

General numerical framework to derive structure preserving reduced order models for thermodynamically consistent reversible-irreversible PDEs

Zengyan Zhang, Jia Zhao *

Department of Mathematics and Statistics, Binghamton University, Binghamton, NY, USA

ARTICLE INFO

Keywords:

Model order reduction

Phase field

Thermodynamically consistent

Structure preserving

ABSTRACT

In this paper, we propose a general numerical framework to derive structure-preserving reduced-order models for thermodynamically consistent PDEs. Our numerical framework has two primary features: (a) a systematic way to extract reduced-order models for thermodynamically consistent PDE systems while maintaining their inherent thermodynamic principles and (b) a general process to derive accurate, efficient, and structure-preserving numerical algorithms to solve these reduced-order models. The platform's generality extends to various PDE systems governed by embedded thermodynamic laws, offering a unique approach from several perspectives. First, it utilizes the generalized Onsager principle to transform the thermodynamically consistent PDE system into an equivalent form, where the free energy of the transformed system takes a quadratic form in terms of the state variables. This transformation is known as energy quadratization (EQ). Through EQ, we gain a novel perspective on deriving reduced-order models that continue to respect the energy dissipation law. Secondly, our proposed numerical approach automatically provides algorithms to discretize these reduced-order models. The proposed algorithms are always linear, easy to implement and solve, and uniquely solvable. Furthermore, these algorithms inherently ensure the thermodynamic laws. Our platform offers a distinctive approach for deriving structure-preserving reduced-order models for a wide range of PDE systems with underlying thermodynamic principles.

1. Introduction

Partial differential equation (PDE) models play an important role in modeling physical phenomena in various scientific and engineering fields. They have been used to model complex systems, multidimensional phenomena, and various interdisciplinary problems. As a coarse-grained modeling approach, ensuring these PDE models remain consistent with underlying thermodynamic laws is essential, particularly when reversible and irreversible transitions occur, making the models both physically plausible and providing reliable predictions. PDEs that respect the thermodynamic laws are known as thermodynamically consistent PDEs. In particular, when temperature fluctuation can be ignored, the thermodynamically consistent PDE system will respect laws of free energy dissipation, i.e., the free energy for the system will not increase in time. These models have numerous applications in various science and engineering fields, along with interdisciplinary applications. There are many broadly used PDE models that fall into the

* Corresponding author.

E-mail address: jiazhao@binghamton.edu (J. Zhao).

thermodynamically consistent category, to name a few: the Allen-Cahn equation [2], the Cahn-Hilliard equation [8], the phase field crystal equation [16,19], the Navier-Stokes equation, and the Ericksen-Leslie model for liquid crystals [6,31].

The thermodynamically consistent PDEs are usually coupled systems with nonlinearity. Analytical approaches often fall short, making the numerical approach necessary. Unfortunately, they are generally hard to solve numerically due to the stiffness in nonlinear terms. Additionally, any numerical deviation from underlying thermodynamic laws can result in physically implausible solutions, against the purpose of the original modeling intention [18]. Therefore, the numerical analysis community has been actively deriving numerical algorithms that preserve the thermodynamic structure, namely structure-preserving numerical algorithms. In particular, numerical algorithms that respect the free energy dissipation laws are usually named energy-stable algorithms [17]. If numerical energy stability is independent of the choices of the time step size, they are called unconditionally energy stable [19]. In this context, many research results have been achieved recently, including the stabilized approach [36], the convex splitting method [17,24,40,45], the energy quadratization method [46,48], the scalar auxiliary variable approach [26,34,35], the SVM method [20] and Lagrangian multiplier approach [12].

Moreover, solving phase field models can be computationally intensive due to spatial and temporal multi scales that have to be resolved, especially when long-time dynamics are desired. A naive way to tackle the computational complexity of thermodynamically consistent models is to employ the reduced-order model (ROM) or model order reduction (MOR) techniques, which aim to reduce the spatial-temporal complexity. ROMs have gained significant attention in recent years for their capability to simplify complex PDE systems with reasonable approximation errors, enabling faster computational speed and more efficient simulations. The way for ROM to save computational cost is by projecting the problem from a high-dimensional system into a much lower-dimensional subspace while maintaining a small approximation error. In particular, the ROMs are usually required to conserve the properties and characteristics of the full-order model, i.e., the original model. The broadly used MOR techniques include proper orthogonal decomposition (POD) methods [30,43], reduced basis methods [32], balancing methods [5], and nonlinear manifold methods [13], as well as various projection-based reductions [9]. Among these approaches, the POD Galerkin method will be the major focus of this paper. The POD method computes an optimal subspace to fit the empirical data. It was first introduced to study turbulence by the fluid-dynamics community [30] as a way to decompose the random vector field representing turbulent fluid motion into a set of deterministic functions that each capture some portion of the total fluctuating kinetic energy in the flow [43]. The POD method has been widely used in computational fluid dynamics and structure analysis ever since. In general, the empirical data is generated, and POD modes are computed during the offline stage. In the online stage, the POD-ROM is solved in real time instead of the full model. Some special techniques are needed to handle the nonlinearity to ensure the ROM is fully independent of the full model's dimension. These techniques include the trajectory piecewise linear approximation [33], missing point estimation [3], gappy POD [44], empirical interpolation [4], and the discrete empirical interpolation method (DEIM) [10].

When it comes to developing POD-ROMs for thermodynamically consistent PDE systems, the challenge lies in not only maintaining the reduced computational complexity and controlled projection error to the original PDEs but also respecting their embedded thermodynamic laws. Unfortunately, a direct application of the classical POD-ROM strategies to the thermodynamically consistent PDEs will usually destroy the thermodynamic structure of the full-order model, which is problematic since these numerical solutions from the POD-ROM will violate the thermodynamic laws. Particular attention is needed to adjust the classical POD-ROM strategies for thermodynamically consistent PDEs. Over the years, some seminal ideas have been proposed to design structure-preserving ROMs for the Hamiltonian (reversible) systems, but there is still little work on dissipative (irreversible) systems. This motivates our research in this paper. Here, we provide a brief summary of the existing work in the literature. A general structure-preserving reduced-order modeling approach for gradient systems is proposed in [47]. The authors use the symmetric interior penalty discontinuous Galerkin (SIPG) method for spatial discretization and the average vector field (AVF) method for temporal discretization. The nonlinear terms are taken care of using the discrete empirical interpolation method (DEIM). The major drawback of this approach is that the resulting system is fully implicit and nonlinear. A similar idea is applied to design structure-preserving integration and model order reduction of the skew-gradient reaction-diffusion systems [27]. A structure-preserving Galerkin POD reduced-order modeling for the Hamiltonian systems is introduced in [21]. The major idea is to introduce a modified skew-symmetric operator in the reduced-order model such that the Hamiltonian is preserved. A further extension introduced in [42] overcomes the high computational complexity with non-polynomial nonlinearities in the Hamiltonian by the discrete empirical interpolation method (DEIM). There is also some work in designing reduced-order models for the dissipative systems, in particular, the phase field models or gradient flow models. For instance, the authors introduce a reduced-order model for the Allen-Cahn equation by embracing the idea of the scalar auxiliary variable (SAV) method in [50]. An alternative approach to develop the reduced-order model for the Allen-Cahn equation is introduced in [29], where the authors use the stabilized semi-implicit scheme. A rigorous numerical analysis is also provided. A finite difference approach in space and IMEX Runge-Kutta approach in time is introduced for developing reduced-order models of the Allen-Cahn equation in [37]. Similar techniques of [21] are exploited to develop structure-preserving reduced-order modeling for the Korteweg-de Vries equation in [38]. A method is introduced for the dissipative Hamiltonian systems, but it only applies to dissipative Hamiltonian systems with a quadratic Hamiltonian [1]. Meanwhile, there are several other research directions. The authors consider the inclusion of spatial adaptivity for the snapshot computation in the offline phase of model order reduction [23]. Using deep learning techniques, particularly deep neural networks, for model order reductions has also been considered [14].

In this paper, we introduce a numerical platform that can systematically derive ROMs for reversible-irreversible thermodynamically consistent PDE models by embracing several novel techniques. First of all, we utilize the energy quadratization (EQ) technique [46,49] to reformulate the generic thermodynamically consistent PDE models into the quadratic Onsager form. By this equivalent model reformulation, the reversible-irreversible thermodynamic structures of the original PDEs are fully disclosed. Then, thanks to the quadratic structure in our reformulated system, we are able to slightly modify the classical POD-ROM, which was inspired by

[15], to derive structure-preserving POD-ROM. There are several unique advantages of our POD-ROM framework compared with existing results in the literature: (1) First of all, our approach is general in that the POD-ROM framework applies to many existing thermodynamically consistent models and respects their thermodynamic structures after model order reduction; (2) Secondly, the structure-preserving numerical integration of our POD-ROM framework is linear, making it easy to implement and cheap to compute; (3) Thirdly, given the linearized nature of the POD-ROM framework, all the nonlinear terms in the ROM can be treated explicitly (using many existing techniques such as DEIM) while still preserving the thermodynamic structure. These properties make our POD-ROM framework thermodynamically consistent and widely applicable.

The rest of this paper is structured as follows. Section 2 provides a comprehensive view of reformulating the thermodynamically consistent PDE models via the generalized Onsager principle, particularly with the Onsager triplet. This allows us to reformulate the system using the EQ method, resulting in a system where free energy is expressed in a quadratic form in terms of the state variables. In Section 3, we elaborate on the procedure of designing structure-preserving reduced-order models in the context of the EQ reformulation. Afterward, we introduce the numerical platforms to develop structure-preserving numerical algorithms for the reduced-order models in Section 4. In Section 5, we present a range of numerical examples for specific thermodynamically consistent PDE models, demonstrating the effectiveness of the proposed computational framework. We then wrap this paper with brief concluding remarks and future work in the last section.

2. Thermodynamically consistent reversible-irreversible PDE models based on the generalized Onsager principle

2.1. Generalized Onsager principle

Consider a domain Ω , and denote the thermodynamic variable ϕ . We recall the generalized Onsager principle. It consists of three key ingredients: the state or thermodynamic variable ϕ , the free energy \mathcal{E} , and the kinetic equation dictated by a mobility matrix (or operator) \mathcal{G} . In this paper, we name it

the Onsager triplet: $(\phi, \mathcal{G}, \mathcal{E})$. (2.1)

The kinetic equation, deriving from the Onsager linear response theory, is given by

$$\partial_t \phi(\mathbf{x}, t) = -\mathcal{G} \frac{\delta \mathcal{E}}{\delta \phi} \text{ in } \Omega, \quad (2.2a)$$

$$\mathcal{B}(\phi(\mathbf{x}, t)) = g(\mathbf{x}, t), \text{ on } \partial\Omega, \quad (2.2b)$$

where \mathcal{B} is an operator for the boundary condition, and \mathcal{G} is the mobility operator that contains two parts:

$$\mathcal{G} = \mathcal{G}_a + \mathcal{G}_s. \quad (2.3)$$

\mathcal{G}_s is symmetric and positive semi-definite to ensure thermodynamically consistency, \mathcal{G}_a is skew-symmetric, and $\frac{\delta \mathcal{E}}{\delta \phi}$ is the variational derivative of \mathcal{E} , known as the chemical potential. Then, the triplet $(\phi, \mathcal{G}, \mathcal{E})$ uniquely defines a thermodynamically consistent model. One intrinsic property of (2.2) owing to the thermodynamical consistency is the energy dissipation law

$$\frac{d\mathcal{E}}{dt} = \dot{\mathcal{E}}_{bulk} + \dot{\mathcal{E}}_{surf}, \quad (2.4a)$$

$$\dot{\mathcal{E}}_{bulk} = -\left(\frac{\delta \mathcal{E}}{\delta \phi}, \mathcal{G}_s \frac{\delta \mathcal{E}}{\delta \phi} \right) \leq 0, \quad (2.4b)$$

$$\left(\frac{\delta \mathcal{E}}{\delta \phi}, \mathcal{G}_a \frac{\delta \mathcal{E}}{\delta \phi} \right) = 0, \quad \dot{\mathcal{E}}_{surf} = \int_{\partial\Omega} g_b ds, \quad (2.4c)$$

where \mathcal{E}_{bulk} denotes the bulk free energy and $\mathcal{E}_{surface}$ represents the surface free energy. Additionally, the inner product is defined by

$$(f, g) = \int_{\Omega} f g d\mathbf{x}, \quad \forall f, g \in L^2(\Omega),$$

and $\dot{\mathcal{E}}_{surf}$ results from the boundary contribution, and g_b is the boundary integrand. When $\mathcal{G}_a = 0$, (2.2) is a purely dissipative system. When $\mathcal{G}_s = 0$, it is a purely dispersive system. $\dot{\mathcal{E}}_{surf}$ vanishes only for suitable boundary conditions, which include periodic and certain physical boundary conditions. When the mass, momentum, and total energy conservation are present in hydrodynamic models, these conservation laws are viewed as constraints imposed on the hydrodynamic variables. Then, the energy dissipation rate will have to be calculated subject to the constraints.

2.2. Model reformulation with the energy quadratization (EQ) method

Now, we illustrate the idea of the energy quadratization method. Denote the total energy as

$$\mathcal{E}(\phi) = \int_{\Omega} e d\mathbf{x}, \quad (2.5)$$

with e the energy density function. We denote \mathcal{L}_0 as a positive semi-definite linear operator that can be separated from e . Introduce the auxiliary variable

$$q = \sqrt{2\left(e - \frac{1}{2}|\mathcal{L}_0^{\frac{1}{2}}\phi|^2 + \frac{A_0}{|\Omega|}\right)}, \quad (2.6)$$

where $A_0 > 0$ is a constant so that q is a well-defined real variable. Then, we rewrite the energy as

$$\mathcal{E}(\phi, q) = \frac{1}{2}(\phi, \mathcal{L}_0\phi) + \frac{1}{2}(q, q) - A_0. \quad (2.7)$$

With the EQ approach above, we transform the free energy density into a quadratic one by introducing an auxiliary variable to “remove” the nonlinear terms from the energy density. Assuming $q = q(\phi)$ and denoting $g(\phi) = \frac{\partial q}{\partial \phi}$, we reformulate (2.2) into an equivalent form

$$\partial_t \phi = -(\mathcal{G}_a + \mathcal{G}_s)(\mathcal{L}_0\phi + qg(\phi)), \quad (2.8a)$$

$$\partial_t q = g(\phi) : \partial_t \phi, \quad (2.8b)$$

$$q|_{t=0} = \sqrt{2\left(e - \frac{1}{2}|\mathcal{L}_0^{\frac{1}{2}}\phi|^2 + \frac{A_0}{|\Omega|}\right)} \Big|_{t=0}. \quad (2.8c)$$

Now, instead of dealing with (2.2) directly, we develop structure-preserving schemes for (2.8). The advantage of using model (2.8) over model (2.2) is that the energy density is transformed into a quadratic form in (2.8).

Denote $\Psi = \begin{bmatrix} \phi \\ q \end{bmatrix}$. We rewrite (2.8) into a vector form

$$\partial_t \Psi = -\mathcal{N}(\Psi)\mathcal{L}\Psi, \quad (2.9)$$

where $\mathcal{N}(\Psi)$ is the mobility operator, and \mathcal{L} is a linear operator.

$$\mathcal{N}(\Psi) = \mathcal{N}_s(\Psi) + \mathcal{N}_a(\Psi), \quad (2.10a)$$

$$\mathcal{N}_a(\Psi) = \mathcal{N}_0^* \mathcal{G}_a \mathcal{N}_0, \quad (2.10b)$$

$$\mathcal{N}_s(\Psi) = \mathcal{N}_0^* \mathcal{G}_s \mathcal{N}_0, \quad (2.10c)$$

where

$$\mathcal{N}_0 = [\mathbf{I} \quad g(\phi)], \quad \mathcal{L} = \begin{bmatrix} \mathcal{L}_0 & \\ & \mathbf{I} \end{bmatrix},$$

with the identity matrix denoted by \mathbf{I} , and \mathcal{N}_0^* is the adjoint operator of \mathcal{N}_0 . We name it the Onsager-Q form, where

$$\frac{d\mathcal{E}(\Psi)}{dt} = \left(\frac{\delta \mathcal{E}}{\delta \Psi} \frac{\partial \Psi}{\partial t}, 1 \right) = -\left(\mathcal{L}\Psi, \mathcal{N}(\Psi)\mathcal{L}\Psi \right) = -\left(\mathcal{N}_0^* \mathcal{L}\Psi, \mathcal{G}_s \mathcal{N}_0 \mathcal{L}\Psi \right) \leq 0, \quad (2.11)$$

where the energy of (2.7) is reformulated in a vector form as

$$\mathcal{E}(\Psi) = \frac{1}{2}(\Psi, \mathcal{L}\Psi) - A_0. \quad (2.12)$$

Note that the energy in (2.7) or (2.12) is quadratized so that one can easily derive linear and energy stable numerical schemes for the model.

Remark 2.1. The EQ approach is also applicable when $q = q(\phi, \nabla\phi)$ and denoting

$$g(\phi) = \frac{\partial q}{\partial \phi}, \quad \mathbf{G}(\phi) = \frac{\partial q}{\partial \nabla\phi}, \quad (2.13)$$

the kinetic equation (2.2) can be reformulated into an equivalent form

$$\partial_t \phi = -(\mathcal{G}_a + \mathcal{G}_s)(\mathcal{L}_0\phi + qg(\phi) - \nabla \cdot (q\mathbf{G}(\phi))), \quad (2.14a)$$

$$\partial_t q = g(\phi) : \partial_t \phi + \mathbf{G}(\nabla\phi) : \nabla \partial_t \phi, \quad (2.14b)$$

$$q|_{t=0} = \sqrt{2\left(e - \frac{1}{2}|\mathcal{L}_0^{\frac{1}{2}}\phi|^2 + \frac{A_0}{|\Omega|}\right)} \Big|_{t=0}. \quad (2.14c)$$

For the simplicity of notation, the approach presented in this paper derives structure-preserving ROMs for (2.8) and it is also applicable for (2.14).

3. Structure-preserving reduced-order models (ROMs)

In this section, we present the general structure-preserving framework for deriving reduced-order models (ROMs). We emphasize that the transformed model (2.8) is equivalent to (2.2). Thus, we focus on designing the ROMs for (2.8), which in turn is a good surrogate model for (2.2).

3.1. Model order reduction (MOR)

Consider the solution for (2.8) as

$$\Psi(\mathbf{x}, t) = \sum_{k=1}^{\infty} \mathbf{a}_k(t) \psi_k(\mathbf{x}),$$

where $\{\psi_k(\mathbf{x})\}_{k=1}^{\infty}$ are the spatial modes and $\{\mathbf{a}_k(t)\}_{k=1}^{\infty}$ are the corresponding time coefficients. For the ROM, we look for a low-dimensional projections

$$\Psi_r(\mathbf{x}, t) = \sum_{k=1}^r \mathbf{a}_k(t) \psi_k(\mathbf{x}), \quad (3.1)$$

such that it provides an accurate approximation to $\Psi(\mathbf{x}, t)$ and $\psi_k(\mathbf{x})$ are the optimal basis modes when r is small. In general, we would like the modes orthonormal, i.e.,

$$\int_{\Omega} \psi_i(\mathbf{x}) \psi_j(\mathbf{x}) d\mathbf{x} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Plugging the ROM solution of (3.1) into (2.9), we have

$$\sum_{k=1}^r \frac{d}{dt} \mathbf{a}_k(t) \psi_k(\mathbf{x}) = -\mathcal{N} \left(\sum_{k=1}^r \mathbf{a}_k(t) \psi_k(\mathbf{x}) \right) \mathcal{L} \left[\sum_{k=1}^r \mathbf{a}_k(t) \psi_k(\mathbf{x}) \right]. \quad (3.2)$$

It is clear that the ROM in (3.2) preserves the original structure. In fact, if we denote the reduced-order energy as

$$\mathcal{E}_r(\Psi_r(\mathbf{x}, t)) = \frac{1}{2} \left(\Psi_r(\mathbf{x}, t), \mathcal{L} \Psi_r(\mathbf{x}, t) \right) - A_0, \quad (3.3)$$

we can have the following energy dissipation law

$$\begin{aligned} \frac{d\mathcal{E}_r(\Psi_r(\mathbf{x}, t))}{dt} &= \left(\mathcal{L} \Psi_r(\mathbf{x}, t), \frac{d}{dt} \Psi_r(\mathbf{x}, t) \right) \\ &= \left(\mathcal{L} \left[\sum_{k=1}^r \mathbf{a}_k(t) \psi_k(\mathbf{x}) \right], \sum_{k=1}^r \frac{d}{dt} \mathbf{a}_k(t) \psi_k(\mathbf{x}) \right) \\ &= - \left(\mathcal{L} \left[\sum_{k=1}^r \mathbf{a}_k(t) \psi_k(\mathbf{x}) \right], \mathcal{N} \left(\sum_{k=1}^r \mathbf{a}_k(t) \psi_k(\mathbf{x}) \right) \mathcal{L} \left[\sum_{k=1}^r \mathbf{a}_k(t) \psi_k(\mathbf{x}) \right] \right) \\ &\leq 0. \end{aligned}$$

3.2. POD-Galerkin method

For structure-preserving spatial discretizations, either finite difference, finite elements, or spectral methods can be used. For simplicity, we assume periodic boundary conditions in the rest of this paper and use the pseudo-spectral method for space discretization. Then, we focus on the proper orthogonal decomposition (POD) for the selection of optimal spatial modes and the construction of structure-preserving ROMs in combination with Galerkin projection.

3.2.1. Pseudo-spectral spatial discretization

We consider a rectangular domain $\Omega = [l_x, r_x] \times [l_y, r_y]$, and denote $L_x = r_x - l_x$ and $L_y = r_y - l_y$. For Fourier spectral method, we discretize the domain into equally distanced rectangular meshes with $h_x = \frac{L_x}{N_x}$ and $h_y = \frac{L_y}{N_y}$, with N_x, N_y the number of meshes in x and y directions respectively, and h_x, h_y the corresponding mesh sizes. Then, we have the discrete coordinates

$$(x_m, y_n) = (l_x + m h_x, l_y + n h_y), \quad m = 0, 1, 2, \dots, N_x - 1, \quad n = 0, 1, 2, \dots, N_y - 1. \quad (3.4)$$

Given the domain is periodic, we have $x_{N_x} = x_0, y_{N_y} = y_0$. Assume N_x and N_y are even numbers with $N_x = 2K_x$ and $N_y = 2K_y$.

Denote ϕ_{ij} the numerical approximation for $\phi(x_i, y_j)$. Then, we have the discrete Fourier expansion in 2D as

$$\phi_{mn} = \frac{1}{N_x N_y} \sum_{k=-K_x+1}^{K_x} \sum_{l=-K_y+1}^{K_y} \hat{\phi}_{kl} \exp \left(2\pi i \left(k \frac{x_m}{L_x} + l \frac{y_n}{L_y} \right) \right) \quad (3.5)$$

and the corresponding Fourier inverse transform is given as

$$\hat{\phi}_{kl} = \sum_{m=0}^{N_x-1} \sum_{n=0}^{N_y-1} \phi_{mn} \exp\left(-2\pi i(k \frac{x_m}{L_x} + l \frac{y_n}{L_y})\right) \quad (3.6)$$

With the notations for the Fourier transform, we can calculate the first-order and second-order partial derivatives as

$$(\mathcal{D}_{N_x} \phi)_{ij} = \frac{1}{N_x N_y} \sum_{k=-K_x+1}^{K_x} \sum_{l=-K_y+1}^{K_y} \frac{2\pi k i}{L_x} \hat{\phi}_{k,l} \exp\left(2\pi i(kx_i + ly_j)\right) \quad (3.7)$$

and the second-order partial derivative is given as

$$(\mathcal{D}_{N_x}^2 \phi)_{ij} = \frac{1}{N_x N_y} \sum_{k=-K_x+1}^{K_x} \sum_{l=-K_y+1}^{K_y} \left(-\frac{4\pi^2 k^2}{L_x^2}\right) \hat{\phi}_{k,l} \exp\left(2\pi i(kx_i + ly_j)\right) \quad (3.8)$$

Similarly, we can define the differential operators for $(\mathcal{D}_{N_y} \phi)_{ij}$ and $(\mathcal{D}_{N_y}^2 \phi)_{ij}$. Then, we can introduce the discrete Laplacian, gradient, and divergence operators as

$$\begin{aligned} \Delta_N \phi &= (\mathcal{D}_{N_x}^2 + \mathcal{D}_{N_y}^2) \phi, \\ \nabla_N \phi &= \begin{pmatrix} \mathcal{D}_{N_x} \phi \\ \mathcal{D}_{N_y} \phi \end{pmatrix}, \\ \nabla_N \cdot \begin{pmatrix} \phi \\ \psi \end{pmatrix} &= \mathcal{D}_{N_x} \phi + \mathcal{D}_{N_y} \psi. \end{aligned} \quad (3.9)$$

Then, we finally can introduce the discrete operator

$$\mathcal{L}_h = -\varepsilon^2 \Delta_N + C, \quad C > 0. \quad (3.10)$$

Here, C is a positive constant. The operator \mathcal{L}^{-1} can be defined as

$$(\mathcal{L}_h^{-1} \phi)_{mn} = \frac{1}{N_x N_y} \sum_{k=-K_x+1}^{K_x} \sum_{l=-K_y+1}^{K_y} \frac{1}{\varepsilon^2 \lambda_{k,l} + C} \hat{\phi}_{kl} \exp\left(2\pi i(k \frac{x_m}{L_x} + l \frac{y_l}{L_y})\right), \quad (3.11)$$

with the coefficients calculated as

$$\lambda_{k,l} = \left(\frac{2k\pi}{L_x}\right)^2 + \left(\frac{2l\pi}{L_y}\right)^2.$$

Other exponential operators can be defined in a similar manner. Given the discrete functions in the 2D mesh $\phi, \psi \in \mathbb{R}^{N_x, N_y}$, we can define the inner product and induced l^2 norm as

$$\|\phi\|_2 = \sqrt{(\phi, \phi)}, \quad (\phi, \psi) = h_x h_y \sum_{i=0}^{N_x-1} \sum_{j=0}^{N_y-1} \phi_{ij} \psi_{ij}. \quad (3.12)$$

It can be easily shown that the following summation by parts formula hold

$$(\phi, \Delta_N \psi) = -(\nabla_N \phi, \nabla_N \psi), \quad (\phi, \Delta_N^2 \psi) = (\Delta_N \phi, \Delta_N \psi). \quad (3.13)$$

For more properties, interested readers can refer to [22]. Without loss of generality, we use the notation \mathcal{L}_h to denote the corresponding discrete operator for \mathcal{L} and \mathcal{N}_h for \mathcal{N} using the pseudo-spectral spatial discretization.

3.2.2. POD-Galerkin projection

To give an accurate low-dimensional approximation from a subspace spanned by a set of reduced basis of dimension r in \mathbb{R}^n , we use the proper orthogonal decomposition (POD) to construct a set of global basis, also known as POD modes, from a singular value decomposition (SVD) of some snapshot data of the system. Then we use Galerkin projection for dimension reduction.

Suppose we have the sampling of the phase variable $\phi(\mathbf{x}, t)$ as

$$\Phi = [\Phi_1 \quad \Phi_2 \quad \cdots \quad \Phi_m], \quad (3.14)$$

which might be measurements from simulations or experimental data. Here Φ_k is the data collected from $t = t_k$ on the equally distanced rectangular meshes represented in a vector form, i.e., $\Phi_k \in \mathbb{R}^n$ with $n = N_x \times N_y$. Typically, $n \gg m$. Denote the full singular value decomposition (SVD) of the data matrix $\Phi \in \mathbb{R}^{n,m}$ as

$$\Phi = \hat{\mathbf{U}}_\phi \hat{\Sigma}_\phi \hat{\mathbf{V}}_\phi^T,$$

where $\hat{\mathbf{U}}_\phi \in \mathbb{R}^{n,n}$, and $\hat{\mathbf{V}}_\phi \in \mathbb{R}^{m,m}$ and $\hat{\Sigma}_\phi \in \mathbb{R}^{n,m}$. Our objective is to find $r \ll n$ optimal spatial modes that are sufficient to represent the spatial dynamics accurately. Given a threshold ϵ , we can find r such that $\|\Phi - \Phi_r\| < \epsilon$, where Φ_r is from the reduced SVD for Φ as

$$\Phi \approx \Phi_r = \mathbf{U}_\phi \Sigma_\phi \mathbf{V}_\phi^T,$$

where $\mathbf{U}_\phi \in \mathbb{R}^{n,r}$, $\Sigma_\phi \in \mathbb{R}^{r,r}$ and $\mathbf{V}_\phi \in \mathbb{R}^{m,r}$. And we denote our optimal basis modes

$$\mathbf{U}_\phi := \hat{\mathbf{U}}_\phi[:, 1:r], \quad (3.15)$$

where the truncation preserves the r most dominant modes. The truncated r modes for Φ are then used as the low-rank, orthogonal basis to represent the spatial dynamics. Similarly, for the auxiliary variable $q(\mathbf{x}, t)$, we can obtain the data from the sampling in (3.14), as

$$\mathbf{Q} = [h(\Phi_1) \ h(\Phi_2) \ \dots \ h(\Phi_m)], \quad (3.16)$$

where $h(\Phi_k) \in \mathbb{R}^n$ can be treated as the sampling for $q(\mathbf{x}, t)$ at t_k on the same spatial meshes. In a similar manner, we calculate the SVD for $\mathbf{Q} \in \mathbb{R}^{n,m}$ as

$$\mathbf{Q} = \hat{\mathbf{U}}_q \hat{\Sigma}_q \hat{\mathbf{V}}_q^T,$$

and have the reduced SVD by picking $r \ll n$ as

$$\mathbf{Q}_r = \mathbf{U}_q \Sigma_q \mathbf{V}_q^T. \quad (3.17)$$

By this approach, we obtain the truncated r modes for $q(\mathbf{x}, t)$ as

$$\mathbf{U}_q := \hat{\mathbf{U}}_q[:, 1:r]. \quad (3.18)$$

With (3.15) and (3.18), we have the r modes for the general variable Ψ as

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_\phi & 0 \\ 0 & \mathbf{U}_q \end{bmatrix}. \quad (3.19)$$

Denoting

$$\mathbf{a}(t) = \begin{bmatrix} \mathbf{a}_\phi(t) \\ \mathbf{a}_q(t) \end{bmatrix}, \quad \Psi = \begin{bmatrix} \phi \\ q \end{bmatrix}, \quad (3.20)$$

where $\mathbf{a}_\phi(t) \in \mathbb{R}^r$ and $\mathbf{a}_q(t) \in \mathbb{R}^r$ are the time-dependent coefficient vectors, we approximate the solution using the POD expansion:

$$\Psi = \mathbf{U}\mathbf{a}(t), \text{ i.e., } \phi = \mathbf{U}_\phi \mathbf{a}_\phi(t), \quad q = \mathbf{U}_q \mathbf{a}_q(t). \quad (3.21)$$

Remark 3.1. Note that the POD approach constructs a reduced basis that is optimal in the sense that a certain approximation error concerning the snapshot data is minimized. The minimum 2-norm error from approximating the snapshot data using the POD modes is given by

$$\sum_{j=1}^m \|\phi_j - \Phi \Phi^T \phi_j\|_2^2 = \sum_{j=r+1}^{d_1} \sigma_j^2, \quad \sum_{j=1}^m \|h(\phi_j) - \mathbf{Q} \mathbf{Q}^T h(\phi_j)\|_2^2 = \sum_{j=r+1}^{d_2} \lambda_j^2,$$

where d_1 and d_2 are the rank of snapshot matrix Φ and \mathbf{Q} respectively, and $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{d_1} > 0$ and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{d_2} > 0$ are the nonzero singular values of Φ and \mathbf{Q} respectively. For more details on the POD basis, we refer the reader to [28].

Then, we apply POD-Galerkin projection to (3.2) and derive the POD-ROM as

$$\frac{d\mathbf{a}(t)}{dt} = -\mathbf{U}^T \mathcal{N}_h(\mathbf{U}\mathbf{a}(t)) \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)], \quad (3.22)$$

which we name POD-ROM-vanilla. By solving this system of much smaller dimensions as (3.22), the solution of a high-dimensional nonlinear dynamical system can be approximated.

However, POD-ROM in (3.22) does not necessarily respect the energy dissipation law as the full model in (2.11). Specifically, if we calculate the energy dissipation rate of (3.22) with respect to the discrete energy

$$\mathcal{E}_r(\mathbf{U}\mathbf{a}(t)) = \frac{1}{2} [\mathbf{U}\mathbf{a}(t)]^T \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)] - A_0,$$

we have

$$\frac{d\mathcal{E}_r(\mathbf{U}\mathbf{a}(t))}{dt} = (\mathcal{L}_h[\mathbf{U}\mathbf{a}(t)])^T \mathbf{U} \frac{d\mathbf{a}(t)}{dt} \quad (3.23a)$$

$$= -\left(\mathcal{L}_h[\mathbf{U}\mathbf{a}(t)]\right)^T \mathbf{U}\mathbf{U}^T \mathcal{N}_h(\mathbf{U}\mathbf{a}(t)) \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)], \quad (3.23b)$$

where $\mathbf{U}\mathbf{U}^T \mathcal{N}_h(\mathbf{U}\mathbf{a}(t))$ is not necessarily a positive semi-definite matrix, which might violate the energy dissipation law.

3.3. General POD-ROM structure-preserving framework

To address the issue above, we propose two approaches that can automatically preserve the energy dissipation structure in the POD-ROM.

3.3.1. Approach I: POD-ROM-I

As a first remedy, we modify the matrix $\mathbf{U}\mathbf{U}^T \mathcal{N}_h(\mathbf{U}\mathbf{a}(t))$ to make it positive semi-definite. Inspired by [21], we modify the mobility operator in (3.22) by assuming that there exists a mobility operator, $\hat{\mathcal{N}}_h$, with the same property as \mathcal{N}_h such that,

$$\mathbf{U}^T \mathcal{N}_h = \hat{\mathcal{N}}_h \mathbf{U}^T. \quad (3.24)$$

Then, by multiplying \mathbf{U} on both sides of the equation, we have

$$\hat{\mathcal{N}}_h = \mathbf{U}^T \mathcal{N}_h \mathbf{U}. \quad (3.25)$$

Replacing the mobility operator \mathcal{N}_h in (3.22) with $\hat{\mathcal{N}}_h$, we came up with the following POD-ROM

$$\frac{d\mathbf{a}(t)}{dt} = -\hat{\mathcal{N}}_h(\mathbf{U}\mathbf{a}(t)) \mathbf{U}^T \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)],$$

i.e.,

$$\frac{d\mathbf{a}(t)}{dt} = -\mathbf{U}^T \mathcal{N}_h(\mathbf{U}\mathbf{a}(t)) \mathbf{U}\mathbf{U}^T \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)]. \quad (3.26)$$

We named (3.26) as POD-ROM-I.

Theorem 3.1 (Energy Stability). *The POD-ROM-I in (3.26) preserves the energy dissipation law*

$$\frac{d\mathcal{E}_r(\mathbf{U}\mathbf{a}(t))}{dt} = -\left(\mathbf{U}\mathbf{U}^T \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)], \mathcal{N}_h(\mathbf{U}\mathbf{a}(t)) \mathbf{U}\mathbf{U}^T \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)]\right) \leq 0, \quad (3.27)$$

where the discrete energy is defined as $\mathcal{E}_r(\mathbf{U}\mathbf{a}(t)) = \frac{1}{2} [\mathbf{U}\mathbf{a}(t)]^T \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)] - A_0$.

Proof. As a matter of fact, this energy dissipation law can be easily observed by performing the following calculations

$$\frac{d\mathcal{E}_r(\mathbf{U}\mathbf{a}(t))}{dt} = \left(\mathcal{L}_h[\mathbf{U}\mathbf{a}(t)]\right)^T \mathbf{U} \frac{d\mathbf{a}(t)}{dt} \quad (3.28a)$$

$$= -\left(\mathcal{L}_h[\mathbf{U}\mathbf{a}(t)]\right)^T \mathbf{U}\mathbf{U}^T \mathcal{N}_h(\mathbf{U}\mathbf{a}(t)) \mathbf{U}\mathbf{U}^T \left(\mathcal{L}_h[\mathbf{U}\mathbf{a}(t)]\right) \quad (3.28b)$$

$$= -\left(\mathbf{U}\mathbf{U}^T \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)], \mathcal{N}_h(\mathbf{U}\mathbf{a}(t)) \mathbf{U}\mathbf{U}^T \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)]\right) \quad (3.28c)$$

$$\leq 0. \quad (3.28d)$$

This completes the proof. \square

However, even though the approach in (3.26) respects the energy dissipation property, it has a modified energy dissipation rate, which might lead to inaccurate dissipative dynamics.

3.3.2. Approach II: POD-ROM-II

To overcome these issues, we introduce a new approach that is inspired by [15], where the author develops energy-stable numerical algorithms for dissipative systems. Specifically, we reformulate the reduced order model by applying \mathcal{L}_h^* on both sides of the equation. Then, we obtain the following POD-ROM by Galerkin projection as

$$\mathbf{U}^T \mathcal{L}_h^* \mathbf{U} \frac{d\mathbf{a}(t)}{dt} = -\mathbf{U}^T \mathcal{L}_h^* \mathcal{N}_h(\mathbf{U}\mathbf{a}(t)) \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)], \quad (3.29)$$

which we name POD-ROM-II. Note that $\mathcal{L}_h = \begin{bmatrix} \mathcal{L}_{0,h} & 0 \\ 0 & \mathbf{I} \end{bmatrix}$ is an invertible operator. Thus the problem (3.29) is well-posed.

Theorem 3.2 (Energy Stability). *The POD-ROM-II in (3.29) preserves the energy dissipation law*

$$\frac{d\mathcal{E}_r(\mathbf{U}\mathbf{a}(t))}{dt} = -\left(\mathcal{L}_h[\mathbf{U}\mathbf{a}(t)], \mathcal{N}_h(\mathbf{U}\mathbf{a}(t)) \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)]\right) \leq 0, \quad (3.30)$$

where the discrete energy is defined as $\mathcal{E}_r(\mathbf{U}\mathbf{a}(t)) = \frac{1}{2} \left[\mathbf{U}\mathbf{a}(t) \right]^T \mathcal{L}_h \left[\mathbf{U}\mathbf{a}(t) \right] - A_0$.

Proof. As a matter of fact, this energy dissipation law can be easily observed by performing the following calculations

$$\frac{d\mathcal{E}_r(\mathbf{U}\mathbf{a}(t))}{dt} = \left(\mathcal{L}_h \left[\mathbf{U}\mathbf{a}(t) \right] \right)^T \mathbf{U} \frac{d\mathbf{a}(t)}{dt} \quad (3.31a)$$

$$= \mathbf{a}(t)^T \mathbf{U}^T \mathcal{L}_h^* \mathbf{U} \frac{d\mathbf{a}(t)}{dt} \quad (3.31b)$$

$$= - \left(\mathcal{L}_h \left[\mathbf{U}\mathbf{a}(t) \right] \right)^T \mathcal{N}_h \left(\mathbf{U}\mathbf{a}(t) \right) \left(\mathcal{L}_h \left[\mathbf{U}\mathbf{a}(t) \right] \right) \quad (3.31c)$$

$$= - \left(\mathcal{L}_h \left[\mathbf{U}\mathbf{a}(t) \right], \mathcal{N}_h \left(\mathbf{U}\mathbf{a}(t) \right) \mathcal{L}_h \left[\mathbf{U}\mathbf{a}(t) \right] \right) \quad (3.31d)$$

$$\leq 0. \quad (3.31e)$$

This completes the proof. \square

Remark 3.2. Note that POD-ROM-I in (3.26) and POD-ROM-II (3.29) are generic, and they apply to the reversible-irreversible models that fit in the generic form. POD-ROM-II in (3.29) is superior to either POD-ROM-vanilla in (3.22) or POD-ROM-I in (3.26), since (3.29) preserves the original energy dissipation structure in (2.11).

We can also expand the notations and rewrite POD-ROM-II in matrix forms. The equation of (3.29) can be written as

$$\begin{bmatrix} \mathbf{U}_\phi^T \mathcal{L}_{0,h}^* \mathbf{U}_\phi & \mathbf{U}_q^T \mathbf{U}_q \end{bmatrix} \frac{d}{dt} \begin{bmatrix} \mathbf{a}_\phi(t) \\ \mathbf{a}_q(t) \end{bmatrix} = \begin{bmatrix} \mathbf{U}_\phi^T & \mathbf{U}_q^T \end{bmatrix} \begin{bmatrix} \mathcal{L}_{0,h}^* & 0 \\ 0 & \mathbf{I} \end{bmatrix} \mathcal{N}_h \left(\begin{bmatrix} \mathbf{U}_\phi \mathbf{a}_\phi(t) \\ \mathbf{U}_q \mathbf{a}_q(t) \end{bmatrix} \right) \begin{bmatrix} \mathcal{L}_{0,h} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{U}_\phi \mathbf{a}_\phi(t) \\ \mathbf{U}_q \mathbf{a}_q(t) \end{bmatrix}.$$

The structure-preserving model of (3.29) reads

$$\mathbf{U}_\phi^T \mathcal{L}_{0,h}^* \mathbf{U}_\phi \frac{d\mathbf{a}_\phi(t)}{dt} = -\mathbf{U}_\phi^T \mathcal{L}_{0,h}^* \mathcal{G} \left(\mathcal{L}_{0,h} \mathbf{U}_\phi \mathbf{a}_\phi(t) + g(\mathbf{U}_\phi \mathbf{a}_\phi(t)) \mathbf{U}_q \mathbf{a}_q(t) \right), \quad (3.32a)$$

$$\frac{d\mathbf{a}_q(t)}{dt} = -\mathbf{U}_q^T g(\mathbf{U}_\phi \mathbf{a}_\phi(t))^T \mathcal{G} \left(\mathcal{L}_{0,h} \mathbf{U}_\phi \mathbf{a}_\phi(t) + g(\mathbf{U}_\phi \mathbf{a}_\phi(t)) \mathbf{U}_q \mathbf{a}_q(t) \right). \quad (3.32b)$$

Furthermore, the POD-ROM can be simplified as

$$A_0 \frac{d\mathbf{a}_\phi(t)}{dt} = -A_1 \mathbf{a}_\phi(t) - A_2 \mathbf{a}_q(t), \quad (3.33a)$$

$$\frac{d\mathbf{a}_q(t)}{dt} = -A_3 \mathbf{a}_\phi(t) - A_4 \mathbf{a}_q(t), \quad (3.33b)$$

with the following notations for the operators

$$A_0 = \mathbf{U}_\phi^T \mathcal{L}_{0,h}^* \mathbf{U}_\phi, \quad (3.34a)$$

$$A_1 = \mathbf{U}_\phi^T \mathcal{L}_{0,h}^* \mathcal{G} \mathcal{L}_{0,h} \mathbf{U}_\phi, \quad (3.34b)$$

$$A_2(t) = \mathbf{U}_\phi^T \mathcal{L}_{0,h}^* \mathcal{G} g(\mathbf{U}_\phi \mathbf{a}_\phi(t)) \mathbf{U}_q, \quad (3.34c)$$

$$A_3(t) = \mathbf{U}_q^T g(\mathbf{U}_\phi \mathbf{a}_\phi(t))^T \mathcal{G} \mathcal{L}_{0,h} \mathbf{U}_\phi, \quad (3.34d)$$

$$A_4(t) = \mathbf{U}_q^T g(\mathbf{U}_\phi \mathbf{a}_\phi(t))^T \mathcal{G} g(\mathbf{U}_\phi \mathbf{a}_\phi(t)) \mathbf{U}_q. \quad (3.34e)$$

The model in (3.33)-(3.34) will provide guidance on designing structure-preserving numerical algorithms for solving the POD-ROMs.

4. Linear time-stepping algorithms for the POD-ROM

With our framework, the POD-ROM in (3.26) and (3.29) can be solved. Consider the time domain $[0, T]$. We discretize it into equally distanced intervals $0 = t_0 < t_1 < \dots < t_N = T$ with $h = \frac{T}{N}$, i.e. $t_i = ih$. Then we denote \mathbf{a}^n as the numerical approximation of $\mathbf{a}(t_n)$. The initial value can be obtained as

$$\mathbf{a}_\phi^0 = \mathbf{U}_\phi^T \phi_0, \quad \mathbf{a}_q^0 = \mathbf{U}_q^T h(\phi_0).$$

With these notations, we discuss the numerical integration for the POD-ROM.

To integrate the ROM in time, we need a structure-preserving time-integrator. The widely used time integration for structure-preserving ROM is the average vector field (AVF) method. However, its drawback is obvious. It is highly nonlinear and computationally expensive, and some nonlinear iterative methods have to be used in each time step (which are also not easy to implement). The existence and uniqueness of the solution strongly limit the time-step size.

4.1. Linear semi-implicit structure-preserving time integration

Among the advantages of this work, our proposed POD-ROMs can be easily integrated in time with linear numerical schemes. If we take the semi-implicit Crank-Nicolson (CN) time discretization, we will have the following schemes for the POD-ROM-II in (3.29).

Scheme 4.1 (CN scheme for the POD-ROM-II). After we calculated \mathbf{a}^{n-1} and \mathbf{a}^n , we can obtain $\mathbf{a}^{n+1} := \begin{bmatrix} \mathbf{a}_\phi^{n+1} \\ \mathbf{a}_q^{n+1} \end{bmatrix}$ via the following linear scheme

$$\mathbf{U}^T \mathcal{L}_h^* \mathbf{U} \frac{\mathbf{a}^{n+1} - \mathbf{a}^n}{\delta t} = -\mathbf{U}^T \mathcal{L}_h^* \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathcal{L}_h[\mathbf{U} \mathbf{a}^{n+\frac{1}{2}}], \quad (4.1)$$

with the notations $\bar{\mathbf{a}}^{n+\frac{1}{2}} = \frac{3}{2}\mathbf{a}^n - \frac{1}{2}\mathbf{a}^{n-1}$ and $\mathbf{a}^{n+\frac{1}{2}} = \frac{1}{2}\mathbf{a}^{n+1} + \frac{1}{2}\mathbf{a}^n$.

Theorem 4.1. The Scheme 4.1 is linear, and it preserves the discrete energy dissipation law

$$\mathcal{E}_r(\mathbf{U} \mathbf{a}^{n+1}) - \mathcal{E}_r(\mathbf{U} \mathbf{a}^n) = -\delta t \left(\mathcal{L}_h[\mathbf{U} \mathbf{a}^{n+\frac{1}{2}}] \right)^T \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \left(\mathcal{L}_h[\mathbf{U} \mathbf{a}^{n+\frac{1}{2}}] \right) \quad (4.2)$$

where the discrete energy \mathcal{E}_r is defined as $\mathcal{E}_r(\mathbf{U} \mathbf{a}) = \frac{1}{2} \left(\mathbf{U} \mathbf{a} \right)^T \mathcal{L}_h(\mathbf{U} \mathbf{a}) - A_0$.

Proof. It is easy to observe that only a linear system needs to be solved at each time step in (4.1), i.e., Scheme 4.1 is linear. The energy stability equality is obtained by multiplying $\delta t (\mathbf{a}^{n+\frac{1}{2}})^T$ on both sides of (4.1). \square

Similarly, if we use the semi-implicit second-order backward differential formula (BDF2) for the temporal discretization, we arrive at the following second-order linear numerical algorithm.

Scheme 4.2 (BDF2 scheme for the POD-ROM-II). After we calculated \mathbf{a}^{n-1} and \mathbf{a}^n , we can obtain $\mathbf{a}^{n+1} := \begin{bmatrix} \mathbf{a}_\phi^{n+1} \\ \mathbf{a}_q^{n+1} \end{bmatrix}$ via the following linear scheme

$$\mathbf{U}^T \mathcal{L}_h^* \mathbf{U} \frac{3\mathbf{a}^{n+1} - 4\mathbf{a}^n + \mathbf{a}^{n-1}}{2\delta t} = -\mathbf{U}^T \mathcal{L}_h^* \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+1}) \mathcal{L}_h[\mathbf{U} \mathbf{a}^{n+1}], \quad (4.3)$$

with the notation $\bar{\mathbf{a}}^{n+1} = 2\mathbf{a}^n - \mathbf{a}^{n-1}$.

Theorem 4.2. The Scheme 4.2 is linear, and it satisfies the following discrete energy dissipation law

$$\hat{\mathcal{E}}_r(\mathbf{U} \mathbf{a}^{n+1}, \mathbf{U} \mathbf{a}^n) - \hat{\mathcal{E}}_r(\mathbf{U} \mathbf{a}^n, \mathbf{U} \mathbf{a}^{n-1}) \leq -\delta t \left(\mathcal{L}_h[\mathbf{U} \mathbf{a}^{n+1}] \right)^T \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+1}) \left(\mathcal{L}_h[\mathbf{U} \mathbf{a}^{n+1}] \right), \quad (4.4)$$

where the modified discrete free energy is defined as

$$\hat{\mathcal{E}}_r(\mathbf{U} \mathbf{a}_1, \mathbf{U} \mathbf{a}_2) = \frac{1}{4} \left[\mathbf{U} \mathbf{a}_1 \right]^T \mathcal{L}_h \left[\mathbf{U} \mathbf{a}_1 \right] + \frac{1}{4} \left[\mathbf{U} (2\mathbf{a}_1 - \mathbf{a}_2) \right]^T \mathcal{L}_h \left[\mathbf{U} (2\mathbf{a}_1 - \mathbf{a}_2) \right] - A_0.$$

Proof. Notice the inequality

$$\begin{aligned} \left(\frac{3a - 4b + c}{2}, a \right) &= \frac{1}{4} \left[(a, a) + (2a - b, 2a - b) - (b, b) - (2b - c, 2b - c) + (a - 2b + c, a - 2b + c) \right] \\ &\geq \frac{1}{4} \left[(a, a) + (2a - b, 2a - b) - (b, b) - (2b - c, 2b - c) \right]. \end{aligned}$$

If we multiply $\delta t (\mathbf{a}^{n+1})^T$ on both sides of (4.3), the energy stability inequality is obtained. \square

Theorem 4.3. The Scheme 4.1 and Scheme 4.2 are both linear and uniquely solvable.

Proof. Given the similarity between Scheme 4.1 and Scheme 4.2, we only show the proof for Scheme 4.1. After some basic algebraic calculation, we can rewrite Scheme 4.1 as

$$\left[\frac{1}{\delta t} \mathbf{U}^T \mathcal{L}_h^* \mathbf{U} + \frac{1}{2} \mathbf{U}^T \mathcal{L}_h^* \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathcal{L}_h \mathbf{U} \right] \mathbf{a}^{n+1} = \frac{1}{\delta t} \mathbf{U}^T \mathcal{L}_h^* [\mathbf{U} \mathbf{a}^n] - \frac{1}{2} \mathbf{U}^T \mathcal{L}_h^* \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathcal{L}_h [\mathbf{U} \mathbf{a}^n]. \quad (4.5)$$

To show the existence and uniqueness of the solution for (4.1), we only need to show there is only a zero solution for

$$A \mathbf{a}^{n+1} = 0, \quad \text{where } A = \frac{1}{\delta t} \mathbf{U}^T \mathcal{L}_h^* \mathbf{U} + \frac{1}{2} \mathbf{U}^T \mathcal{L}_h^* \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathcal{L}_h \mathbf{U}.$$

If \mathbf{a} is a solution to $A\mathbf{a}^{n+1} = 0$, we have

$$0 = \mathbf{a}^T A \mathbf{a} = \frac{\mathbf{a}^T \mathbf{U}^T \mathcal{L}_h^* \mathbf{U} \mathbf{a}}{\delta t} + \frac{1}{2} \mathbf{a}^T \left[\mathbf{U}^T \mathcal{L}_h^* \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathcal{L}_h \mathbf{U} \right] \mathbf{a} \geq \frac{\mathbf{a}^T \mathbf{U}^T \mathcal{L}_h^* \mathbf{U} \mathbf{a}}{\delta t},$$

i.e. $\mathbf{a} = 0$, by noticing $\mathbf{U}^T \mathcal{L}_h^* \mathbf{U}$ is a positive definite matrix. So, A is invertible, and there is a unique solution for Scheme 4.1. Furthermore, we have shown that A is a positive definite matrix. \square

For the POD-ROM-I in (3.26), similar numerical techniques can be applied. Specifically, we have the following time-marching linear numerical schemes that will also be used for comparisons in the numerical result section.

Scheme 4.3 (CN scheme for POD-ROM-I). The CN scheme for (3.26) reads

$$\frac{\mathbf{a}^{n+1} - \mathbf{a}^n}{\delta t} = -\mathbf{U}^T \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathbf{U} \mathbf{U}^T \mathcal{L}_h[\mathbf{U} \mathbf{a}^{n+\frac{1}{2}}]. \quad (4.6)$$

Scheme 4.4 (BDF2 scheme for POD-ROM-I). The BDF2 scheme for (3.26) reads

$$\frac{3\mathbf{a}^{n+1} - 4\mathbf{a}^n + \mathbf{a}^{n-1}}{2\delta t} = -\mathbf{U}^T \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+1}) \mathbf{U} \mathbf{U}^T \mathcal{L}_h[\mathbf{U} \mathbf{a}^{n+1}]. \quad (4.7)$$

From the discussion above, we conclude that Scheme 4.1 and Scheme 4.2 are linear, uniquely solvable, and preserve the energy dissipation structure. However, they both have one defect. We recall the original definition for q in (2.6), denoting $q := h(\phi)$. After temporal discretization, the numerical result $\mathbf{U}_q \mathbf{a}_q^{n+1}$ from Scheme 4.1 or Scheme 4.2 is not necessarily equal to $h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1})$ anymore, which leads to numerical errors.

4.2. Relaxation technique to improve the accuracy

One remedy to fix the issue above is to utilize the relaxation technique we proposed to improve the accuracy and stability of EQ and SAV methods [26,48]. Namely, we can relax the numerical solution $\mathbf{U}_q \mathbf{a}_q^{n+1}$ from Scheme 4.1 and Scheme 4.2 such that the numerical error $\|\mathbf{U}_q \mathbf{a}_q^{n+1} - h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1})\|$ will be damped to zero gradually. Therefore, we develop the following relaxed schemes, which will be used in this paper to conduct numerical simulations.

Scheme 4.5 (Relaxed CN scheme for the POD-ROM-II). After we calculated \mathbf{a}^{n-1} and \mathbf{a}^n , we can obtain \mathbf{a}^{n+1} via the following three steps:

- Step 1. Obtain $\hat{\mathbf{a}}^{n+1} := \begin{bmatrix} \mathbf{a}_\phi^{n+1} \\ \hat{\mathbf{a}}_q^{n+1} \end{bmatrix}$ via the linear scheme

$$\mathbf{U}^T \mathcal{L}_h^* \mathbf{U} \frac{\hat{\mathbf{a}}^{n+1} - \mathbf{a}^n}{\delta t} = -\mathbf{U}^T \mathcal{L}_h^* \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}], \quad (4.8)$$

with the notations $\bar{\mathbf{a}}^{n+\frac{1}{2}} = \frac{3}{2}\mathbf{a}^n - \frac{1}{2}\mathbf{a}^{n-1}$ and $\hat{\mathbf{a}}^{n+\frac{1}{2}} = \frac{1}{2}\hat{\mathbf{a}}^{n+1} + \frac{1}{2}\mathbf{a}^n$.

- Step 2. Update $\mathbf{U}_q \mathbf{a}_q^{n+1}$ via the relaxation strategy

$$\mathbf{U}_q \mathbf{a}_q^{n+1} = \xi_0 \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} + (1 - \xi_0) h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}), \quad \xi_0 \in [0, 1],$$

where ξ_0 is a solution for the optimization problem:

$$\xi_0 = \min_{\xi \in [0, 1]} \xi, \quad \text{s.t. } \frac{1}{2} (\mathbf{U}_q \mathbf{a}_q^{n+1}, \mathbf{U}_q \mathbf{a}_q^{n+1}) - \frac{1}{2} (\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}, \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}) \leq \delta t \eta \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}], \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}] \right), \quad (4.9)$$

with an artificial parameter $\eta \in [0, 1]$ that can be assigned.

- Step 3. Update \mathbf{a}^{n+1} as $\mathbf{a}^{n+1} := \begin{bmatrix} \mathbf{a}_\phi^{n+1} \\ \mathbf{a}_q^{n+1} \end{bmatrix}$.

Remark 4.1. The Scheme 4.5 relaxes the solution $\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}$ to get $\mathbf{U}_q \mathbf{a}_q^{n+1}$ which is closer to the original definition of q and then we obtain \mathbf{a}_q^{n+1} by $\mathbf{U}_q^T \mathbf{U}_q \mathbf{a}_q^{n+1}$. Since $\mathbf{U}_q^T \mathbf{U}_q = \mathbf{I}_r$, the relaxation strategy in Step 2 is equivalent to relaxing the numerical solution $\hat{\mathbf{a}}_q^{n+1}$ via

$$\mathbf{a}_q^{n+1} = \xi_0 \hat{\mathbf{a}}_q^{n+1} + (1 - \xi_0) \mathbf{U}_q^T h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}).$$

Remark 4.2 (Optimal choice for ξ_0). The optimization problem in (4.9) can be simplified as the following algebraic optimization problem

$$\xi_0 = \min_{\xi \in [0,1], a\xi^2 + b\xi + c \leq 0} \xi,$$

where the coefficients are given by

$$\begin{aligned} a &= \frac{1}{2} \left(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}), \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}) \right), \\ b &= \left(h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}), \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}) \right), \\ c &= \frac{1}{2} \left(h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}), h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}) \right) - \frac{1}{2} \left(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}, \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} \right) - \delta t \eta \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}], \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}] \right). \end{aligned}$$

The solution set is nonempty since $\xi = 1$ is in the feasible domain. Also, notice that $a + b + c > 0$ and $\delta t \eta \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}], \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}] \right) > 0$. With $a > 0$, the optimization problem in (4.9) can be solved as

$$\xi_0 = \max \{0, \frac{-b - \sqrt{b^2 - 4ac}}{2a}\}.$$

Theorem 4.4. *The Scheme 4.5 is unconditionally energy stable.*

Proof. According to the Theorem 4.1, the first step of Scheme 4.5 gives us the following energy dissipation law

$$\begin{aligned} &\frac{1}{2} \left[\left(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}, \mathcal{L}_{0,h} \mathbf{U}_\phi \mathbf{a}_\phi^{n+1} \right) + \left(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}, \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} \right) \right] - \frac{1}{2} \left[\left(\mathbf{U}_\phi \mathbf{a}_\phi^n, \mathcal{L}_{0,h} \mathbf{U}_\phi \mathbf{a}_\phi^n \right) + \left(\mathbf{U}_q \mathbf{a}_q^n, \mathbf{U}_q \mathbf{a}_q^n \right) \right] \\ &= -\delta t \left(\mathcal{L}_h(\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}) \right)^T \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}] \right). \end{aligned}$$

Also, from (4.9), we know

$$\frac{1}{2} \left(\mathbf{U}_q \mathbf{a}_q^{n+1}, \mathbf{U}_q \mathbf{a}_q^{n+1} \right) - \frac{1}{2} \left(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}, \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} \right) \leq \delta t \eta \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}], \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}] \right).$$

Adding the above two equations together gives us the following energy dissipation law

$$\begin{aligned} &\frac{1}{2} \left[\left(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}, \mathcal{L}_{0,h} \mathbf{U}_\phi \mathbf{a}_\phi^{n+1} \right) + \left(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}, \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} \right) \right] - \frac{1}{2} \left[\left(\mathbf{U}_\phi \mathbf{a}_\phi^n, \mathcal{L}_{0,h} \mathbf{U}_\phi \mathbf{a}_\phi^n \right) + \left(\mathbf{U}_q \mathbf{a}_q^n, \mathbf{U}_q \mathbf{a}_q^n \right) \right] \\ &\leq -\delta t (1 - \eta) \left(\mathcal{L}_h(\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}) \right)^T \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+\frac{1}{2}}) \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+\frac{1}{2}}] \right) \leq 0. \quad \square \end{aligned} \tag{4.10}$$

Scheme 4.6 (Relaxed BDF2 scheme for the POD-ROM-II). After we calculated \mathbf{a}^{n-1} and \mathbf{a}^n , we can obtain \mathbf{a}^{n+1} via the following three steps:

- Step 1. Obtain $\hat{\mathbf{a}}^{n+1} := \begin{bmatrix} \mathbf{a}_\phi^{n+1} \\ \hat{\mathbf{a}}_q^{n+1} \end{bmatrix}$ via the linear scheme

$$\mathbf{U}^T \mathcal{L}_h^* \mathbf{U} \frac{3\hat{\mathbf{a}}^{n+1} - 4\mathbf{a}^n + \mathbf{a}^{n-1}}{2\delta t} = -\mathbf{U}^T \mathcal{L}_h^* \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+1}) \mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+1}], \tag{4.11}$$

with the notation $\bar{\mathbf{a}}^{n+1} = 2\mathbf{a}^n - \mathbf{a}^{n-1}$.

- Step 2. Update $\mathbf{U}_q \mathbf{a}_q^{n+1}$ via the relaxation strategy

$$\mathbf{U}_q \mathbf{a}_q^{n+1} = \xi_0 \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} + (1 - \xi_0) h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}), \quad \xi_0 \in [0, 1],$$

where ξ_0 is a solution for the optimization problem:

$$\begin{aligned} \xi_0 &= \min_{\xi \in [0,1]} \xi, \quad \text{s.t. } \frac{1}{4} \left(\left(\mathbf{U}_q \mathbf{a}_q^{n+1}, \mathbf{U}_q \mathbf{a}_q^{n+1} \right) + \left(2\mathbf{U}_q \mathbf{a}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n, 2\mathbf{U}_q \mathbf{a}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n \right) \right) \\ &\quad - \frac{1}{4} \left(\left(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}, \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} \right) + \left(2\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n, 2\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n \right) \right) \\ &\leq \delta t \eta \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+1}], \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+1}) \mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+1}] \right), \end{aligned} \tag{4.12}$$

with an artificial parameter $\eta \in [0, 1]$ that can be assigned.

- Step 3. Update \mathbf{a}^{n+1} as $\mathbf{a}^{n+1} := \begin{bmatrix} \mathbf{a}_\phi^{n+1} \\ \mathbf{a}_q^{n+1} \end{bmatrix}$.

Remark 4.3 (Optimal choice for ξ_0). In a similar manner, the optimization problem in (4.12) can be simplified as the following algebraic optimization problem

$$\xi_0 = \min_{\xi \in [0,1], a\xi^2 + b\xi + c \leq 0} \xi,$$

where the coefficients are given by

$$\begin{aligned} a &= \frac{5}{4} \left(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}), \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}) \right), \\ b &= \frac{1}{2} \left(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}), 5h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}) - 2\mathbf{U}_q \mathbf{a}_q^n \right), \\ c &= \frac{1}{4} \left[\left(h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}), h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}) \right) + \left(2h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}) - \mathbf{U}_q \mathbf{a}_q^n, 2h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}) - \mathbf{U}_q \mathbf{a}_q^n \right) \right. \\ &\quad \left. - \left(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}, \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} \right) - \left(2\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n, 2\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n \right) \right] - \delta t \eta \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+1}], \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+1}) \mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+1}] \right). \end{aligned}$$

The solution set is nonempty since $\xi = 1$ is in the feasible domain. Also, notice that $a + b + c > 0$ and

$$\delta t \eta \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+1}], \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+1}) \mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+1}] \right) > 0.$$

With $a > 0$, the optimization problem in (4.12) can be solved as

$$\xi_0 = \max \{ 0, \frac{-b - \sqrt{b^2 - 4ac}}{2a} \}.$$

Theorem 4.5. *The Scheme 4.6 is unconditionally energy stable.*

Proof. According to the Theorem 4.2, the first step of Scheme 4.6 gives us the following energy dissipation law

$$\begin{aligned} &\frac{1}{4} \left[\left(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}, \mathcal{L}_{0,h} \mathbf{U}_\phi \mathbf{a}_\phi^{n+1} \right) + \left(2\mathbf{U}_\phi \mathbf{a}_\phi^{n+1} - \mathbf{U}_\phi \mathbf{a}_q^n, \mathcal{L}_{0,h} (2\mathbf{U}_\phi \mathbf{a}_\phi^{n+1} - \mathbf{U}_\phi \mathbf{a}_\phi^n) \right) \right. \\ &\quad + \left(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}, \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} \right) + \left(2\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n, 2\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n \right) \left. \right] \\ &\quad - \frac{1}{4} \left[\left(\mathbf{U}_\phi \mathbf{a}_\phi^n, \mathcal{L}_{0,h} \mathbf{U}_\phi \mathbf{a}_\phi^n \right) + \left(2\mathbf{U}_\phi \mathbf{a}_\phi^n - \mathbf{U}_\phi \mathbf{a}_q^{n-1}, \mathcal{L}_{0,h} (2\mathbf{U}_\phi \mathbf{a}_\phi^n - \mathbf{U}_\phi \mathbf{a}_\phi^{n-1}) \right) \right. \\ &\quad + \left(\mathbf{U}_q \mathbf{a}_q^n, \mathbf{U}_q \mathbf{a}_q^n \right) + \left(2\mathbf{U}_q \mathbf{a}_q^n - \mathbf{U}_q \mathbf{a}_q^{n-1}, 2\mathbf{U}_q \mathbf{a}_q^n - \mathbf{U}_q \mathbf{a}_q^{n-1} \right) \left. \right] \\ &\leq -\delta t \eta \left(\mathcal{L}_h(\mathbf{U} \hat{\mathbf{a}}^{n+1}) \right)^T \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+1}) \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+1}] \right). \end{aligned}$$

Also, from (4.12), we know

$$\begin{aligned} &\frac{1}{4} \left(\left(\mathbf{U}_q \mathbf{a}_q^{n+1}, \mathbf{U}_q \mathbf{a}_q^{n+1} \right) + \left(2\mathbf{U}_q \mathbf{a}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n, 2\mathbf{U}_q \mathbf{a}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n \right) \right) \\ &\quad - \frac{1}{4} \left(\left(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}, \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} \right) + \left(2\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n, 2\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n \right) \right) \\ &\leq \delta t \eta \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+1}], \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+1}) \mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+1}] \right). \end{aligned}$$

Adding the above two equations together gives us the following energy dissipation law

$$\begin{aligned} &\frac{1}{4} \left[\left(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}, \mathcal{L}_{0,h} \mathbf{U}_\phi \mathbf{a}_\phi^{n+1} \right) + \left(2\mathbf{U}_\phi \mathbf{a}_\phi^{n+1} - \mathbf{U}_\phi \mathbf{a}_q^n, \mathcal{L}_{0,h} (2\mathbf{U}_\phi \mathbf{a}_\phi^{n+1} - \mathbf{U}_\phi \mathbf{a}_\phi^n) \right) \right. \\ &\quad + \left(\mathbf{U}_q \mathbf{a}_q^{n+1}, \mathbf{U}_q \mathbf{a}_q^{n+1} \right) + \left(2\mathbf{U}_q \mathbf{a}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n, 2\mathbf{U}_q \mathbf{a}_q^{n+1} - \mathbf{U}_q \mathbf{a}_q^n \right) \left. \right] \\ &\quad - \frac{1}{4} \left[\left(\mathbf{U}_\phi \mathbf{a}_\phi^n, \mathcal{L}_{0,h} \mathbf{U}_\phi \mathbf{a}_\phi^n \right) + \left(2\mathbf{U}_\phi \mathbf{a}_\phi^n - \mathbf{U}_\phi \mathbf{a}_q^{n-1}, \mathcal{L}_{0,h} (2\mathbf{U}_\phi \mathbf{a}_\phi^n - \mathbf{U}_\phi \mathbf{a}_\phi^{n-1}) \right) \right. \\ &\quad + \left(\mathbf{U}_q \mathbf{a}_q^n, \mathbf{U}_q \mathbf{a}_q^n \right) + \left(2\mathbf{U}_q \mathbf{a}_q^n - \mathbf{U}_q \mathbf{a}_q^{n-1}, 2\mathbf{U}_q \mathbf{a}_q^n - \mathbf{U}_q \mathbf{a}_q^{n-1} \right) \left. \right] \\ &\leq -\delta t (1 - \eta) \left(\mathcal{L}_h(\mathbf{U} \hat{\mathbf{a}}^{n+1}) \right)^T \mathcal{N}_h(\mathbf{U} \bar{\mathbf{a}}^{n+1}) \left(\mathcal{L}_h[\mathbf{U} \hat{\mathbf{a}}^{n+1}] \right). \quad \square \end{aligned} \tag{4.13}$$

Similar relaxation techniques can be applied to the numerical schemes for POD-ROM-I in Scheme 4.3 and Scheme 4.4. For instance, we can have the following relaxed CN scheme.

Scheme 4.7 (Relaxed CN scheme for the POD-ROM-I). After we calculated \mathbf{a}^{n-1} and \mathbf{a}^n , we can obtain \mathbf{a}^{n+1} via the following three steps:

- Step 1. Obtain $\hat{\mathbf{a}}^{n+1} := \begin{bmatrix} \mathbf{a}_\phi^{n+1} \\ \hat{\mathbf{a}}_q^{n+1} \end{bmatrix}$ via the linear scheme

$$\frac{\hat{\mathbf{a}}^{n+1} - \mathbf{a}^n}{\delta t} = -\mathbf{U}^T \mathcal{N}_h(\mathbf{U}\hat{\mathbf{a}}^{n+\frac{1}{2}}) \mathbf{U} \mathbf{U}^T \mathcal{L}_h[\mathbf{U}\hat{\mathbf{a}}^{n+\frac{1}{2}}], \quad (4.14)$$

with the notations $\bar{\mathbf{a}}^{n+\frac{1}{2}} = \frac{3}{2}\mathbf{a}^n - \frac{1}{2}\mathbf{a}^{n-1}$ and $\hat{\mathbf{a}}^{n+\frac{1}{2}} = \frac{1}{2}\hat{\mathbf{a}}^{n+1} + \frac{1}{2}\mathbf{a}^n$.

- Step 2. Update $\mathbf{U}_q \mathbf{a}_q^{n+1}$ via the relaxation strategy

$$\mathbf{U}_q \mathbf{a}_q^{n+1} = \xi_0 \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1} + (1 - \xi_0) h(\mathbf{U}_\phi \mathbf{a}_\phi^{n+1}), \quad \xi_0 \in [0, 1],$$

where ξ_0 is a solution for the optimization problem:

$$\begin{aligned} \xi_0 = \min_{\xi \in [0, 1]} \xi, \quad \text{s.t. } & \frac{1}{2}(\mathbf{U}_q \mathbf{a}_q^{n+1}, \mathbf{U}_q \mathbf{a}_q^{n+1}) - \frac{1}{2}(\mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}, \mathbf{U}_q \hat{\mathbf{a}}_q^{n+1}) \\ & \leq \delta t \eta \left(\mathbf{U} \mathbf{U}^T \mathcal{L}_h[\mathbf{U}\hat{\mathbf{a}}^{n+\frac{1}{2}}], \mathcal{N}_h(\mathbf{U}\bar{\mathbf{a}}^{n+\frac{1}{2}}) \mathbf{U} \mathbf{U}^T \mathcal{L}_h[\mathbf{U}\hat{\mathbf{a}}^{n+\frac{1}{2}}] \right), \end{aligned} \quad (4.15)$$

with an artificial parameter $\eta \in [0, 1]$ that can be assigned.

- Step 3. Update \mathbf{a}^{n+1} as $\mathbf{a}^{n+1} := \begin{bmatrix} \mathbf{a}_\phi^{n+1} \\ \mathbf{a}_q^{n+1} \end{bmatrix}$.

The proof for the energy stability to Scheme 4.7 is similar to the one for Scheme 4.5.

4.3. Discrete empirical interpolation method for nonlinear terms

For the nonlinear term, we can utilize the discrete empirical interpolation method (DEIM) [10] to reduce the computational costs. Mainly, as we can tell from (3.29) and (2.10), the only nonlinearity comes from the discrete mobility operator $\mathcal{N}_h(\mathbf{U}\mathbf{a}(t))$, which is the term $g(\mathbf{U}_\phi \mathbf{a}_\phi(t))$. When a general nonlinearity is present, the cost to evaluate the nonlinear terms still depends on the dimension of the original system since $\mathbf{U}_\phi \mathbf{a}_\phi(t) \in \mathbb{R}^n$. The DEIM approach [10] provides an interpolation-based projection to approximate the nonlinear term where the interpolation indices are selected to limit the growth of an error bound. With the data (3.14), we can have

$$\mathbf{N} = [g(\phi_1) \ g(\phi_2) \ \cdots \ g(\phi_{m-1}) \ g(\phi_m)]. \quad (4.16)$$

The approximation from projecting $g(\mathbf{U}_\phi \mathbf{a}_\phi(t))$, denoted as $g(t)$, onto the subspace obtained from the POD approach is of the form

$$g(t) \approx \mathbf{W} \mathbf{c}(t), \quad (4.17)$$

where $\mathbf{W} = [\mathbf{w}_1 \ \cdots \ \mathbf{w}_k] \in \mathbb{R}^{n,k}$ with $k \ll n$ consists of the first k columns of the left singular matrix from SVD, and a time-dependent coefficient vector $\mathbf{c}(t) \in \mathbb{R}^k$. To determine $\mathbf{c}(t)$ which is a highly overdetermined system, we can construct a basis dependent interpolation matrix, $P = [\mathbf{e}_{\gamma_1} \ \cdots \ \mathbf{e}_{\gamma_k}] \in \mathbb{R}^{n,k}$, in an algorithmic way given in [10] where $\mathbf{e}_{\gamma_i} \in \mathbb{R}^n$ is the γ_i th column of the identity matrix $\mathbf{I}_n \in \mathbb{R}^{n,n}$ for $i = 1, \dots, k$. Suppose $P^T \mathbf{W} \in \mathbb{R}^{k,k}$ is nonsingular, then the coefficient vector $\mathbf{c}(t)$ is determined uniquely from

$$P^T g(t) = P^T \mathbf{W} \mathbf{c}(t). \quad (4.18)$$

Finally, $g(t)$ can be approximated by

$$g(t) \approx \mathbf{W} \mathbf{c}(t) = \mathbf{W} (P^T \mathbf{W})^{-1} P^T g(t). \quad (4.19)$$

If g is a component-wise function, we can further have

$$g(t) \approx \mathbf{W} (P^T \mathbf{W})^{-1} g(P^T \mathbf{U}_\phi \mathbf{a}_\phi(t)). \quad (4.20)$$

Notice the fact $\mathbf{W} (P^T \mathbf{W})^{-1} \in \mathbb{R}^{n,k}$ which is precomputed and $P^T \mathbf{U}_\phi \in \mathbb{R}^{k,r}$, which are independent of n of the full-order system.

Here, we briefly explain how DEIM can be applied to the POD-ROM-II in (3.29) with the following nonlinear term,

$$\begin{aligned} \mathcal{N}_h(\mathbf{U}\mathbf{a}(t)) &= \mathcal{N}_0^* \mathcal{G} \mathcal{N}_0 \\ &= \begin{bmatrix} \mathbf{I} \\ g(\mathbf{U}_\phi \mathbf{a}_\phi(t))^T \end{bmatrix} \mathcal{G} [\mathbf{I} \ g(\mathbf{U}_\phi \mathbf{a}_\phi(t))] \\ &= \begin{bmatrix} \mathcal{G} \\ g(\mathbf{U}_\phi \mathbf{a}_\phi(t))^T \mathcal{G} & \mathcal{G} g(\mathbf{U}_\phi \mathbf{a}_\phi(t)) \\ g(\mathbf{U}_\phi \mathbf{a}_\phi(t))^T \mathcal{G} & g(\mathbf{U}_\phi \mathbf{a}_\phi(t))^T \mathcal{G} g(\mathbf{U}_\phi \mathbf{a}_\phi(t)) \end{bmatrix}. \end{aligned} \quad (4.21)$$

We apply DEIM to $g(\mathbf{U}_\phi \mathbf{a}_\phi(t))$ where the nonlinearity lies and we approximate it by $g(\mathbf{U}_\phi \mathbf{a}_\phi(t)) \approx \mathbf{W} (P^T \mathbf{W})^{-1} g(P^T \mathbf{U}_\phi \mathbf{a}_\phi(t))$, where $\mathbf{W} \in \mathbb{R}^{n,k}$ with $k \ll n$ and $P \in \mathbb{R}^{n,k}$. Then for the right-hand side term, $\mathbf{U}^T \mathcal{L}_h^* \mathcal{N}_h(\mathbf{U}\mathbf{a}(t)) \mathcal{L}_h[\mathbf{U}\mathbf{a}(t)]$, in (3.29), we need to compute

(3.34), where $\mathbf{U}_\phi \in \mathbb{R}^{n,r}$ with $r \ll n$ and $\mathbf{U}_q \in \mathbb{R}^{n,r}$. Then A_0 and A_1 are $r \times r$ matrix which is pre-computed only once in the offline stage. As for A_2 , A_3 and A_4 , we plug in the approximation of $g(\mathbf{U}_\phi \mathbf{a}_\phi(t))$ and we can get

$$A_2(t) = \mathbf{U}_\phi^T \mathcal{L}_{0,h}^* \mathcal{G} g(\mathbf{U}_\phi \mathbf{a}_\phi(t)) \mathbf{U}_q = \mathbf{U}_\phi^T \mathcal{L}_{0,h}^* \mathcal{G} \mathbf{W} (P^T \mathbf{W})^{-1} g(P^T \mathbf{U}_\phi \mathbf{a}_\phi(t)) \mathbf{U}_q,$$

where $\mathbf{U}_\phi^T \mathcal{L}_{0,h}^* \mathcal{G} \mathbf{W} (P^T \mathbf{W})^{-1} \in \mathbb{R}^{r,k}$ and $P^T \mathbf{U}_\phi \in \mathbb{R}^{k,r}$ can be pre-computed only once in the offline stage and $g(P^T \mathbf{U}_\phi \mathbf{a}_\phi(t))$ is computed in the online stage where we have a $k \times r$ matrix multiplied by $\mathbf{a}_\phi(t) \in \mathbb{R}^{r,1}$ and g is a function applied element-wise to a $k \times 1$ vector and finally we have $g(P^T \mathbf{U}_\phi \mathbf{a}_\phi(t)) = [g_1 \ g_2 \ \dots \ g_k]^T$. Similar arguments apply to A_3 and A_4 .

For simplicity, the DEIM is not utilized in the later numerical examples. But we emphasize that the application of DEIM doesn't affect our theoretical results in the previous sections since the nonlinearity in our POD-ROM is dealt explicitly as shown in Scheme 4.5 and Scheme 4.6.

5. Numerical examples

So far, we have provided a unified platform to develop structure-preserving ROMs for thermodynamically consistent PDE models and their structure-preserving numerical approximations. The main idea is to transform the thermodynamically consistent PDE models into an equivalent form using the energy quadratization approach. The free energy of the transformed system is in the quadratic form of the state variables, which we exploit to develop the ROMs.

To illustrate the effectiveness of our numerical platform, we present several examples in this section. Specifically, we apply the general framework to several thermodynamically consistent phase field models.

5.1. Allen-Cahn equation

We start with the Allen-Cahn (AC) equation

$$\partial_t \phi = -M(-\varepsilon^2 \Delta \phi + \phi^3 - \phi), \quad (\mathbf{x}, t) \in \Omega \times (0, T], \quad (5.1a)$$

$$\phi(\mathbf{x}, 0) = \phi_0(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (5.1b)$$

with periodic boundary conditions. Here Ω is a smooth domain, $M > 0$ is the mobility constant, and ε is a parameter to control the interfacial thickness. The Onsager triplet for the AC equation is

$$(\phi, \mathcal{G}, \mathcal{E}) := \left(\phi, M, \int_{\Omega} \left[\frac{\varepsilon^2}{2} |\nabla \phi|^2 + \frac{1}{4} (\phi^2 - 1)^2 \right] d\mathbf{x} \right).$$

If we introduce the auxiliary variables

$$q := \frac{\sqrt{2}}{2} (\phi^2 - 1 - \gamma_0), \quad g(\phi) := \frac{\partial q}{\partial \phi} = \sqrt{2} \phi,$$

we can rewrite the equation as

$$\partial_t \phi = -M \left(-\varepsilon^2 \Delta \phi + \gamma_0 \phi + q g(\phi) \right), \quad (5.2a)$$

$$\partial_t q = g(\phi) \partial_t \phi. \quad (5.2b)$$

Denote $\Psi = \begin{bmatrix} \phi \\ q \end{bmatrix}$, $\mathcal{G}_s = M$, $\mathcal{L}_0 = -\varepsilon^2 \Delta + \gamma_0$, $\mathcal{L} = \begin{bmatrix} \mathcal{L}_0 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathcal{N}_0 = [\mathbf{I} \ g(\phi)]$, such that the equation is written as

$$\partial_t \Psi = -\mathcal{N}(\Psi) \mathcal{L} \Psi, \quad \text{where} \quad \mathcal{N}(\Psi) = \mathcal{N}_0^T \mathcal{G}_s \mathcal{N}_0. \quad (5.3)$$

Hence, the structure-preserving POD-ROMs can be derived from the techniques introduced in previous sections. Now, we consider a specific numerical example.

Example 1. We choose the model parameters $M = 1$ and $\varepsilon = 0.02$. The domain Ω is set as $[0, 1]^2$. The initial condition is

$$\phi_0(x, y) = 2 \left[\sum_{i=1}^7 \frac{1}{2} \left(1 - \tanh \frac{\sqrt{(x - X(i))^2 + (y - Y(i))^2}}{\varepsilon} - R(i) \right) \right] - 1, \quad (5.4)$$

where the parameters are

$$X = [1/4 \ 1/8 \ 1/4 \ 1/2 \ 3/4 \ 1/2 \ 3/4], \quad (5.5a)$$

$$Y = [1/4 \ 3/8 \ 5/8 \ 1/8 \ 1/8 \ 1/2 \ 3/4], \quad (5.5b)$$

$$R = [1/20 \ 1/16 \ 1/12 \ 1/12 \ 1/10 \ 1/8 \ 1/8]. \quad (5.5c)$$

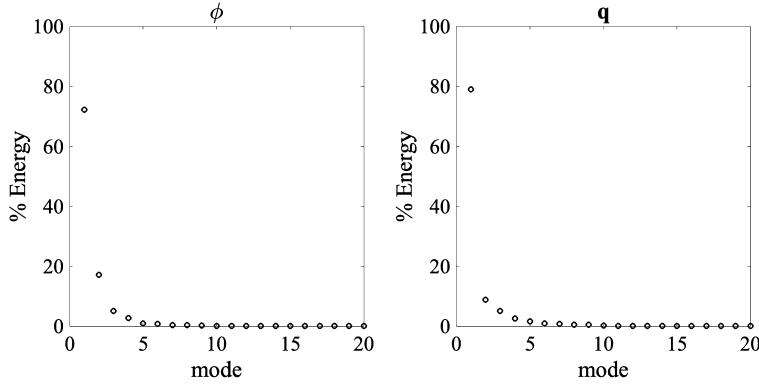


Fig. 1. Singular value distributions for the data collected from the Allen-Cahn equation.

This represents seven disks of various sizes in different domain locations. Consider $T = 15$. We first generate the data using an accurate numerical solver. In this case, we use the Crank-Nicolson scheme based on the EQ method to solve the equation with adequate small time step $\delta t = 10^{-4}$, which is used as the “accurate” solution. Then, we sample the following data $[\Phi_1 \ \Phi_2 \ \dots \ \Phi_m]$, where $\Phi_k \in \mathbf{R}^n$ with $n = N_x N_y$ is the numerical solution at $t = 0.1k$ in a vector form. We use $N_x = N_y = 128$ and $m = 150$ in this example. Then, we follow the POD-ROM numerical framework proposed in previous sections. The numerical parameters are $\delta t = 10^{-3}$ and $\gamma_0 = 1$. The distribution of the singular values is summarized in Fig. 1, where the percentage of each singular value relative to the total sum of singular values is calculated.

The numerical results using the ROM and proposed schemes are summarized in Fig. 2. The figure shows that the numerical approximation of the reduced-order model, even with just $r=4$ modes, provides a reasonably good approximation compared to results from the full-order model. Moreover, the reduced-order model with $r=10$ modes offers an even more accurate approximation.

The energy dissipation across different modes is illustrated in Fig. 3(a), which illustrates that the reduced-order model with $r=4$ modes effectively captures the energy dissipation trends, while the one with $r=10$ modes captures the energy dissipation more accurately. Additionally, we present the energy evolution using both approach I and approach II with $r=10$ modes, as shown in Fig. 3(b). It appears that approach II provides more accurate results. As we discussed, approach I is not as accurate as approach II since it preserves an energy dissipation law with a modified energy dissipation rate. The relative energy error and error for ϕ under different reduced order modes are also shown in Fig. 3(c)-(d). The relative error for the energy is calculated by $\frac{E_{ROM}(t) - E_{FOM}(t)}{E_{FOM}(t)}$, where $E_{ROM}(t)$ and $E_{FOM}(t)$ represent the energy for the ROM and full order model at time t respectively. The relative l^2 error for ϕ is calculated by $\frac{\|\phi_{ROM} - \phi_{FOM}\|}{\|\phi_{FOM}\|}$. The results indicate that the error decreases with the increase in the number of modes.

5.2. Cahn-Hilliard equation

In the next problem, we investigate the Cahn-Hilliard (CH) equation

$$\partial_t \phi = M \Delta \mu, \quad (\mathbf{x}, t) \in \Omega \times (0, T], \quad (5.6a)$$

$$\mu = -\epsilon^2 \Delta \phi + \phi^3 - \phi, \quad (\mathbf{x}, t) \in \Omega \times (0, T], \quad (5.6b)$$

$$\phi(\mathbf{x}, 0) = \phi_0(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (5.6c)$$

with periodic boundary conditions. The Onsager triplet for the CH equation is

$$(\phi, \mathcal{G}, \mathcal{E}) := \left(\phi, -M \Delta, \int_{\Omega} \left[\frac{\epsilon^2}{2} |\nabla \phi|^2 + \frac{1}{4} (\phi^2 - 1)^2 \right] d\mathbf{x} \right).$$

Similarly, we can introduce

$$q = \frac{\sqrt{2}}{2} (\phi^2 - 1 - \gamma_0), \quad g(\phi) := \frac{\partial q}{\partial \phi} = \sqrt{2} \phi.$$

And we will have the equation rewritten as

$$\partial_t \phi = M \Delta \left(-\epsilon^2 \Delta \phi + q g(\phi) \right), \quad (5.7a)$$

$$\partial_t q = g(\phi) \partial_t \phi. \quad (5.7b)$$

Denote $\Psi = \begin{bmatrix} \phi \\ q \end{bmatrix}$, $\mathcal{G}_s = -M \Delta$, $\mathcal{L}_0 = -\epsilon^2 \Delta + \gamma_0$, $\mathcal{L} = \begin{bmatrix} \mathcal{L}_0 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathcal{N}_0 = [\mathbf{I} \ g(\phi)]$, such that the equation is written as

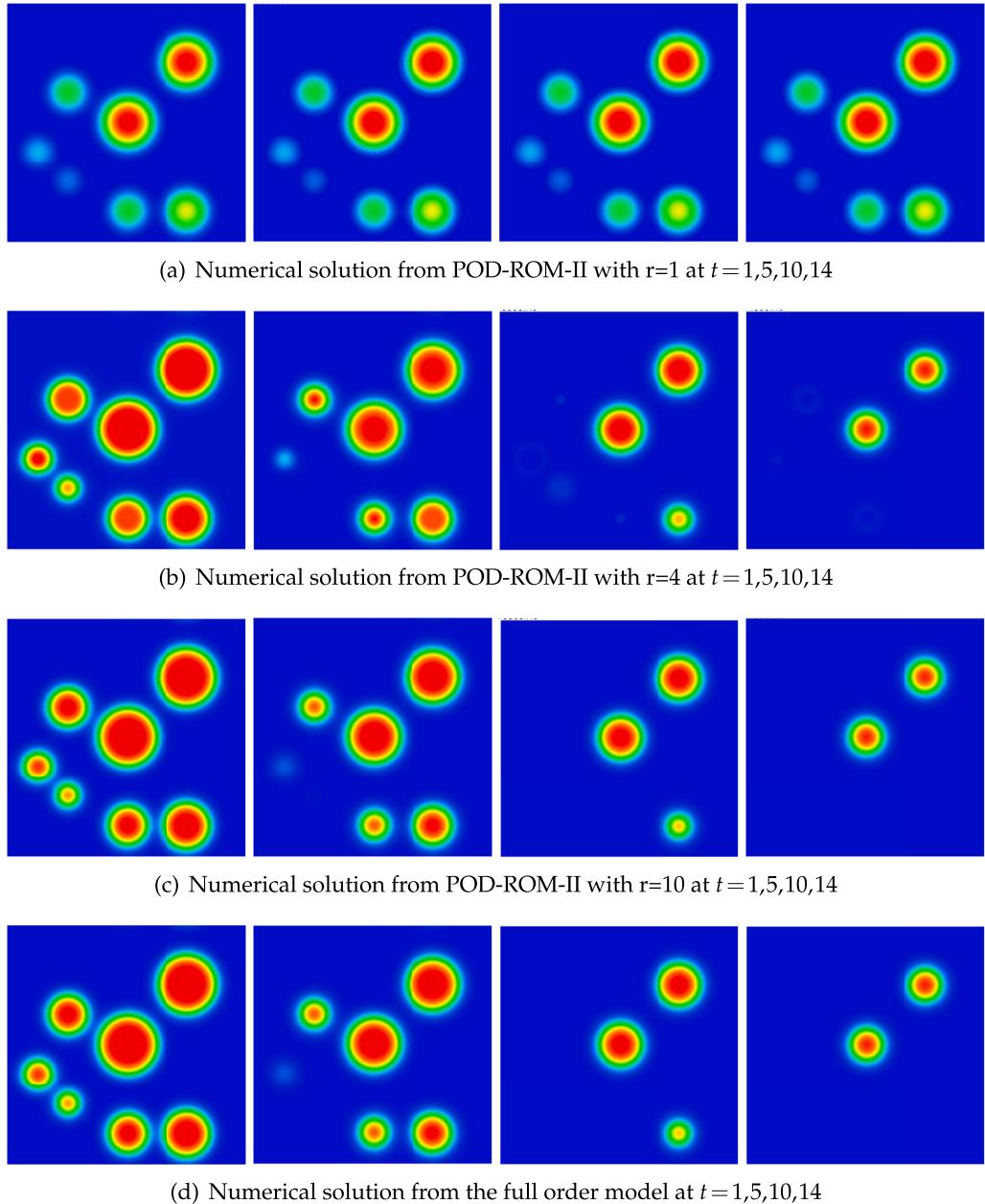


Fig. 2. A comparison of the numerical solutions between the full order model and the POD-ROM-II with various numbers of modes using Scheme 4.5.

$$\partial_t \Psi = -\mathcal{N}(\Psi) \mathcal{L} \Psi, \quad \text{where} \quad \mathcal{N}(\Psi) = \mathcal{N}_0^T \mathcal{G}_s \mathcal{N}_0. \quad (5.8)$$

Hence, the POD-ROM method introduced in previous sections can be directly applied.

Example 2. In this example, we choose the model parameters $M = 0.01$ and $\varepsilon = 0.02$. The domain $[0, 1]^2$ is considered. The numerical parameters used are $\gamma_0 = 2$ and $N_x = N_y = 128$. The initial condition is chosen the same as (5.4). This represents seven disks of various sizes in different locations of the domain. Consider $T = 15$, and $\delta t = 10^{-3}$. We first generate the data using an accurate numerical solver. In this case, we use a classical convex splitting scheme to solve the equation with adequate time step $\delta t = 10^{-3}$, which is used as the “accurate” solution. Then we collect the data $[\Phi_1 \ \Phi_2 \ \dots \ \Phi_m]$, where $\Phi_k \in \mathbb{R}^n$ with $n = N_x N_y$ is the numerical solution at $t = 0.1k$ in a vector form. In this example, we choose $N_x = N_y = 128$ and $m = 150$. The singular value distribution is summarized in Fig. 4, where the percentage of each singular value relative to the total sum of singular values is calculated.

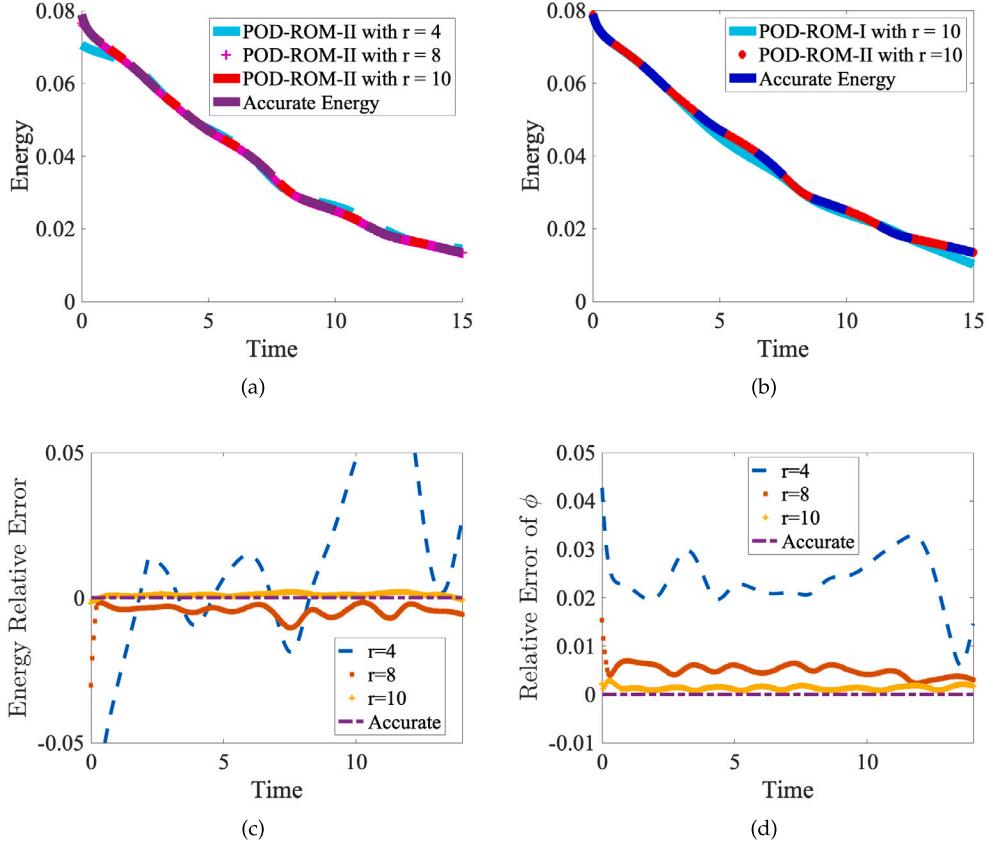


Fig. 3. A comparison of the energy dissipation for the Allen-Cahn equation. (a) A comparison of the energy dissipation results between the full order model and POD-ROM-II with various modes using Scheme 4.5; (b) A comparison of the energy dissipation results between POD-ROM-I using Scheme 4.7 and POD-ROM-II using Scheme 4.5; (c) the relative error for the energy; and (d) the relative l^2 error for ϕ . This figure illustrates the proficiency of the proposed POD-ROM approach in capturing the thermodynamic structures of the full-order model with only a few modes. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

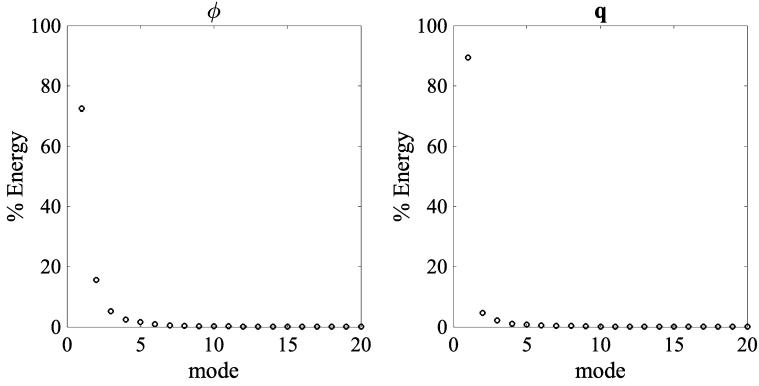


Fig. 4. Singular value distributions for the data collected from the Cahn-Hilliard equation.

Then, we follow the POD-ROM numerical framework proposed in previous sections. The numerical results are summarized in Fig. 5. It illustrates that the reduced-order model with $r = 10$ modes can have an accurate approximation of the coarsening dynamics. The numerical results with $r = 15$ modes show similar dynamics to the results from the full order model.

Furthermore, we have summarized the results of the energy dissipation for POD-ROM-II using different modes in Fig. 6(a). It illustrates that the reduced-order model with $r = 10$ modes can accurately capture the Cahn-Hilliard equation's energy dissipation. Additionally, from Fig. 6(b), we observe that POD-ROM-II provides a more accurate prediction than POD-ROM-I for energy dissipation. Similarly as the example for the Allen-Cahn equation, we also summarize the relative energy error and the error for ϕ in Fig. 6(c) and 6(d). We observe that with the increase in the mode r , the error is decreasing.

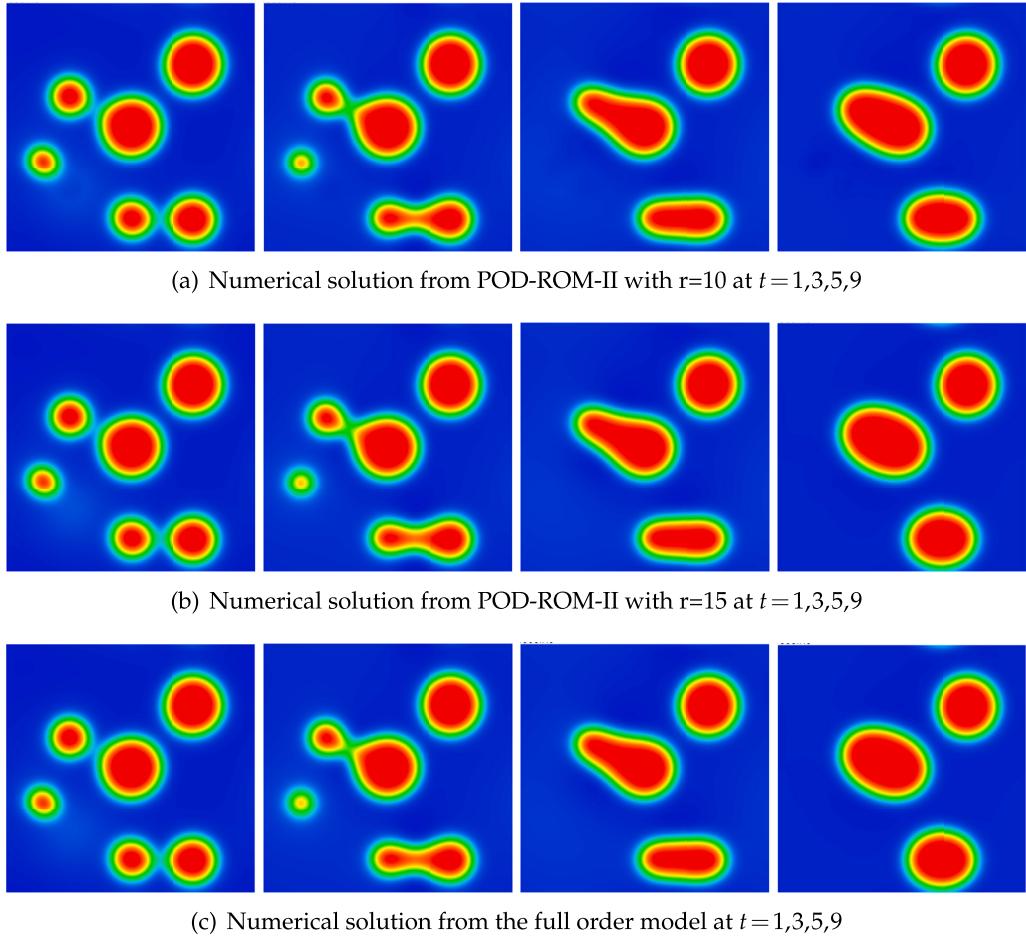


Fig. 5. A comparison between the numerical solutions of the Cahn-Hilliard equation from the full order model and the numerical solutions from the POD-ROM-II with various numbers of modes.

5.3. Phase field crystal equation

Next, we investigate the phase field crystal (PFC) model with our proposed reduced-order model techniques. The PFC model reads as

$$\partial_t \phi = M \Delta \left[(a_0 + \Delta)^2 \phi + f'(\phi) \right], \quad (5.9)$$

where $f(\phi) = \frac{1}{4}\phi^4 - \frac{b_0}{2}\phi^2$. Here a_0 and b_0 are model parameters. The PFC model can also be derived from the generalized Onsager principle with the Onsager triplet

$$(\phi, \mathcal{G}, \mathcal{E}) := \left(\phi, -M \Delta, \int_{\Omega} \left[\frac{1}{2} \phi (-b_0 + (a_0 + \Delta)^2) \phi + \frac{1}{4} \phi^4 \right] d\mathbf{x} \right).$$

Introduce the auxiliary variables

$$q(\mathbf{x}, t) := \frac{\sqrt{2}}{2} (\phi^2 - b_0 - \gamma_0), \quad g(\phi) := \frac{\partial q}{\partial \phi} = \sqrt{2} \phi,$$

the PFC model could be transformed as

$$\partial_t \phi = M \Delta \left((a_0 + \Delta)^2 \phi + \gamma_0 \phi + g(\phi) q \right), \quad (5.10a)$$

$$\partial_t q = g(\phi) \partial_t \phi. \quad (5.10b)$$

Denote $\Psi = \begin{bmatrix} \phi \\ q \end{bmatrix}$, $\mathcal{G}_s = -M \Delta$, $\mathcal{L}_0 = (a_0 + \Delta)^2 + \gamma_0$, $\mathcal{L} = \begin{bmatrix} \mathcal{L}_0 & 0 \\ 0 & \mathbf{I} \end{bmatrix}$ and $\mathcal{N}_0 = [\mathbf{I} \quad g(\phi)]$, such that the equation is written as

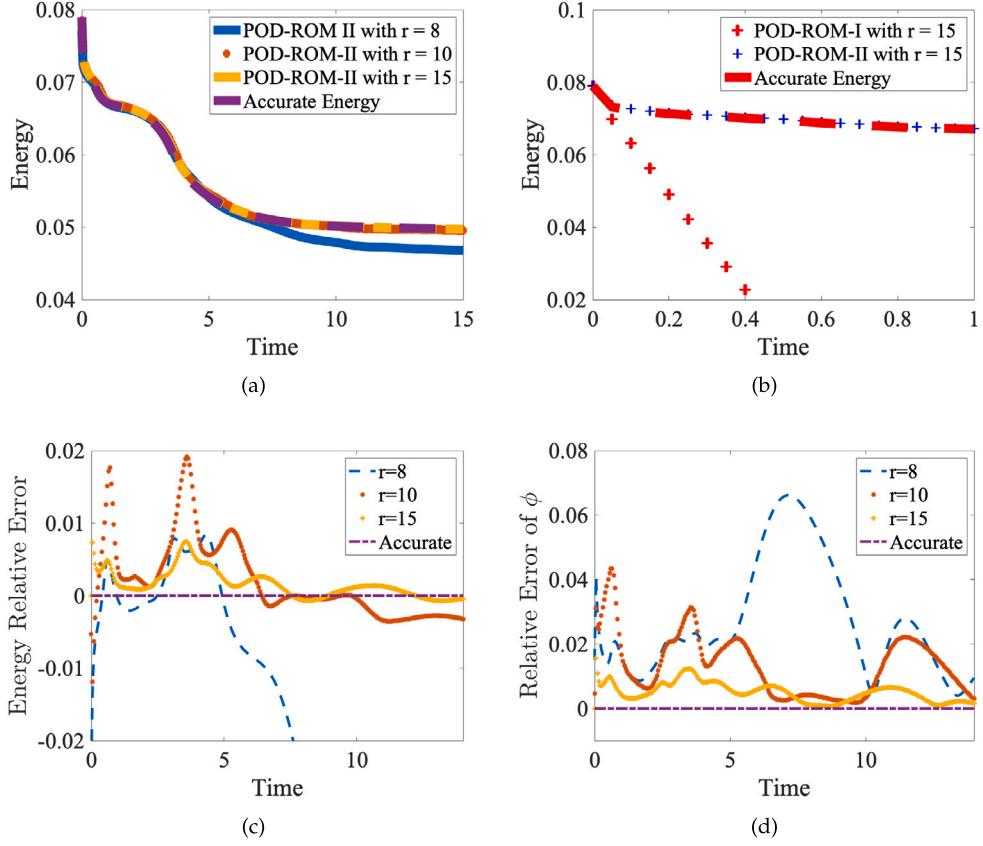


Fig. 6. A comparison of the energy dissipation for the Cahn-Hilliard equation. (a) A comparison of the energy dissipation results between the full order model and the POD-ROM-II with various modes using Scheme 4.5; (b) A comparison of the energy dissipation results between POD-ROM-I using Scheme 4.7 and POD-ROM-II using Scheme 4.5; (c) the relative error for the energy; and (d) the relative error for ϕ .

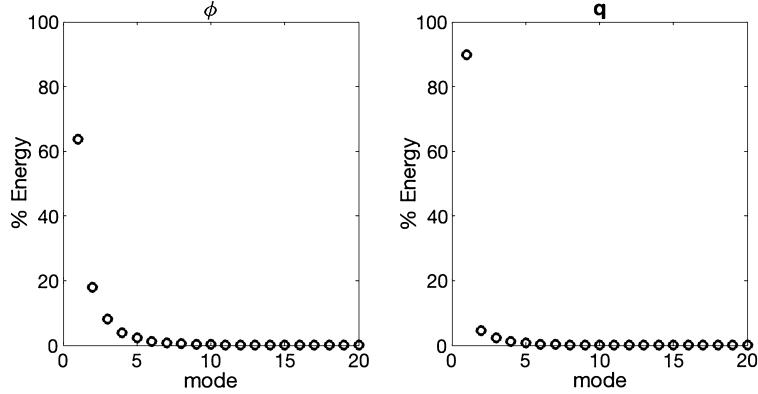


Fig. 7. Singular value distributions for the data collected from the phase field crystal equation.

$$\partial_t \Psi = -\mathcal{N}(\Psi) \mathcal{L} \Psi, \quad \text{where} \quad \mathcal{N}(\Psi) = \mathcal{N}_0^T \mathcal{G}_s \mathcal{N}_0. \quad (5.11)$$

Hence, the POD-ROM method introduced in previous sections can be directly applied.

Example 3. In the numerical example, we choose $a_0 = 1$, $b_0 = 0.325$, $\delta t = 10^{-3}$, $N_x = N_y = 128$, $\gamma_0 = 1$, $M = 1$, and $T = 100$. We choose the domain $\Omega = [0, 100]^2$. We initialize three dots in the domain as the initial profile, following the strategy in [39]. The singular value distribution is summarized in Fig. 7, where the percentage of each singular value relative to the total sum of singular values is calculated.

The results are summarized in Fig. 8. We can observe that with $r = 4$ modes, the ROM is able to capture the dynamics properly. Furthermore, with $r = 8$ modes, the ROM model captures the dynamics accurately.

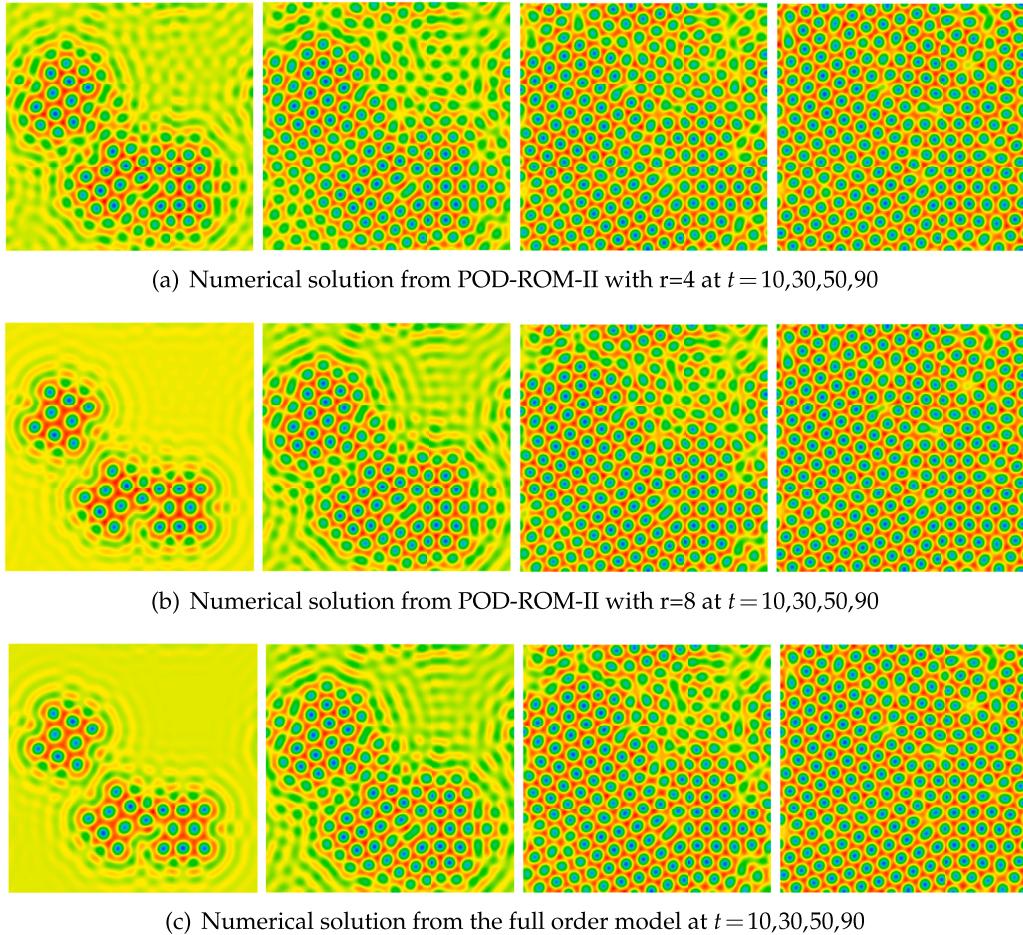


Fig. 8. A comparison between the numerical solutions from the full model and the numerical solutions from POD-ROM-II for the phase field crystal equations with various numbers of modes.

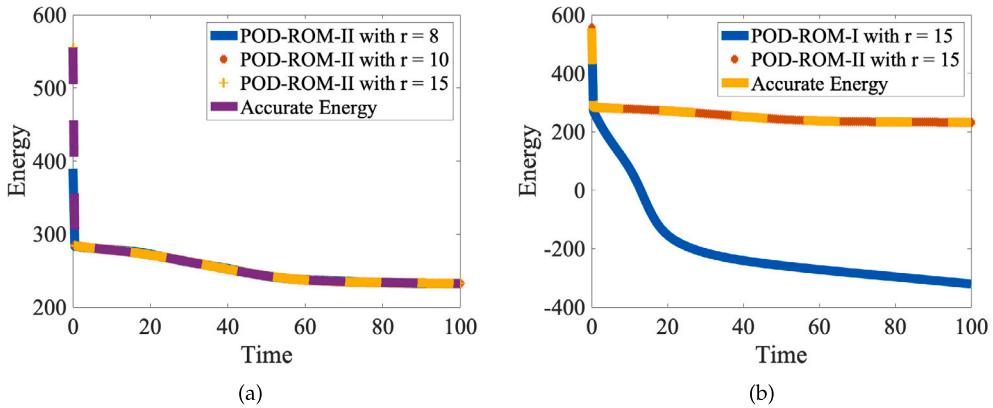


Fig. 9. A comparison of the energy dissipation for the phase field crystal equation. (a) A comparison of the energy dissipation results between the full order model and the POD-ROM-II with various numbers of modes using Scheme 4.5; (b) A comparison of the energy dissipation results between POD-ROM-I using Scheme 4.7 and POD-ROM-II using Scheme 4.5.

Then, we compare the energy dissipation curves for POD-ROM-II with different modes, as shown in Fig. 9(a). It appears that the POD-ROM-II can accurately predict the energy evolution with $r = 10$ modes already. Furthermore, we compare the results between POD-ROM-I and POD-ROM-II shown in Fig. 9(b), and it appears that POD-ROM-II can provide a more accurate prediction for the energy evolution than POD-ROM-I with the same modes. This is reasonable since POD-ROM-I has a modified energy dissipation rate.

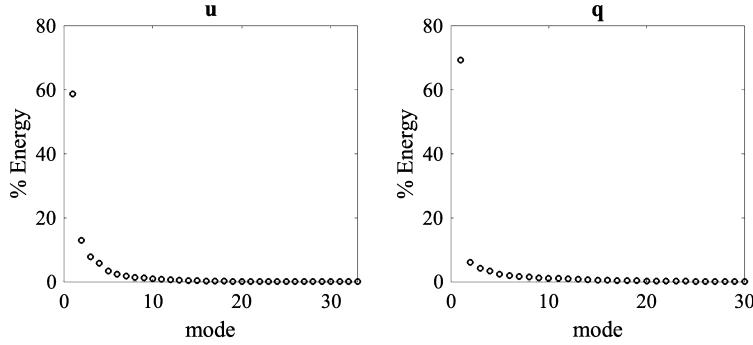


Fig. 10. Singular value distributions for the data collected from the sine-Gordon equation.

5.4. Sine-Gordon equation

So far, we have applied our general framework to several dissipative (irreversible) systems to demonstrate its effectiveness. In the last example, we consider a dispersive (reversible) system. Namely, we consider the following sine-Gordon equation in 2D

$$\partial_{tt}u = \Delta u - \sin(u), \quad (x, y) \in \Omega \subset \mathbb{R}^2, \quad 0 < t \leq T, \quad (5.12a)$$

$$u(x, y, 0) = g_1(x, y), \quad u_t(x, y, 0) = g_2(x, y), \quad (x, y) \in \Omega \subset \mathbb{R}^2, \quad (5.12b)$$

with periodic boundary conditions. It has the following energy conservation law

$$\frac{d}{dt}H(u) = 0, \quad H(u) = \int_{\Omega} \left[\frac{1}{2}u_t^2 + \frac{1}{2}|\nabla u|^2 + 1 - \cos(u) \right] d\mathbf{x}.$$

If we introduce the intermediate variable $v = u_t$, the sine-Gordon equation is rewritten as

$$\partial_t \Phi = -\mathcal{G} \frac{\delta H(\Phi)}{\delta \Phi}, \quad \Phi = \begin{bmatrix} u \\ v \end{bmatrix}, \quad \mathcal{G} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

with the energy $H(\Phi) := H(u, v) = \int_{\Omega} \left[\frac{1}{2}v^2 + \frac{1}{2}|\nabla u|^2 + 1 - \cos(u) \right] d\mathbf{x}$. Furthermore, the sine-Gordon equation can be cast into the generalized Onsager form with the triplet

$$(\phi, \mathcal{G}, \mathcal{E}) := \left(\begin{bmatrix} u \\ v \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \int_{\Omega} \left[\frac{1}{2}v^2 + \frac{1}{2}|\nabla u|^2 + 1 - \cos(u) \right] d\mathbf{x} \right).$$

Then, we introduce the auxiliary variables

$$q = \sqrt{2} \sqrt{1 + A_0 - \cos(u)}, \quad g(u) := \frac{\partial q}{\partial u} = \frac{1}{\sqrt{2}} \frac{\sin(u)}{\sqrt{1 + A_0 - \cos(u)}},$$

with A_0 a regularization constant [11]. The sine-Gordon equation is reformulated into the Onsager-Q form

$$\partial_t \Psi = -\mathcal{N}(\Psi) \mathcal{L} \Psi, \quad \mathcal{N}(\Psi) = \mathcal{N}_0^T \mathcal{G}_a \mathcal{N}_0,$$

where

$$\mathcal{G}_a = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \mathcal{L} = \begin{bmatrix} -\Delta & 0 & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & 0 & \mathbf{I} \end{bmatrix}, \quad \mathcal{N}_0 = \begin{bmatrix} \mathbf{I} & 0 & g(u) \\ 0 & \mathbf{I} & 0 \end{bmatrix},$$

with the modified energy $H(\Psi) = \frac{1}{2}(\Psi, \mathcal{L}\Psi) - A_0|\Omega|$, i.e.,

$$H(u, v, q) = \frac{1}{2} \int_{\Omega} \left[|\nabla u|^2 + v^2 + q^2 \right] d\mathbf{x} - A_0|\Omega|.$$

Therefore, the proposed structure-preserving numerical techniques can be used to solve the sine-Gordon equation. Notice that the relaxation step only applies to the dissipative system. For dispersive systems like the sine-Gordon equation, we can directly utilize Scheme 4.1 and Scheme 4.2.

Example 4. In the numerical example, we follow the setup in [7, 25, 41], by considering the 2D domain $\Omega = [-30, 10] \times [-21, 7]$ with the following initial conditions

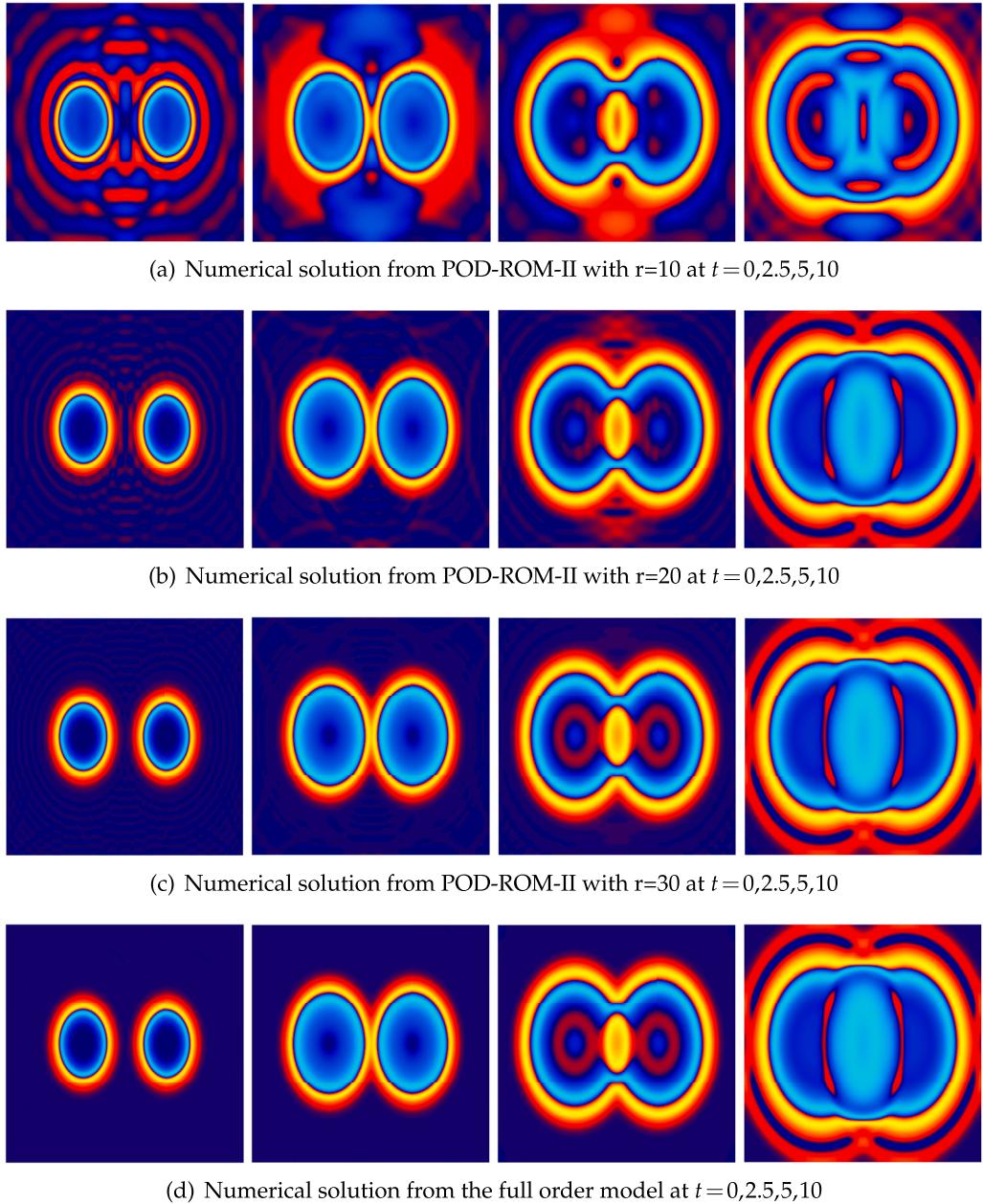


Fig. 11. A comparison between the numerical solutions from the full model and the numerical solutions from POD-ROM-II using Scheme 4.1 for the sine-Gordon equation with various numbers of modes. The profile $\sin(\frac{u}{2})$ are shown at $t = 0, 2.5, 5$ and 10 .

$$g_1(x, y) = 4 \arctan \left[\exp \left(\frac{4 - \sqrt{(x+3)^2 + (y+7)^2}}{0.436} \right) \right], \quad g_2(x, y) = 4.13 \operatorname{sech} \left(\frac{4 - \sqrt{(x+3)^2 + (y+2)^2}}{0.436} \right),$$

with a symmetric extension across $x = -10$ to create the two circular solutions.

We choose spatial meshes $N_x = N_y = 256$, a time step size $\delta t = 10^{-4}$, and $T = 15$. The singular value distribution is summarized in Fig. 10, where the percentage of each singular value relative to the total sum of singular values is calculated.

The numerical results are summarized in Fig. 11. We can observe that with $r = 20$ modes, the ROM is able to capture a similar pattern of the dynamics. With $r = 30$ modes, the ROM model captures the dynamics accurately.

To further investigate the numerical accuracy of the POD-ROM II approach, we present the relative ℓ^2 error for the numerical solutions u and v under various number of modes. As shown in Fig. 12, with the number of mode r increasing, the error is decreasing and getting close to zero.

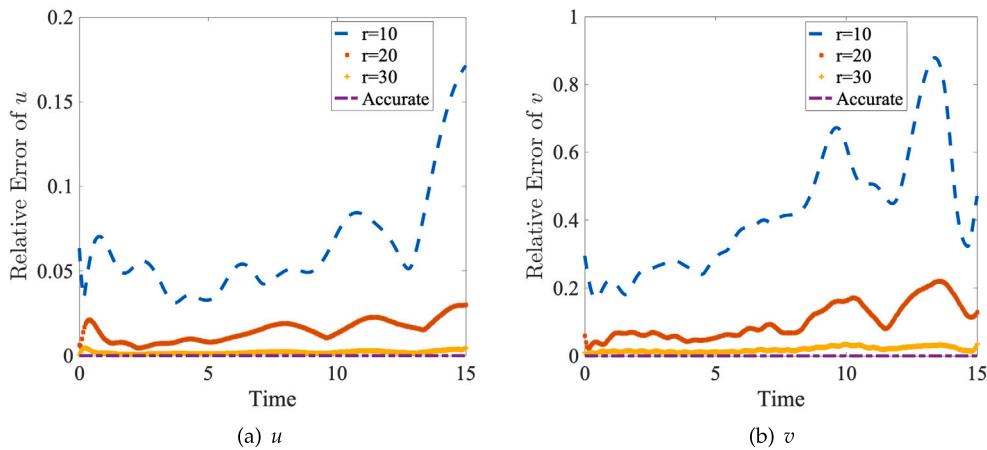


Fig. 12. A comparison of the relative l^2 error of the numerical solutions u and v under various modes: $r = 10, 20, 30$.

6. Conclusion and future work

In this paper, we introduce a general numerical framework to develop structure-preserving reduced-order models (ROMs) for thermodynamically consistent reversible-irreversible PDEs. Our numerical framework is general and provides a unified approach to develop structure-preserving reduced-order models for PDE systems with free energy dissipation laws. An extension of the current approach to investigate non-isothermal thermodynamic PDE models is possible. Meanwhile, there are still several open questions. For instance, the maximum principle for the Allen-Cahn equation is well-studied. How to derive a reduced-order model to preserve the maximum principle while preserving the energy dissipation law is still unclear. Additionally, for some models, mass conservation is essential, saying the Cahn-Hilliard equation, which is the main reason making it different from the Allen-Cahn equation. The current framework can't guarantee mass conservation and energy dissipation simultaneously. These open questions are to be addressed in our subsequent works.

CRediT authorship contribution statement

Zengyan Zhang: Writing – original draft, Visualization, Validation, Investigation, Formal analysis. **Jia Zhao:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Zengyan Zhang and Jia Zhao would like to acknowledge the support from the National Science Foundation with grant NSF-DMS-2111479/2405605. They would also like to acknowledge NVIDIA Corporation for the donation of GPUs to conduct some of the numerical simulations in this paper.

Data availability

No data was used for the research described in the article.

References

- [1] B. Afkham, J. Hesthaven, Structure-preserving model-reduction of dissipative Hamiltonian systems, *J. Sci. Comput.* 81 (2019) 3–21.
- [2] S.M. Allen, Ground state structures in ordered binary alloys with second neighbor interactions, *Acta Metall.* 20 (1972) 423–433.
- [3] P. Astrid, Reduction of process simulation models: a proper orthogonal decomposition approach, PhD thesis, Eindhoven University of Technology, 2004.
- [4] M. Barrault, Y. Maday, N.C. Nguyen, A.T. Patera, An “empirical interpolation” method: application to efficient reduced-basis discretization of partial differential equations, *C. R. Math.* 339 (9) (2004) 667–672.
- [5] P. Benner, T. Breiten, Model order reduction based on system balancing, in: *Model Reduction and Approximation*, 2017, pp. 261–295.
- [6] A.N. Beris, B. Edwards, *Thermodynamics of Flowing Systems: with Internal Microstructure*, Oxford Science Publications, New York, 1994.
- [7] A.G. Bratsos, The solution of the two-dimensional sine-Gordon equation using the method of lines, *J. Comput. Appl. Math.* 206 (2007) 251–277.
- [8] J.W. Cahn, J.E. Hilliard, Free energy of a nonuniform system. I. Interfacial free energy, *J. Chem. Phys.* 28 (1958) 258–267.

[9] K. Carlberg, C. Bou-Mosleh, C. Farhat, Efficient non-linear model reduction via a least-squares Petrov-Galerkin projection and compressive tensor approximations, *Int. J. Numer. Methods Eng.* 86 (2) (2011) 155–181.

[10] S. Chaturantabut, D.C. Sorensen, Nonlinear model reduction via discrete empirical interpolation, *SIAM J. Sci. Comput.* 32 (5) (2010) 2737–2764.

[11] L. Chen, J. Zhao, X. Yang, Regularized linear schemes for the molecular beam epitaxy model with slope selection, *Appl. Numer. Math.* 128 (2018) 138–156.

[12] Q. Cheng, C. Liu, J. Shen, A new Lagrange multiplier approach for gradient flows, *Comput. Methods Appl. Mech. Eng.* (2020) 113070.

[13] Y. Choi, Nonlinear manifold-based reduced order model, Technical report, Lawrence Livermore National Lab. (LLNL), Livermore, CA (United States), 2020.

[14] T. Daniel, F. Casenave, N. Akkari, D. Ryckelynck, Model order reduction assisted by deep neural networks (ROM-net), *Adv. Model. Simul. Eng. Sci.* 7 (16) (2020) 1–27.

[15] H. Egger, Structure preserving approximation of dissipative evolution problems, *Numer. Math.* 143 (2019) 85–106.

[16] K.R. Elder, M. Grant, Modeling elastic and plastic deformations in nonequilibrium processing using phase field crystals, *Phys. Rev. E* 70 (051605) (2004).

[17] D. Eyre, Unconditionally gradient stable time marching the Cahn-Hilliard equation, in: *Computational and Mathematical Models of Microstructural Evolution*, vol. 529, San Francisco, CA, 1998, 1998, pp. 39–46.

[18] D. Furihata, T. Matsuo, *Discrete Variational Derivative Method*, CR2C Press, 2010.

[19] H. Gomez, X. Nogueira, An unconditionally energy-stable method for the phase field crystal equation, *Comput. Methods Appl. Mech. Eng.* 249 (2012) 52–61.

[20] Y. Gong, Q. Hong, Q. Wang, Supplementary variable method for thermodynamically consistent partial differential equations, *Comput. Methods Appl. Mech. Eng.* (2021) 321–339.

[21] Y. Gong, Q. Wang, Z. Wang, Structure-preserving Galerkin POD reduced-order modeling of Hamiltonian systems, *Comput. Methods Appl. Mech. Eng.* 315 (2017) 780–798.

[22] Y. Gong, J. Zhao, Q. Wang, Linear second order in time energy stable schemes for hydrodynamic models of binary mixtures based on a spatially pseudospectral approximation, *Adv. Comput. Math.* 44 (2018) 1573–1600.

[23] C. Grable, M. Hinze, Pod reduced-order modeling for evolution equations utilizing arbitrary finite element discretizations, *Adv. Comput. Math.* 44 (2018) 1941–1978.

[24] J. Guo, C. Wang, S. Wise, X. Yue, An H^2 convergence of a second-order convex-splitting, finite difference scheme for the three-dimensional Cahn-Hilliard equation, *Commun. Math. Sci.* 14 (2) (2015) 489–515.

[25] C. Jiang, W. Cai, Y. Wang, A linearly implicit and local energy-preserving scheme for the sine-Gordon equation based on the invariant energy quadratization approach, *J. Sci. Comput.* 80 (2019) 1629.

[26] M. Jiang, Z. Zhang, J. Zhao, Improving the accuracy and consistency of the scalar auxiliary variable (SAV) method with relaxation, *J. Comput. Phys.* 456 (110954) (2022).

[27] B. Karasozan, T. Kucukseyhan, M. Uzunca, Structure preserving integration and model order reduction of skew-gradient reaction–diffusion systems, *Ann. Oper. Res.* 258 (2017) 79–106.

[28] Karl Kunisch, Stefan Volkwein, Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics, *SIAM J. Numer. Anal.* 40 (2) (2002) 492–515.

[29] H. Li, D. Wang, Z. Song, F. Zhang, Numerical analysis of an unconditionally energy-stable reduced-order finite element method for the Allen-Cahn phase field model, *Comput. Math. Appl.* 96 (2021) 67–76.

[30] J.L. Lumley, The structure of inhomogeneous turbulent flows, in: *Atmospheric Turbulence and Radio Wave*, 1967.

[31] H.C. Ottinger, *Beyond Equilibrium Thermodynamics*, Wiley, 2005.

[32] J.S. Peterson, The reduced basis method for incompressible viscous flow calculations, *SIAM J. Sci. Stat. Comput.* 10 (4) (1989) 777–786.

[33] M. Rewienski, J. White, Model order reduction for nonlinear dynamical systems based on trajectory piecewise-linear approximations, *Linear Algebra Appl.* 415 (2–3) (2006) 426–454.

[34] J. Shen, J. Xu, J. Yang, A new class of efficient and robust energy stable schemes for gradient flows, *arXiv:1710.01331*, 2017.

[35] J. Shen, J. Xu, J. Yang, The scalar auxiliary variable (SAV) approach for gradient flows, *J. Comput. Phys.* 353 (2018) 407–416.

[36] J. Shen, X. Yang, Numerical approximation of Allen-Cahn and Cahn-Hilliard equations, *Discrete Contin. Dyn. Syst., Ser. B* 28 (4) (2010) 1669–1691.

[37] H. Song, L. Liang, Q. Li, A reduced order method for Allen-Cahn equations, *J. Comput. Appl. Math.* 292 (2016) 213–229.

[38] M. Uzunca, B. Karasozan, S. Yildiz, Structure-preserving reduced-order modeling of Korteweg-de Vries equation, *Math. Comput. Simul.* 188 (2021) 193–211.

[39] P. Vignal, L. Dalcin, D. Brown, N. Collier, V. Calo, An energy stable convex splitting for the phase-field crystal equation, *Comput. Struct.* 158 (2015) 355–368.

[40] C. Wang, X. Wang, S. Wise, Unconditionally stable schemes for equations of thin film epitaxy, *Discrete Contin. Dyn. Syst.* 28 (1) (2010) 405–423.

[41] J. Wang, Q. Huang, A family of effective structure-preserving schemes with second-order accuracy for the undamped sine–Gordon equation, *Comput. Math. Appl.* 90 (2021) 38–45.

[42] Z. Wang, Structure-preserving Galerkin POD-DEIM reduced-order modeling of Hamiltonian systems, *arXiv:2103.00388v2*, 2021.

[43] Julien Weiss, A tutorial on the proper orthogonal decomposition, in: *AIAA Aviation 2019 Forum*, 2019.

[44] K. Willcox, Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition, *Comput. Fluids* 35 (2006) 208–226.

[45] S. Wise, C. Wang, J.S. Lowengrub, An energy-stable and convergent finite-difference scheme for the phase field crystal equation, *SIAM J. Numer. Anal.* 47 (3) (2009) 2269–2288.

[46] X. Yang, J. Zhao, Q. Wang, Numerical approximations for the molecular beam epitaxial growth model based on the invariant energy quadratization method, *J. Comput. Phys.* 333 (2017) 102–127.

[47] T. Yildiz, M. Uzunca, B. Karasozan, Structure preserving reduced order modeling for gradient systems, *Appl. Math. Comput.* 347 (2019) 194–209.

[48] J. Zhao, A revisit of the energy quadratization method with a relaxation technique, *Appl. Math. Lett.* 120 (107331) (2021).

[49] J. Zhao, X. Yang, Y. Gong, X. Zhao, J. Li, X. Yang, Q. Wang, A general strategy for numerical approximations of thermodynamically consistent nonequilibrium models-part I: thermodynamical systems, *Int. J. Numer. Anal. Model.* 15 (6) (2018) 884–918.

[50] X. Zhou, M. Zaiez, C. Xu, Reduced-order modelling for the Allen-Cahn equation based on scalar auxiliary variable approaches, *J. Math. Study* 53 (3) (2019) 258–276.