. This enables them to generalize across varying input conditions, resolutions, or boundary conditions without retraining. While *supervised data-driven methods* (Bhattacharya et al., 2021; Anandkumar et al., 2020; Lu et al., 2019; Li et al., 2020; Patel et al., 2021) have succeeded in cases such as the Allen-Cahn equation (Raonić et al., 2023) and ice crystal growth (Tseng et al., 2023), they often suffer from overfitting and long-time instability, which motivates hybrid neural-classical alternatives (Oommen et al., 2024).

Meanwhile, *unsupervised physics-informed* learning satisfies governing equations without requiring large datasets. Physics-Informed Neural Networks (PINNs) (Raissi et al., 2019; Cuomo et al., 2022) minimize PDE residuals directly and have been applied to the Allen-Cahn equation (Wight & Zhao, 2020; Mattey & Ghosh, 2022). However, PINNs require retraining for each new boundary or initial conditions.

Physics-Informed Neural Operators (PINOs) merge these concepts to generalize across the PDE parameter space. Recent works include finite-element frameworks for heat diffusion (Yamazaki et al., 2025), finite-difference schemes for the Allen-Cahn equation (Geng et al., 2024), and variational energy minimization with DeepONets (Lu et al., 2019) for phase-field models Li et al. (2023). Despite these advances, PINO training often fails to converge for highly nonlinear problems.

Main Contributions We introduce an energy-based neural operator approach for gradient flows, inspired by the Deep Ritz method (Yu et al., 2018) and energy splitting.

In Section 2, we propose *DeepRitzSplit*, a neural operator-based semigroup framework predicting the next time step solution from any input state via semi-implicit energy splitting. By learning the associated energy minimization, our approach ensures unconditional energy dissipation regardless of the chosen timestep, which is a property generally not guaranteed by finite-difference schemes, finite-element schemes, or conventional PINO architectures. Inspired by Bretin et al. (2022), we also design a Reaction-Diffusion Neural Operator (RDNO) tailored for semi-implicit schemes.

In Section 3, we apply our method to anisotropic dendritic growth (Karma & Rappel, 1998). We demonstrate that our approach is faster than conventional spectral schemes while accurately preserving morphological characteristics, such as solid phase fraction and tip velocity. This establishes DeepRitzSplit as a physically consistent and efficient surrogate for large-scale phase-field simulations. Despite a trade-off in tip radius accuracy, training on one to three grains allows generalization to multi-grain configurations. Finally, DeepRitzSplit can serve as a preconditioner for training neural operators by minimizing scheme residuals, which significantly reduce the training time that would otherwise be required.

2. Deep Ritz Neural Operators via energy splitting

Operator splitting schemes have been successful for the numerical approximation of a wide range of problems, they allow to break a problem into subproblems that are easier to solve separately.

In this section we introduce a semi-implicit splitting scheme based on Eyre (1998); Zaitzeff et al. (2021) for general gradient flows that preserves the energy dissipation property. Next, we propose *DeepRitzSplit* — an approach to train neural operators by minimizing a variational formulation of the splitting scheme. The training is unsupervised and relies exclusively on the model equations. Finally, we also propose a *Reaction-Diffusion Neural Operator* (RDNO) architecture based on the convex-concave splitting scheme, for efficient training and improved interpretability.

2.1. An enegetically stable semi-implicit scheme for gradient flows

Consider the following evolution equation as L^2-gradient flow for a given energy functional $E : V \to \mathbb{R}$,

$$\partial_t u = -\nabla_{L^2} E(u) \quad \text{on} \quad \Omega \times (0, T] \tag{1}$$

for some $T > 0$, where $\nabla_{L^2} E(u)$ is the Gateaux derivative of E at $u \in V$, Ω is a domain of $\mathbb{R}^2$ and V is a Hilbert subspace of $L^2(\Omega)$.

Solutions to the gradient flow (1) have the energy dissipation property:

$$\frac{d}{dt} E(u) = \langle \nabla_{L^2} E(u), \partial_t u \rangle_{L^2} = -\|\nabla_{L^2} E(u)\|_2^2 \leq 0 \tag{2}$$

where $\langle \cdot\,, \cdot \rangle_{L^2}$ is the scalar product on $L^2(\Omega)$.

Initially introduced by Eyre (1998) for the Allen-Cahn equation and later generalized by Zaitzeff et al. (2021) for general gradient flows, the semi-implicit *splitting* scheme for gradient flows (1) first consists of splitting the energy E into

$$E(u) = E_1(u) + E_2(u) \quad \forall u \in V.$$

Then in the numerical discretization scheme we handle E_1 implicitly and E_2 explicitly:

$$\frac{u_{n+1} - u_n}{\Delta t} = -\nabla_{L^2} E_1(u_{n+1}) - \nabla_{L^2} E_2(u_n) \tag{3}$$

with the given time step size $\Delta t > 0$. The above scheme corresponds to the optimization problem

$$u_{n+1} := \arg\min_{u \in V} \Big(\frac{1}{2\Delta t} \|u - u_n\|_2^2 + E_1(u) + E_2(u_n) + \langle \nabla_{L^2} E_2(u_n), u - u_n \rangle_{L^2} \Big). \tag{4}$$

In particular, in the case where E_2 is concave, that is, for every $u, v \in V$, the second-order Gateaux differential verifies

$$D^2 E_2(u)(v, v) \leq 0,$$

and where the optimization problem (4) is well-posed, the energy dissipation property (5) holds for the corresponding numerical scheme (3) and for every time step $\Delta t > 0$:

$$E(u_{n+1}) \leq E(u_n). \tag{5}$$

Such a numerical scheme is said to be *unconditionally energetically stable*.

2.2. Deep Ritz method with energy splitting

The *Deep Ritz method*, introduced by Yu et al. (2018), is a neural network-based approach for solving PDEs arising from variational problems. Rather than solving the PDE directly, Deep Ritz approximates solutions by minimizing an energy functional, replacing traditional basis functions with deep neural networks.

Unlike PINNs, which minimize PDE residuals using expensive automatic differentiation, Deep Ritz employs a variational formulation suited for gradient flows. By eliminating higher-order derivatives in the loss function, it significantly reduces computation time and avoids related derivative errors in exchange for numerical integration errors.

In this work, we propose *DeepRitzSplit*, which extends the Deep Ritz method to gradient flows by incorporating the functional (4) of the splitting scheme (3) into the loss formulation. The free energy is decomposed into convex and concave parts; at each time step, the concave part is linearized around the current state, while the convex part is retained in its original form. The loss follows the variational structure (4) of the unconditionally energy-dissipative splitting scheme. This ensures the neural operator preserves the energy dissipation property (2) regardless of time step size.

Namely, we train a neural operator $\mathcal{NN}(\cdot\,;\boldsymbol{\theta}^*)$ parameterized by a set of optimal parameters $\boldsymbol{\theta}^*$, such that for every $u_n \in V$, the output of the neural operator approximates the solutition corresponding to the next time step:

$$\mathcal{NN}(u_n;\boldsymbol{\theta}^*) \simeq u_{n+1} = \arg\min_{u\in V} \mathcal{L}_{\text{split}}(u, u_n) \tag{6}$$

where $\mathcal{L}_{\text{split}}$ is the loss function defined by

$$\mathcal{L}_{\text{split}}(u, u_n) := \Bigg(\frac{1}{2\Delta t}\|u - u_n\|_{L^2}^2 + E_1(u) + E_2(u_n) + \langle \nabla_{L^2} E_2(u_n), u - u_n\rangle_{L^2} \Bigg). \tag{7}$$

for every $u_n \in V$. Or, in other words, we want to find

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \mathbb{E}_{u\sim\mu}(\mathcal{L}_{\text{split}}(\mathcal{NN}_{\boldsymbol{\theta}}(u), u))$$

where u is sampled from a given data distribution μ on V.

In Section 3, we will train the neural operators on the dendritic growth model, using the variational formulation (6) in DeepRitzSplit as the loss function to ensure that the energy dissipation property (2) is *approximately intrinsically* preserved. We investigate its performance as function of the neural network architecture and compare it to conventional data-driven learning approaches.

Various approaches incorporate the energy dissipation property (2), particularly for the Allen-Cahn equation. For instance, Kütük & Yücel (2024) enforce dissipation directly in the loss function as a penalty term, while Zhang et al. (2024) introduce a modified energy functional to ensure it. However, these methods rely on conventional data-driven training, whereas our approach is unsupervised and requires only the initial conditions of the system.

2.3. Reaction-Diffusion Neural Operator

In addition to the DeepRitzSplit method introduced in Section 2.2, we also propose here a novel Reaction-Diffusion-based Neural Operator (abbreviated as RDNO throughout the paper), inspired by the structure of the splitting scheme (3). Thanks to the splitting scheme (3), we can write

$$\mathcal{L}(u_{n+1}) = u_n - \Delta t \nabla E_2(u_n) \tag{8}$$

where $\mathcal{L} : V \to V$ is defined by

$$\mathcal{L}(u) := u + \Delta t \nabla E_1(u). \tag{9}$$

If there exists a left inverse operator $\mathcal{D} : V \to V$ of $\mathcal{L}$, that is,

$$\mathcal{D} \circ \mathcal{L} = \mathrm{Id}_V \tag{10}$$

where Id_V is the identity operator on V, then together with Eqs. (8) and (9), we obtain that

$$u_{n+1} = \mathcal{D} \circ \mathcal{R}(u_n) \tag{11}$$

with $\mathcal{R}(u_n) := u_n - \Delta t \nabla E_2(u_n)$, for every $u_n \in V$.

The above observation leads to consider the following Reaction-Diffusion Neural Operator

$$\mathcal{NN}(\cdot\,;\boldsymbol{\theta}) = D_{\boldsymbol{\theta}} \circ R_{\boldsymbol{\theta}}, \tag{12}$$

which is composed of two main blocks, as illustrated in Fig. 1. The first block $\mathcal{R}_{\boldsymbol{\theta}}$, called the *reaction block*, usually consists of a stack of convolutional layers designed to approximate the reaction operator $\mathcal{R}$ as defined in (12). The second block, the *diffusion block*, represents the diffusion operator $\mathcal{D}$ in (10).

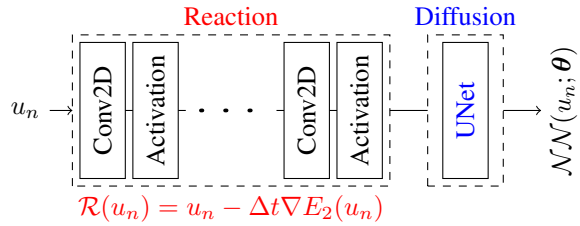

Figure 1. An illustration of Reaction-Diffusion Neural Operator (RDNO) with a UNet as diffusion block. Alternatively, one can also simply prescribe the reaction block by explicitly encoding $\mathcal{R}$ defined in (11).

Various architectures can be considered for the diffusion block. For instance, Fig. 1 illustrates a UNet (Ronneberger et al., 2015) in the RDNO architecture. UNet consists of a contracting path (encoder) for context via downsampling and a symmetric expanding path (decoder) for localization via upsampling, using skip connections to retain spatial information.

Alternatively, the reaction block can be replaced by explicitly encoding $\mathcal{R}$ in (11) rather than parameterizing it via a neural network $\mathcal{R}_{\boldsymbol{\theta}}$. The approach enhances **interpretability**, as the operator approximates the (left) inverse of an operator for which an analytical expression may be difficult to derive.

Inspired by Bretin et al. (2022), this architecture mirrors those used for the Allen-Cahn equation. However, while they used a supervised data-driven approach based on the

splitting scheme (3) samples, our method differs in structure. We invert their order of reaction and diffusion blocks and replace their multi-layer perceptron reaction block with a convolutional one. This aligns more naturally with the splitting scheme and reduces trainable parameters, resulting in faster training.

3. Neural operators for anisotropic dendritic growth

Compared to studies on the isotropic Allen-Cahn equation, the use of neural networks for anisotropic dendritic growth during solidification is relatively unexplored due to the mathematical complexity of the anisotropy term. Prior works (Tseng et al., 2023; Oommen et al., 2024) employ exclusively supervised data-driven approaches. In contrast, we use the unsupervised physics-informed DeepRitzSplit neural operator. Furthermore, we provide a comprehensive analysis of predicted dendritic morphology, including the dendrite tip velocities and radii.

We consider the growth of dendritic crystal grains from a pure liquid using the quantitative phase-field model of Karma & Rappel (1998). The variational formulation of the model is given by the system of two coupled PDEs that describe the phase evolution and the heat transfer:

$$\begin{cases} \tau \partial_t \phi &= -\nabla_\phi E_\mathrm{d}(\phi, U) \\ \partial_t U &= D\nabla^2 U + K h'(\phi)\partial_t \phi \,, \end{cases} \tag{13}$$

where ϕ is the phase of the material such that $\{\phi = -1\}$ defines the liquid state and $\{\phi = 1\}$ defines the solid state, and U is the *dimensionless* temperature, governed by heat diffusion and by the release of latent heat at the interface during solidification, such that $U = 0$ is the dimensionless melting temperature. $\tau > 0$ is the relaxation time, $D > 0$ is the diffusion coefficient and $K > 0$ is the latent heat parameter. The evolution of ϕ corresponds to the gradient flow of the free energy E_d, defined by

$$E_\mathrm{d}(\phi, U) = \int_\Omega \left(\frac{1}{2}|a(\phi)^2||\nabla\phi|^2 + \frac{\lambda_0}{2\epsilon K}U^2 + \frac{W(\phi)}{\epsilon^2} + \frac{\lambda_0}{\epsilon}h(\phi)U \right) d\boldsymbol{x} \tag{14}$$

with $h(s) := s^5/5 - 2s^3/3 + s$. The double-well potential is defined as $W = h'/4$ and the interface thickness as $\epsilon > 0$. The interface energy anisotropy is given by $a(\phi) = 1 + \sigma\cos(m\Theta)$, where $\Theta = \arctan(n_2/n_1)$ is the in-plane azimuthal angle. This angle is determined by the normal field $(n_1, n_2) = \vec{n} = \nabla\phi/|\nabla\phi|$, which specifies the interface orientation relative to the crystalline axis (aligned here with the horizontal axis). Furthermore, $m \in \mathbb{N}$ denotes the symmetry order of crystallization and $\sigma > 0$ the anisotropy strength. Finally, we set $\lambda_0 = D\tau/(0.6267\epsilon)$ to accurately reproduce negligible interface kinetics in the thin-interface limit (Karma & Rappel, 1998). The well-posedness of the system (13) including a regularization term in the free energy is established in Miranville (2015).

Two examples of dendrite growth for $\sigma = 0.01$ and $\sigma = 0.05$ are shown in Fig. 2, illustrating growth evolution and the effect of anisotropy strength. [1] For all following simulations, initial conditions consist of a solid disk of radius 5ϵ and an uniform initial temperatures of $U_0 = \kappa = -0.3$ in the liquid and $U_0 = 0$ in the solid. Boundary conditions are chosen periodic.

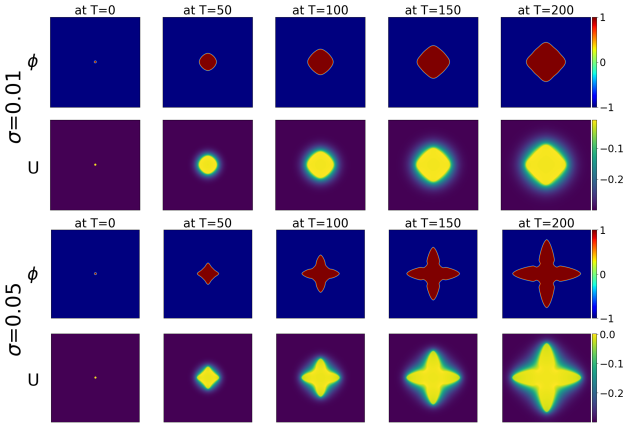

Figure 2. Dendritic growth simulations with fourfold symmetry ($m = 4$) and anisotropy strengths $\sigma = 0.01$ (top) and $\sigma = 0.05$ (bottom) on a unit square. Simulation with the SAV method (Guo et al., 2024).

In order to apply the DeepRitzSplit method (introduced in Section 2.2) to the phase-field model (13), we split the free energy E_d, defined by (14), into two parts:

$$E_\mathrm{d}(\phi, U) = E_1(\phi, U) + E_2(\phi, U)\,,$$

where $E_1(\phi, U) = \int_\Omega \left(\frac{1}{2}a(\phi)^2|\nabla\phi|^2 + \beta\frac{\phi^2}{2\epsilon^2} \right) d\boldsymbol{x}$ and

$$E_2(\phi, U) = \int_\Omega \left(\frac{\lambda_0}{2\epsilon K}U^2 + \frac{W(\phi)}{\epsilon} + \frac{\lambda_0}{\epsilon}h(\phi)U - \frac{\beta}{2\epsilon^2}\phi^2 \right) d\boldsymbol{x} \tag{15}$$

for some $\beta > 0$.

Then, the corresponding solution scheme, employing the semi-implicit splitting scheme for the phase-field equation and a fully implicit scheme for the heat diffusion equation,

[1] We restrict our analysis to $\sigma < 1/15 \approx 0.067$. For $\sigma \geq 1/15$, the free energy E_d defined in (14) no longer remains convex, leading to the formation of *faceted dendrites* which are outside the scope of the present paper.

is written as:

$$\begin{cases} \tau \dfrac{\phi_{n+1} - \phi_n}{dt} = -\nabla_\phi E_1(\phi_{n+1}, U_n) - \nabla_\phi E_2(\phi_n, U_n) \\ \dfrac{U_{n+1} - U_n}{dt} = D\nabla^2 U_{n+1} + Kh'(\phi_{n+1}) \dfrac{\phi_{n+1} - \phi_n}{\Delta t}. \end{cases} \tag{16}$$

with initial conditions $\phi_0, U_0 \in H^1(\Omega) \cap L^\infty(\Omega)$. For physical consistency, we also impose $\|\phi_0\|_\infty \leq 1$ and $\kappa \leq U_0 \leq 0$ with κ the initial dimensionless temperature of the liquid.

The system (16) leads to the following minimization problem:

$$\phi_{n+1} := \arg\min_{\phi \in V} \mathcal{L}(\phi, \phi_n, U_n) \tag{17}$$

where $\mathcal{L}$ is defined by

$$\mathcal{L}(\phi, \psi, U) := \Bigg(\frac{1}{2\Delta t} \|\phi - \psi\|_2^2 + E_1(\phi, U) + E_2(\psi, U) + \langle \nabla_\phi E_2(\psi, U), \phi - \psi \rangle_{L^2} \Bigg) \tag{18}$$

for every $\phi, \psi, U \in H^1(\Omega) \cap L^\infty(\Omega)$.

Note that this minimization problem has a solution for all ϕ_n, U_n since the functional in Eq. (16) is C-elliptic with $C \geq (1-\sigma)^2/4 > 0$. Thanks to the maximum principle, we have that $\|\phi(t)\|_\infty \leq 1$ and $\kappa \leq U(t) \leq 0$ for all $t \geq 0$, which gives

$$\|D_\phi^2 E_2\|_2 \leq \frac{1}{\epsilon}(2 + 4\lambda_0 \epsilon \kappa (\sqrt{1/3}^3 - \sqrt{1/3}) - \beta) \tag{19}$$

where D_ϕ^2 is the second-order Gateaux differential at ϕ on $L^2(\Omega)$. Therefore, with a well-chosen $\beta > 0$, depending only on κ, λ_0 and ϵ, E_2 becomes concave, so that the numerical scheme (16) is unconditionally stable and thus the energy dissipation property (2) is retrieved.

The presence of the anisotropic term makes solving the system (16) with a direct numerical solution method a notoriously difficult challenge in the classical sense. This motivates the use of a neural operator to solve the underlying optimization problem (17). More precisely, we aim to construct a neural operator $\mathcal{NN}(\cdot\,; \boldsymbol{\theta})$ such that

$$\mathcal{NN}(\phi_n, U_n; \boldsymbol{\theta}) \approx \phi_{n+1}\,, \tag{20}$$

where ϕ_{n+1} is a minimizer of the problem (17) and thus the solution of the scheme given by (16). This leads to minimizing the following empirical loss:

$$\frac{1}{|K|} \sum_{(\phi_n, U_n) \in K} \mathcal{L}(\mathcal{NN}(\phi_n, U_n; \boldsymbol{\theta}), \phi_n, U_n)\,, \tag{21}$$

where training configuration set K is a finite subset of $(H^1(\Omega) \cap L^\infty(\Omega))^2$.

To benchmark the neural operator for dendritic growth, we use the benchmark problem from the Phase Field Community Hub. All parameters are given in Table 1. Note that the spatial coordinate scaling differs from the published formulation, we use the square domain side length as the characteristic length rather than the interface thickness ϵ. Finally, thanks to Eq. (19), we choose $\beta = 14.4$ to ensure the concavity of E_2 as $\|D_\phi^2 E_2\| \lesssim -\frac{1}{\epsilon} 11.66 < 0$.

Table 1. Parameters for numerical experiments of the phase-field model of dendritic growth. All parameters are dimensionless.

Anisotropy strength	σ	0.01 or 0.05
Symmetry order	m	4
Interface thickness	ϵ	$1/400$
Relaxation time	τ	$1.6 \cdot 10^5$
Coupling parameter	λ_0	638
Diffusion coefficient	D	$6.25 \cdot 10^{-5}$
Latent heat parameter	K	0.5
Initial temperature	κ	-0.3
Splitting parameter	β	14.4
Timestep	Δt	0.04
Grid spacing	h	$1/400$

3.1. Numerical experiments

Training configurations We consider dendrites growing in a square unit *periodic* domain from $n \in \{1, 2, 3\}$ initial nuclei of radius 5ϵ into an undercooled liquid at $U = \kappa$. With $c_i \in [0, 1]^2_{\text{per}}$, the initial conditions are:

$$\phi_0(x, y) = \max_{i=1,\ldots,n} \left\{ \tanh\left(\frac{5\epsilon - \|(x, y) - c_i\|}{\sqrt{2}\epsilon} \right) \right\}, \qquad U_0 = \begin{cases} 0 & \text{on } \{\phi_0 > 0\} \\ \kappa & \text{otherwise.} \end{cases} \tag{22}$$

For training, we use 6,000 sample configurations per anisotropy strength σ. These are derived from 12 initial conditions: four random arrangements each for one, two, and three grains, with 500 temporal snapshots extracted from each evolution.

The neural operators are trained using all 6,000 (ϕ_n, U_n) pairs rather than just the initial state (ϕ_0, U_0). In the supervised data-driven training, the network minimizes the discrepancy between its prediction and the next time-step solution ϕ_{n+1}, generated by applying the SAV method (see Experiments Settings below) to the current state (ϕ_n, U_n).

Experiments Settings The reference for comparing all neural methods is the Scalar Auxiliary Variable (SAV) scheme for the Karma-Rappel phase-field model of anisotropic dendritic growth, recently developed by Guo et al. (2024). We use Fourier spectral spatial discretization

for the SAV method, for the reader's convenience, technical details are provided in Appendix A.1.

We first evaluate the unsupervised DeepRitzSplit method using the loss function in Eq. (21) against a conventional supervised approach. The supervised training utilizes the same dataset as DeepRitzSplit, minimizing the discrepancy between neural operator predictions and SAV reference solutions. Both methods employ the identical neural operator architecture illustrated in Fig. 1.

We consider two neural architectures for our benchmark:

- RDNO: The diffusion block uses a UNet with 4 downscaling/upscaling levels, 8 hidden channels, width 3, periodic padding, and tanh activation. The reaction block explicitly encodes

$$\mathcal{R}(\phi_n, U_n) := \phi_n - \frac{\Delta t}{\tau}\nabla E_2(\phi_n, U_n).$$

- UNet: the same structure as the RDNO diffusion block for the entire operator.

The RDNO architecture enhances training efficiency and interpretability as it aligns with the scheme (11), whereas the standard UNet functions as a black-box regression model.

All the training and evaluations were performed using a RTX6000 ADA GPU. The early stopping is applied once the loss fluctuates by less than 3‰.

Evaluation metrics To assess DeepRitzSplit approximations of the dendritic growth model, we evaluate the neural operator's prediction of single-grain evolution. A solid particle is placed in the unit square $[0, 1]^2$ at solidification temperature $U = 0$, surrounded by liquid at a uniform temperature $\kappa < 0$ (initial condition (22) for $n = 1$ and $c_1 = (0.5, 0.5)$).

We compare predictions against analytical solutions for growth speed and tip radius given by the dendritic growth theory (Barbieri & Langer, 1989; Dantzig & Rappaz, 2016). This theory shows that the tip of an isothermal dendrite branch growing into an infinite liquid settles into a constant growth speed, V_{tip}, and a parabolic tip shape with radius of curvature ρ_{tip} at the vertex. The solution of the Stefan problem for such a parabola (Ivantsov, 1947) shows that the tip speed and radius are related to the dimensionless temperature difference that drives its growth, κ, by

$$\sqrt{\pi Pe}\exp(Pe)\text{erfc}(\sqrt{Pe}) + \kappa = 0, \tag{23}$$

where $Pe = \rho_{\text{tip}}V_{\text{tip}}/(2D)$ is the tip Péclet number, D is the thermal diffusivity, and erfc is the complementary error function. A tip growing sufficiently far from the domain boundaries approximates this model, and we use Pe given by Eq. (23) as reference value.

For the comparisons, we also track the approximated perimeters, given by $\int_\Omega \left(\frac{|\nabla\phi|^2}{2} + \frac{W(\phi)}{\epsilon^2}\right) d\boldsymbol{x}$ (Modica, 1977), the free energy, E_{d}, given by Eq. (14), and the evolution of the solid phase fraction, defined by $\int_\Omega \left(\frac{1+\phi^\epsilon}{2}\right) d\boldsymbol{x}$.

Results Comparisons of single-grain results for $\sigma = 0.05$ using various neural operators are shown in Fig. 6 (Appendix A.2). We observe that the neural operator trained via a conventional data-driven approach fails to capture the solution. Conversely, DeepRitzSplit using both RDNO and UNet performs relatively well, with RDNO providing a better growth rate approximation and better symmetry. Similar results for $\sigma = 0.01$ are provided in Fig. 7 (Appendix A.2).

Training and evaluation times for the DeepRitzSplit model are listed in Table 2. Neural operators require approximately 8 hours for the loss convergence. Evaluation times for both architectures show that the neural approach is faster than the SAV direct numerical scheme.

Table 2. Training time for neural operators with DeepRitzSplit and comparison of evaluation times to the direct numerical solution with SAV over 5000 iterations.

	Trainable Parameters	**Time**	
		Train(hrs)	Eval.(s)
RDNO	589326	≈ 8	184
UNet	589326	≈ 8	183
SAV	n/a	n/a	269

Relative errors for the evolution of the solid phase fraction, perimeter, and free energy at various time points are presented in Table 3. All errors remain within a few percent. The largest errors occur in the perimeter, resulting from the somewhat more slender dendrite branches produced by the DeepRitzSplit-RDNO at large times.

Table 3. Relative errors (%) of DeepRitzSplit for solid fraction, perimeter, and energy compared to the SAV reference in single-grain configurations.

σ	**Time**	**Model**	**Solid Frac.**	**Perimeter**	**Energy**
0.01	$T_1 = 150$	RDNO	2.13	−0.77	−0.96
		UNet	−0.65	−2.12	−0.48
	$T_2 = 450$	RDNO	1.80	−2.12	−2.37
		UNet	−1.46	−4.59	0.71
0.05	$T_1 = 100$	RDNO	2.03	2.29	−0.81
		UNet	−2.08	−3.11	−0.37
	$T_2 = 200$	RDNO	1.89	5.92	−2.17
		UNet	−1.49	−1.15	−0.30

Next, we evaluate whether these predictions capture the underlying physics by analyzing tip velocities and radii at steady-state. The evolutions of V_{tip} and ρ_{tip} are shown in Fig. 3 for $\sigma = 0.05$ and in Fig. 8 (Appendix A.2) for

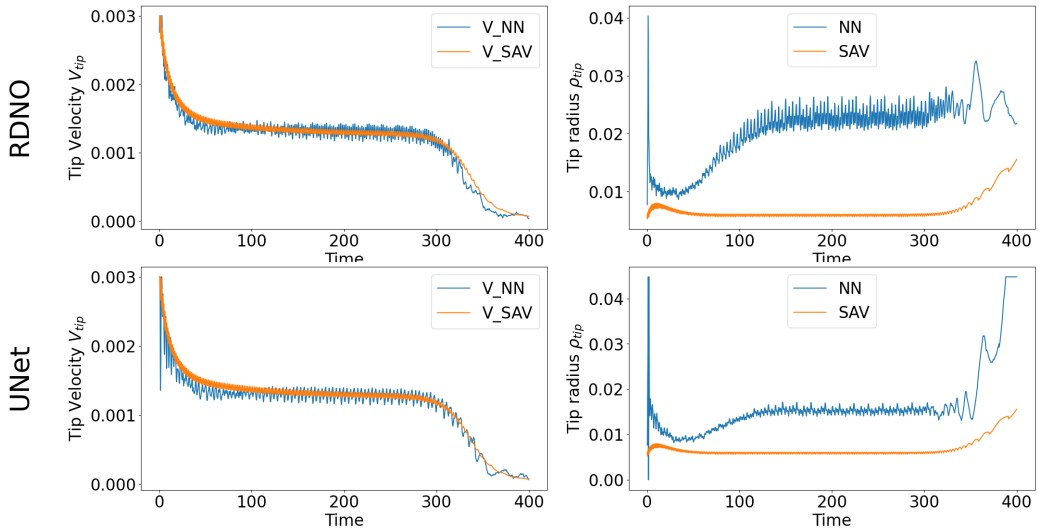

Figure 3. Evolution of tip velocity V_{tip} and radius ρ_{tip} for RDNO (top) and UNet (bottom) compared to the SAV numerical reference for $\sigma = 0.05$. Note that the growth slows down when the tip approaches the domain boundary.

$\sigma = 0.01$. DeepRitzSplit predictions shows high accuracy, closely matching the velocity evolution of the SAV reference. However, discrepancies become pronounced regarding the tip radius, where the evolution depends strongly on both the solution method and the specific neural architecture.

The results on the steady-state dendrite tip operating state (velocity and radius), presented in Table 4 indicate that while the SAV scheme and both neural operators yield accurate tip velocity predictions, they *all* fail to properly characterize the dendrite tip radius.

Table 4. Dendrite tip velocity and radius of curvature during steady-state growth. Results compare DeepRitzSplit (RDNO/UNet) against numerical (SAV) and analytical references. V_{tip} and ρ_{tip} are averaged at $T = 400$ for $\sigma = 0.01$ and at $T = 200$ for $\sigma = 0.05$.

σ	Metric	Reference		DeepRitzSplit		Rel. Error (%)	
		Analyt.	SAV	RDNO	UNet	RDNO	UNet
0.01	V_{tip} $(\cdot 10^{-4})$	n/a	5.45	5.57	4.94	2.2	9.3
	ρ_{tip} $(\cdot 10^{-2})$	n/a	7.20	7.93	0.65	10.2	39.4
	Pe $(\cdot 10^{-1})$	0.45	3.14	3.53	2.56	12.6	18.5
0.05	V_{tip} $(\cdot 10^{-3})$	n/a	1.29	1.32	1.30	-0.6	-1.2
	ρ_{tip} $(\cdot 10^{-2})$	n/a	0.59	2.25	1.52	280.2	157.0
	Pe $(\cdot 10^{-2})$	4.48	6.12	23.8	1.58	283.9	158.7

3.2. Application to multi-grain simulations

To assess out-of-distribution performance, we test the trained neural operators on simulations featuring more grains than the training set. This evaluation determines if training on limited configurations with few grains generalizes to simulate a wide variety of grain ensembles and spatial arrangements. Results for 16 grains across three arrangements (regular, uniform, and cluster) are shown in Fig. 10 (Appendix A). Dimensionless temperature differences for RDNO and UNet relative to the SAV method are illustrated in Fig. 5, with corresponding maximum temperatures summarized in Table 5.

Table 5. Maximum dimensionless temperatures for multi-grain simulations across all methods and spatial arrangements at anisotropy strength $\sigma = 0.05$.

Spatial Arrangements	**Evaluation methods**		
	SAV	RDNO	UNet
Regular	0	$7.37 \cdot 10^{-3}$	$3.16 \cdot 10^{-1}$
Uniform	$3.06 \cdot 10^{-2}$	$4.47 \cdot 10^{-1}$	$5.57 \cdot 10^{-1}$
Cluster	$2.85 \cdot 10^{-2}$	$2.93 \cdot 10^{-1}$	$5.52 \cdot 10^{-1}$

One can observe that RDNO predictions align more closely with the SAV scheme regarding dendrite morphology and dimensionless temperature distribution. While all methods generally remain consistent with the SAV reference, discrepancies are observed in local morphological features and temperature overshoot, most remarkably during grain merging. The underlying reasons for these differences require further investigation.

3.3. Refinement with DeepRitzSplit as a preconditioner

Since tip velocity depends on first-order derivatives and tip radius on second-order derivatives of the phase solution ϕ, DeepRitzSplit approximates tip velocity well but struggles with tip radius. The error shown in Table 4 likely stems from the fact that the variational formulation used in loss function only accounts for first-order terms.

Therefore, the neural operator can be retrained using the

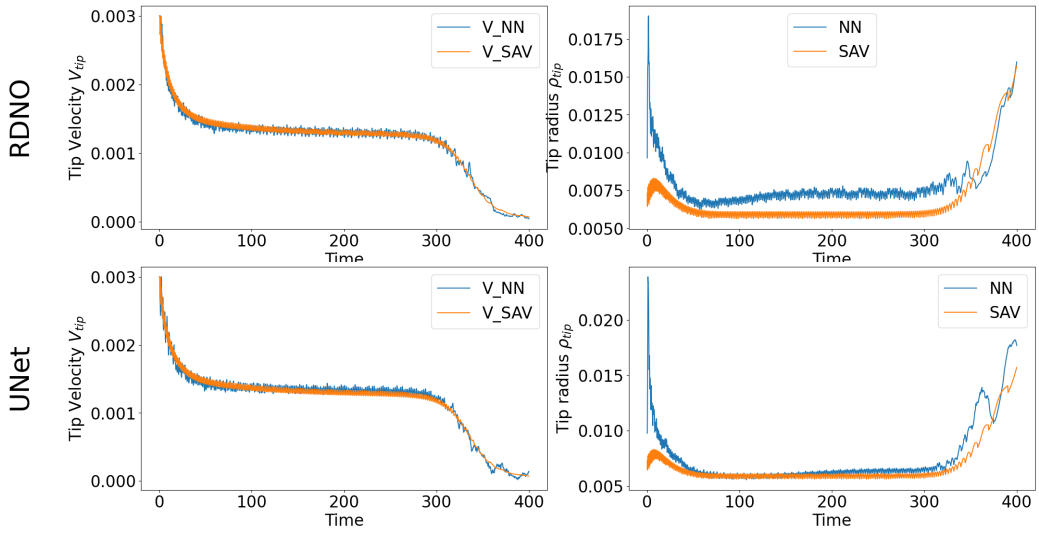

Figure 4. Evolutions of tip velocity V_{tip} and tip radius ρ_{tip} using RDNO and UNet trained with DeepRitzSplit and scheme residuals (24), compared to the numerical reference (SAV) for simulations with a single grain and the anisotropic strength $\sigma = 0.05$.

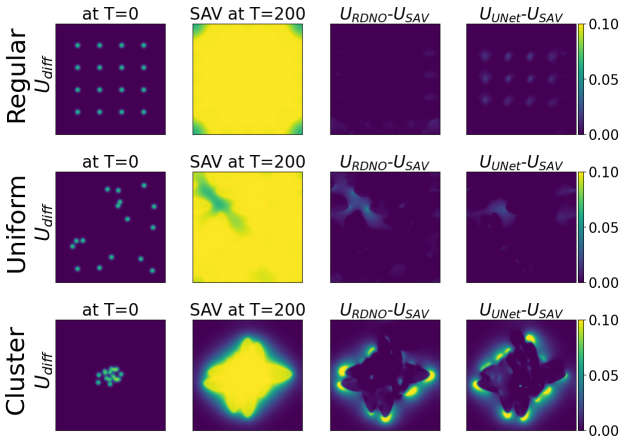

Figure 5. Predictions of multi-grain growth using UNet and RDNO across three spatial arrangements for anisotropy strength $\sigma = 0.05$.

splitting scheme residuals (16) after the DeepRitzSplit training. We define the second loss function:

$$\mathcal{L}_s(\phi, \phi_n, U_n) := \left\| \tau \frac{\phi - \phi_n}{\Delta t} + \nabla_\phi E_1(\phi, U_n) + \nabla_\phi E_2(\phi_n, U_n) \right\|_2^2 \tag{24}$$

However, due to the high nonlinearity of the loss (24), the training convergence exceeds two weeks, compared to a few hours for DeepRitzSplit. This refinement significantly improves the tip radius precision. as shown in Table 6 and Fig. 4.

We observe that if the neural operator is trained directly using the scheme residuals, the training hardly converges and the loss remains unstable. This observation motivates thus the use of DeepRitzSplit as a preconditioner, providing a "warm-start" that enables convergence within a reasonable timeframe.

Table 6. Dendrite tip velocity and radius of curvature during steady-state growth. Neural operators are pre-trained via DeepRitzSplit and refined using the scheme residuals (24). V_{tip} and ρ_{tip} are averaged $T = 400$ for $\sigma = 0.01$ and $T = 200$ for $\sigma = 0.05$.

σ	Metric	Reference		DeepRitzSplit		Rel. Error (%)	
		Analyt.	SAV	RDNO	UNet	RDNO	UNet
0.01	V_{tip} $(\cdot 10^{-4})$	n/a	5.45	5.28	5.28	3.4	3.2
	ρ_{tip} $(\cdot 10^{-2})$	n/a	7.20	7.49	7.11	4.1	1.2
	Pe $(\cdot 10^{-1})$	0.45	3.14	3.16	3.00	0.8	4.3
0.05	V_{tip} $(\cdot 10^{-3})$	n/a	1.29	1.29	1.34	0.3	3.4
	ρ_{tip} $(\cdot 10^{-2})$	n/a	0.59	0.71	0.62	19.3	4.7
	Pe $(\cdot 10^{-2})$	4.48	6.12	7.30	6.65	19.4	8.8

4. Conclusion and Perspective

We incorporated the convex-concave splitting scheme into the Deep Ritz method for gradient flows. Via the phase-field models for dendritic growth, we demonstrated that energy dissipation is preserved and that our approach offers better generalization over conventional data-driven methods. We show that physical properties such as solid fractions and tip velocities are accurately captured. Our approach also serves as a preconditioner method to further enhance morphological properties such as tip curvature. In multi-grain simulations, our proposed RDNO aligns more closely with the numerical scheme than the U-Nets. Further investigation for morphological properties such as tip curvature is required.

As a result, our approach paired with the RDNO can serve as a surrogate for semi-implicit splitting schemes, providing an interpretable alternative to solvers that are difficult to implement classically.

Acknowledgements

This work was supported by the Lorraine University of Excellence through the interdisciplinary project TransPINNO. Experiments presented in this paper were carried out using the Grid'5000 testbed, supported by a scientific interest group hosted by Inria and including CNRS, RENATER and several Universities as well as other organizations (see https://www.grid5000.fr).

Impact Statement

The goal of this paper is to improve neural operator frameworks for modeling complex physical systems. While the long-term application of these models in materials science may have potential societal consequences, the current work focuses on fundamental methodology and does not present ethical concerns.

References

Anandkumar, A., Azizzadenesheli, K., Bhattacharya, K., Kovachki, N., Li, Z., Liu, B., and Stuart, A. Neural operator: Graph kernel network for partial differential equations. In *ICLR 2020 workshop on integration of deep neural models and differential equations*, 2020.

Barbieri, A. and Langer, J. S. Predictions of dendritic growth rates in the linearized solvability theory. *Physical Review A*, 39(10):5314–5325, 1989. ISSN 10502947. doi: 10.1103/PhysRevA.39.5314.

Bhattacharya, K., Hosseini, B., Kovachki, N. B., and Stuart, A. M. Model reduction and neural networks for parametric pdes. *The SMAI journal of computational mathematics*, 7:121–157, 2021.

Bretin, E., Denis, R., Masnou, S., and Terii, G. Learning phase field mean curvature flows with neural networks. *Journal of Computational Physics*, 470:111579, 2022.

Choi, J. Y., Xue, T., Liao, S., and Cao, J. Accelerating phase-field simulation of three-dimensional microstructure evolution in laser powder bed fusion with composable machine learning predictions. *Additive Manufacturing*, 79:103938, 2024.

Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., and Piccialli, F. Scientific machine learning through physics–informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3):88, 2022.

Dantzig, J. A. and Rappaz, M. *Solidification: -Revised and Expanded*. EPFL press, 2016.

Eyre, D. J. Unconditionally gradient stable time marching the cahn-hilliard equation. *MRS online proceedings library (OPL)*, 529:39, 1998.

Fix, G. J. Phase field models for free boundary problems. In Fasano, A. and Primicerio, M. (eds.), *Free Boundary Problems: Theory and Applications*, volume II of *Research Notes in Mathematics*. Pitman Advanced Publishing Program, Boston, 1983. ISBN 978-0-273-08589-8.

Geng, Y., Teng, Y., Wang, Z., and Ju, L. A deep learning method for the dynamics of classic and conservative allen-cahn equations based on fully-discrete operators. *Journal of Computational Physics*, 496:112589, 2024.

Guo, Y., Azaïez, M., and Xu, C. An efficient numerical method for the anisotropic phase field dendritic crystal growth model. *Communications in Nonlinear Science and Numerical Simulation*, 131:107858, 2024.

Ivantsov, G. P. Temperature field around a spherical, cylindrical and needle-like crystal growing in a supercooled melt. *Doklady Akademii Nauk SSSR*, 58(4):567–569, 1947.

Karma, A. and Rappel, W.-J. Quantitative phase-field modeling of dendritic growth in two and three dimensions. *Physical review E*, 57(4):4323, 1998.

Kütük, M. and Yücel, H. Energy dissipation preserving physics informed neural network for allen-cahn equations. *arXiv preprint arXiv:2411.08760*, 2024.

Li, M., Azaiez, M., and Xu, C. New efficient time-stepping schemes for the anisotropic phase-field dendritic crystal growth model. *Computers & Mathematics with Applications*, 109:204–215, 2022.

Li, W., Bazant, M. Z., and Zhu, J. Phase-field deeponet: Physics-informed deep operator neural network for fast simulations of pattern formation governed by gradient flows of free-energy functionals. *Computer Methods in Applied Mechanics and Engineering*, 416:116299, 2023.

Li, Z.-Y., Kovachki, N. B., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A. M., and Anandkumar, A. Fourier neural operator for parametric partial differential equations. *ArXiv*, abs/2010.08895, 2020. URL https://api.semanticscholar.org/CorpusID:224705257.

Lu, L., Jin, P., and Karniadakis, G. E. Deeponet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators. *arXiv preprint arXiv:1910.03193*, 2019.

Mattey, R. and Ghosh, S. A novel sequential method to train physics informed neural networks for allen cahn and cahn hilliard equations. *Computer Methods in Applied Mechanics and Engineering*, 390:114474, 2022.

Miranville, A. On an anisotropic allen-cahn system. *Cubo (Temuco)*, 17(2):73–88, 2015.

Modica, L. Il limite nella γ-convergenza di una famiglia di funzionali ellittici. *(Italian) Boll. Un. Mat. Ital. A (5)*, 14 (3):526–529, 1977.

Oommen, V., Shukla, K., Desai, S., Dingreville, R., and Karniadakis, G. E. Rethinking materials simulations: Blending direct numerical simulations with neural operators. *npj Computational Materials*, 10(1):145, 2024.

Patel, R. G., Trask, N. A., Wood, M. A., and Cyr, E. C. A physics-informed operator regression framework for extracting data-driven continuum models. *Computer Methods in Applied Mechanics and Engineering*, 373:113500, 2021.

Plapp, M. Unified derivation of phase-field models for alloy solidification from a grand-potential functional. *Physical Review E*, 84(3):031601, September 2011. ISSN 1539-3755. doi: 10.1103/PhysRevE.84.031601.

Raissi, M., Perdikaris, P., and Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

Raonić, B., Molinaro, R., De Ryck, T., Rohner, T., Bartolucci, F., Alaifari, R., Mishra, S., and de Bézenac, E. Convolutional neural operators for robust and accurate learning of pdes. *Advances in Neural Information Processing Systems*, 36:77187–77200, 2023.

Ronneberger, O., Fischer, P., and Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pp. 234–241. Springer, 2015.

Takaki, T. Large-scale phase-field simulations for dendrite growth: A review on current status and future perspective. *IOP Conference Series: Materials Science and Engineering*, 1274(1):012009, January 2023. ISSN 1757-8981, 1757-899X. doi: 10.1088/1757-899x/1274/1/012009.

Tourret, D., Liu, H., and LLorca, J. Phase-field modeling of microstructure evolution: Recent applications, perspectives and challenges. *Progress in Materials Science*, 123(May 2020):100810, 2022. ISSN 00796425. doi: 10.1016/j.pmatsci.2021.100810.

Tseng, B.-Y., Guo, C.-W. C., Chien, Y.-C., Wang, J.-P., and Yu, C.-H. Deep learning model to predict ice crystal growth. *Advanced Science*, 10(21):2207731, 2023.

Wight, C. L. and Zhao, J. Solving allen-cahn and cahn-hilliard equations using the adaptive physics informed neural networks. *arXiv preprint arXiv:2007.04542*, 2020.

Yamazaki, Y., Harandi, A., Muramatsu, M., Viardin, A., Apel, M., Brepols, T., Reese, S., and Rezaei, S. A finite element-based physics-informed operator learning framework for spatiotemporal partial differential equations on arbitrary domains. *Engineering with Computers*, 41(1): 1–29, 2025.

Yu, B. et al. The deep ritz method: a deep learning-based numerical algorithm for solving variational problems. *Communications in Mathematics and Statistics*, 6(1):1–12, 2018.

Zaitzeff, A., Esedoḡlu, S., and Garikipati, K. High order, semi-implicit, energy stable schemes for gradient flows. *Journal of Computational Physics*, 447:110688, 2021.

Zhang, J., Zhang, S., Shen, J., and Lin, G. Energy-dissipative evolutionary deep operator neural networks. *Journal of Computational Physics*, 498:112638, 2024.

A. Phase-field models for Dendritic Growth: Technical details and Visualizations

A.1. SAV based method for anisotropic dendritic growth

The SAV method is a numerical scheme designed to handle gradient flows and energy-based PDEs while preserving their inherent energy dissipation properties. One of its main advantages is its *unconditional energy stability*, which allows for larger time steps without sacrificing the physical accuracy of the simulation. Additionally, by reformulating complex nonlinear terms using an auxiliary scalar variable, the SAV method simplifies the numerical treatment and enables the use of efficient implicit time-stepping schemes. However, the introduction of auxiliary variables may increase the computational time, and the method's generalization to highly nonlinear systems can be nontrivial.

Initially proposed in Li et al. (2022) and later improved in Guo et al. (2024), a SAV based splitting method was introduced for the anisotropic phase-field dendritic crystal growth model. The core idea is to first consider the following decoupling of the free energy E:

$$E(\phi, U) = \tilde{E}_1(\phi, U) + \tilde{E}_2(\phi, U)$$

where

$$\tilde{E}_1(\phi, U) = \int_\Omega \left(\frac{1}{2}\alpha|\nabla\phi|^2 + \beta\frac{\phi^2}{2\epsilon^2} \right) dx$$

with $\alpha, \beta > 0$ and

$$\tilde{E}_2(\phi, U) = E(\phi, U) - E_1(\phi, U).$$

The corresponding splitting numerical scheme is therefore given by

$$\begin{cases} \tau\dfrac{\phi_{n+1} - \phi_n}{dt} = -\nabla_\phi \tilde{E}_1(\phi_{n+1}, U_n) - \nabla_\phi \tilde{E}_2(\phi_n, U_n) \\ \qquad\qquad = \alpha\Delta\phi_{n+1} - \beta\epsilon^{-2}\phi_{n+1} - \nabla_\phi \tilde{E}_2(\phi_n, U_n) \\ \dfrac{U_{n+1} - U_n}{dt} = D\Delta U_{n+1} + Kh'(\phi_{n+1})\dfrac{\phi_{n+1} - \phi_n}{dt} \end{cases} \tag{25}$$

However, regardless of the numerical efficiency of solving the system (25), such a numerical scheme is *not* unconditionally stable since the concavity of $\tilde{E}_2$ is not ensured in general. In order to guarantee energy stability, Guo et al. reformulate the scheme (25) by adding a scalar auxiliary variable q_n as an energy relaxation term, see Eqs. (26)-(27) below. Despite requiring more computational time compared to Eq. (25), this approach enables the use of efficient implicit time-stepping schemes while ensuring energy dissipation as described in Eq. (2). Unlike traditional numerical methods that may struggle with stability or require small time steps for accuracy, the SAV method allows for larger time steps without compromising stability.

The SAV method introduced in Guo et al. (2024) consists of two steps. First, for (ϕ_n, U_n, q_n), we seek for a solution $(\bar{\phi}_{n+1}, \bar{U}_{n+1}, \bar{q}_{+1})$ of,

$$\begin{cases} \tau\dfrac{\bar{\phi}_{n+1} - \phi_n}{\mathrm{dt}} = -\nabla_\phi \tilde{E}_1(\bar{\phi}_{n+1}, U_n) - \nabla_\phi \tilde{E}_2(\phi_n, U_n) \\ \qquad\qquad = \alpha\Delta\bar{\phi}_{n+1} - \beta\epsilon^{-2}\bar{\phi}_{n+1} - \nabla_\phi \tilde{E}_2(\phi_n, U_n) \\ \dfrac{\bar{U}_{n+1} - U_n}{dt} = D\Delta\bar{U}_{n+1} + Kh'(\bar{\phi}_{n+1})\dfrac{\bar{\phi}_{n+1} - \phi_n}{\mathrm{dt}} \\ \dfrac{(\bar{q}_{n+1} - q_n)}{\mathrm{dt}} = \xi_{n+1} dE_t(\bar{\phi}_{n+1}, \bar{U}_{n+1}) \end{cases} \tag{26}$$

with $q_0 = E(\phi_0, U_0)$ and $\xi_{n+1} = \frac{\bar{q}_{n+1}}{E(\bar{\phi}_{n+1}, \bar{U}_{n+1})}$. And then we set $\eta_{n+1} = 1 - (1 - \xi_{n+1})^2$ and

$$\begin{cases} \phi_{n+1} = \eta_{n+1}\bar{\phi}_{n+1} \\ U_{n+1} = \eta_{n+1}\bar{U}_{n+1} \\ q_{n+1} = \zeta_{n+1}\bar{q}_{n+1} + (1 - \zeta_{n+1})E(\phi_{n+1}, U_{n+1}) \end{cases} \tag{27}$$

where $\zeta_{n+1} \in [0, 1]$ is chosen such that

$$\frac{(q_{n+1} - \bar{q}_{n+1})}{\mathrm{dt}} \leq -\xi_{n+1} dE_t(\bar{\phi}_{n+1}, \bar{U}_{n+1}) \,. \tag{28}$$

With the help of the auxiliary variable q_n, it is proven in Guo et al. (2024) that the numerical scheme (26)-(27) is energetically stable.

A.2. DeepRitzSplit to anisotropic dendritic growth simulations

Dendritic growth simulations for weak anisotropic strength $\sigma = 0.05$

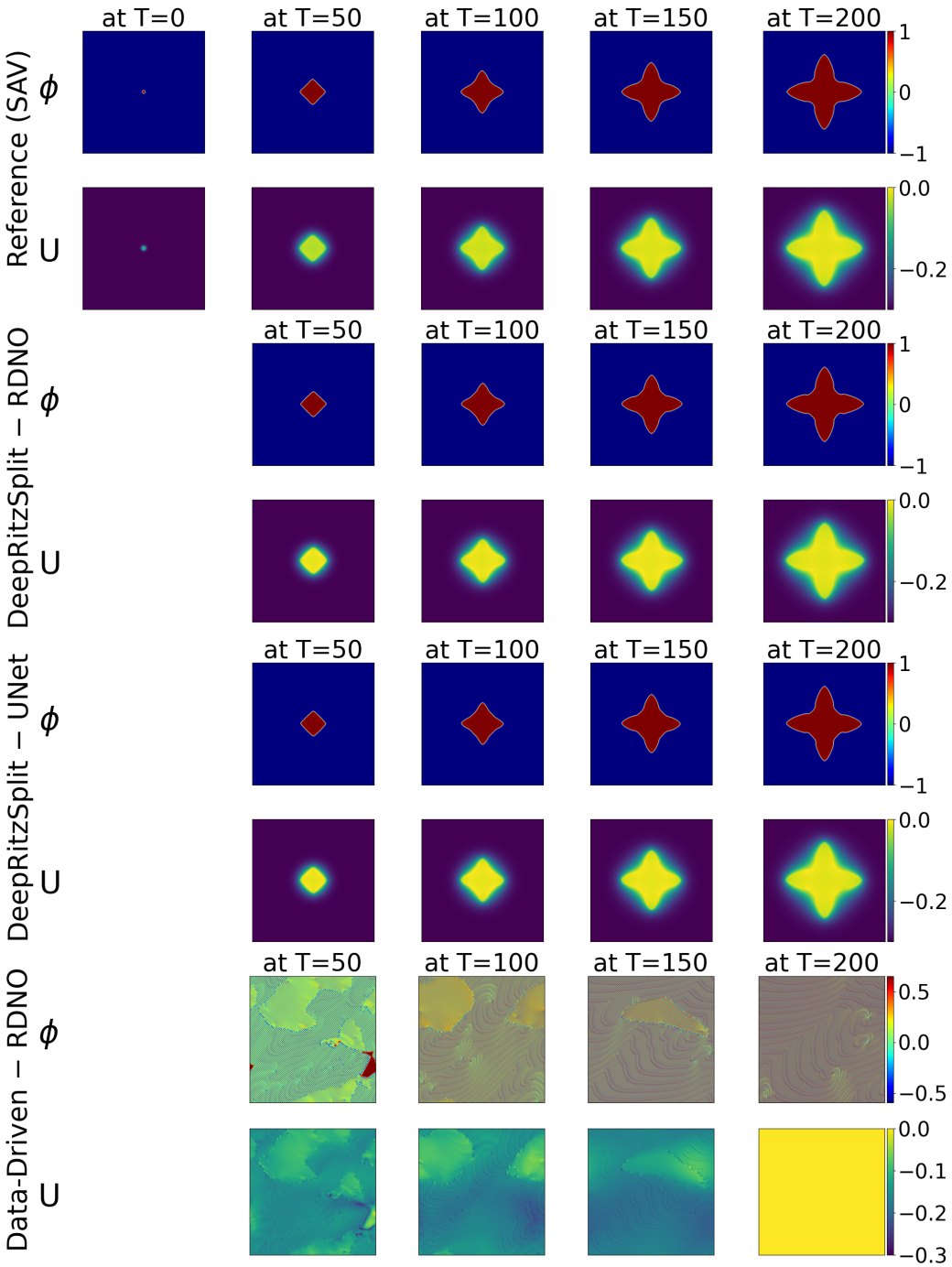

Figure 6. Dendritic growth predictions with anisotropic strength $\sigma = 0.05$

Dendritic growth simulations for weak anisotropic strength $\sigma = 0.01$

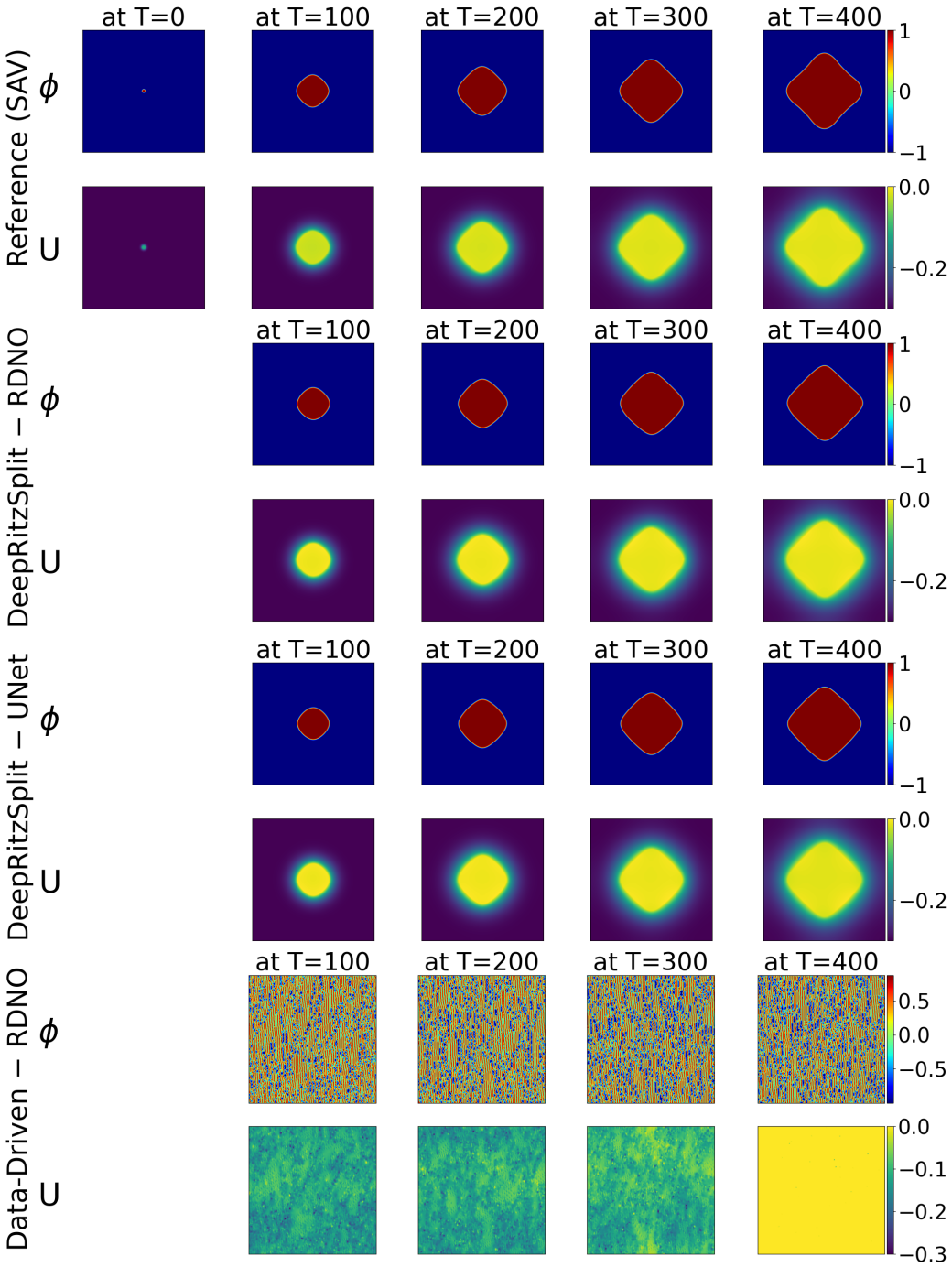

Figure 7. Dendritic growth predictions with anisotropic strength $\sigma = 0.01$

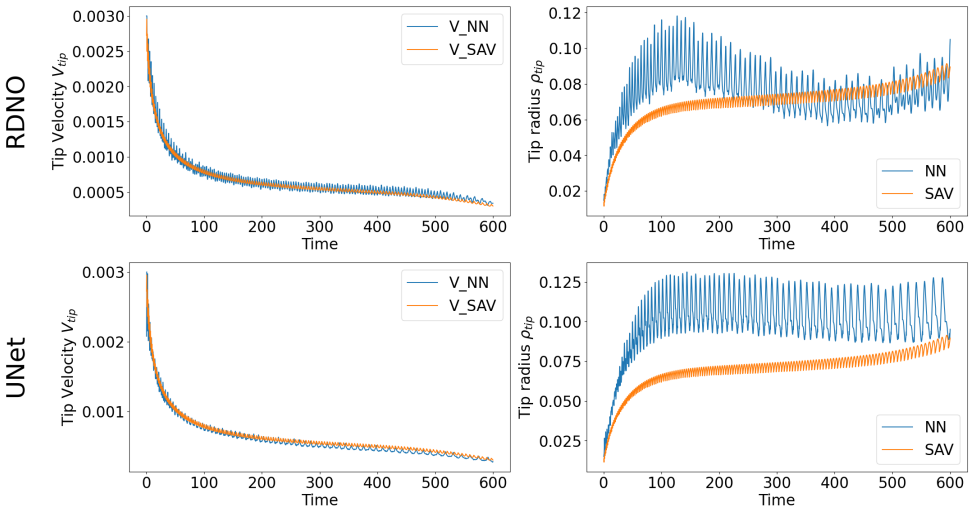

Figure 8. Evolution of tip velocity Vtip and radius ρ_{tip} for RDNO (top) and UNet (bottom) compared to the SAV numerical reference for $\sigma = 0.01$.

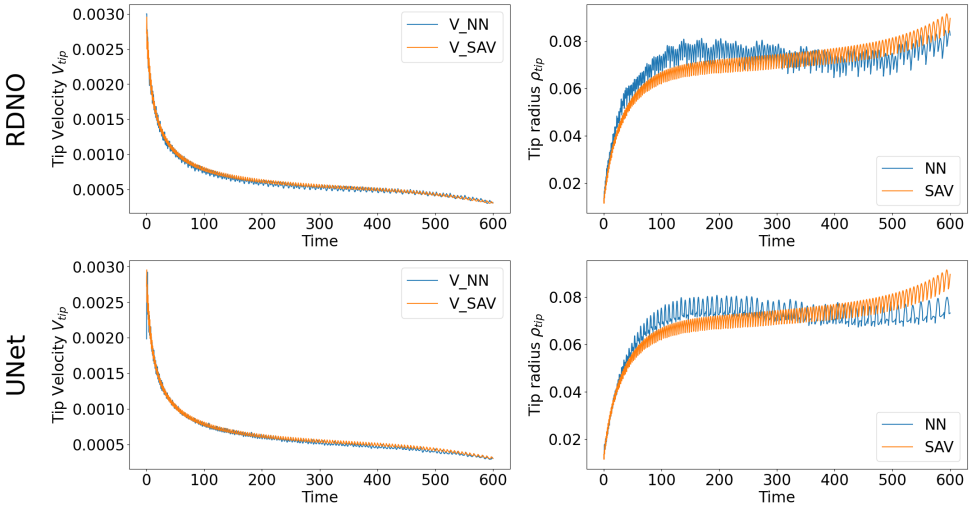

Figure 9. Evolutions of tip velocity V_{tip} and tip radius ρ_{tip} using RDNO and UNet trained with DeepRitzSplit and scheme residuals (24), compared to the numerical reference (SAV) for simulations with a single grain and the anisotropic strength $\sigma = 0.01$.

A.3. Application to multi-grain simulations

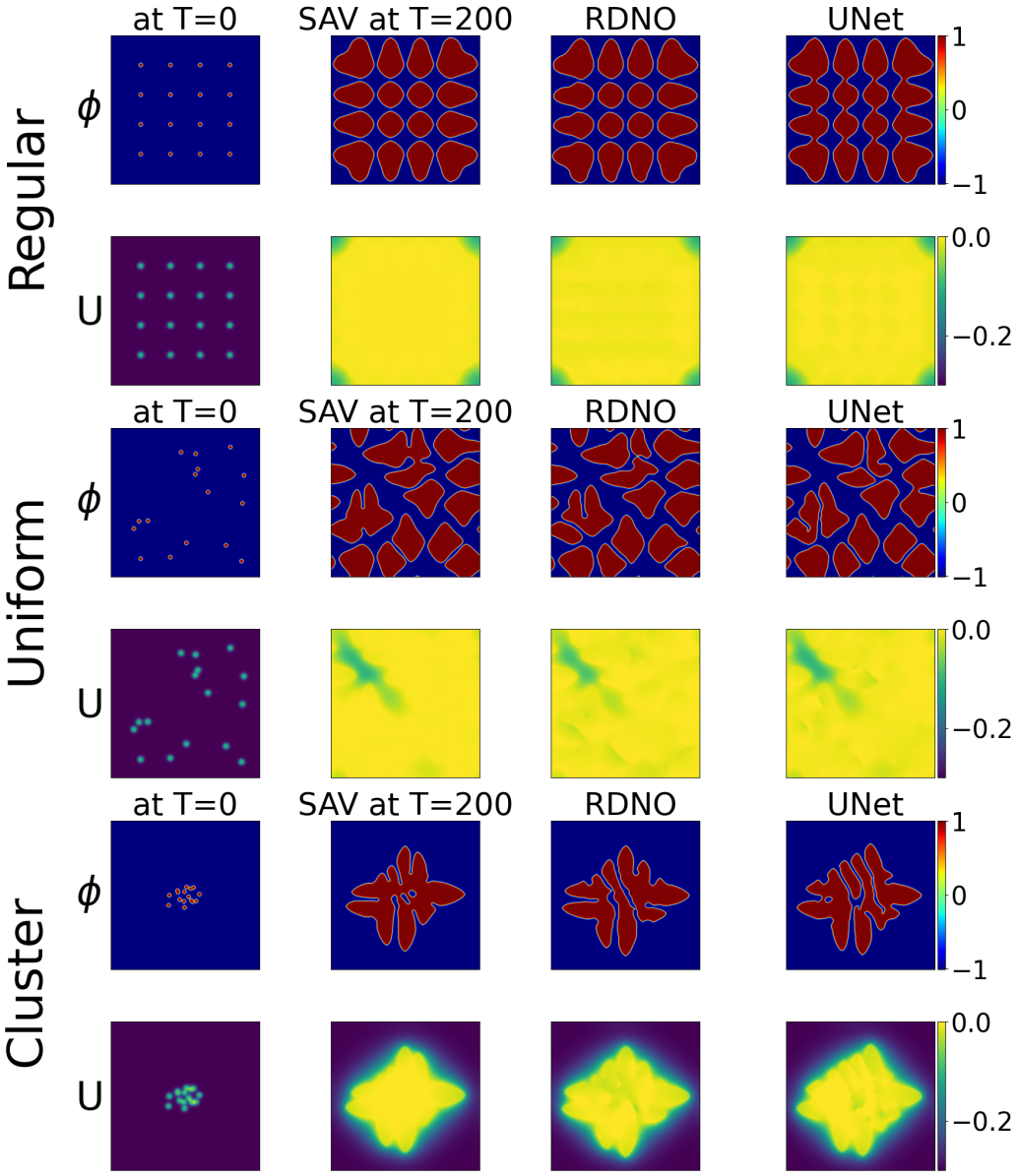

Figure 10. Predictions of multi-grains with UNet and RDNO over 3 different spatial arrangements for the anisotropic strength $\sigma = 0.05$.