

Evolution of nonlinear reduced-order solutions for PDEs with conserved quantities

William Anderson and Mohammad Farazmand*

Department of Mathematics, North Carolina State University,
2311 Stinson Drive, Raleigh, NC 27695-8205, USA

Abstract

Reduced-order models of time-dependent partial differential equations (PDEs) where the solution is assumed as a linear combination of prescribed modes are rooted in a well-developed theory. However, more general models where the reduced solutions depend nonlinearly on time varying parameters have thus far been derived in an ad hoc manner. Here, we introduce Reduced-order Nonlinear Solutions (RONS): a unified framework for deriving reduced-order models that depend nonlinearly on a set of time-dependent parameters. The set of all possible reduced-order solutions are viewed as a manifold immersed in the function space of the PDE. The parameters are evolved such that the instantaneous discrepancy between reduced dynamics and the full PDE dynamics is minimized. This results in a set of explicit ordinary differential equations on the tangent bundle of the manifold. In the special case of linear parameter dependence, our reduced equations coincide with the standard Galerkin projection. Furthermore, any number of conserved quantities of the PDE can readily be enforced in our framework. Since RONS does not assume an underlying variational formulation for the PDE, it is applicable to a broad class of problems. We demonstrate the efficacy of RONS on three examples: an advection-diffusion equation, the nonlinear Schrödinger equation and Euler's equation for ideal fluids.

1 Introduction

Reduced-order models are routinely used to facilitate computational and mathematical analysis of nonlinear partial differential equations (PDEs). The theory is well-developed when the reduced-order solution $\hat{u}(\mathbf{x}, t) = \sum_i q_i(t)u_i(\mathbf{x})$ is a linear combination of time-independent modes u_i (see Refs. [10, 31], for exhaustive reviews). However, a robust mathematical framework is missing when

*Corresponding author's email address: farazmand@ncsu.edu

the model $\hat{u}(\mathbf{x}, \mathbf{q}(t))$ depends nonlinearly on a set of time-dependent variables $\mathbf{q} = (q_1, q_2, \dots, q_n)$.

This is despite the fact that such nonlinear reduced-order solutions are ubiquitously used, e.g., in prediction of rogue waves [17, 23, 33], vortex methods in fluid dynamics [8, 15], tracking shocks in supersonic flows [19], and shape optimization [24, 35], to name a few. In these studies, the reduced solutions depend nonlinearly on amplitudes, length scales, traveling speed, phase, etc. Currently, these parameters are evolved in an ad hoc manner based on domain expertise and familiarity with the underlying PDE. The purpose of this paper is to propose a unified framework that is broadly applicable to time-dependent nonlinear PDEs.

Nonlinear reduced-order solutions $\hat{u}(\mathbf{x}, \mathbf{q}(t))$ are often better suited for efficiently quantifying the dynamics, as compared to their linear counterpart $\sum_i q_i(t)u_i(\mathbf{x})$. As a rudimentary example, consider the heat equation in d dimensions, $\partial_t u = \Delta u$, and its fundamental solution

$$u(\mathbf{x}, \mathbf{q}(t)) = A(t) \exp\left(-\frac{|\mathbf{x}|^2}{L^2(t)}\right), \quad (1)$$

where $A(t) = (4\pi t)^{-d/2}$ and $L(t) = \sqrt{4t}$. Here, the parameters are $\mathbf{q} = (A, L)$ where the solution depends linearly on the amplitude A and nonlinearly on the length scale L . We use this example throughout the paper for illustrative purposes. But, more importantly, it already showcases the potency of reduced-order models that depend nonlinearly on time-dependent parameters. If we were to approximate solution (1) as the linear combination of time-independent modes u_i , i.e., $\sum_i q_i(t)u_i(\mathbf{x})$, several modes would be required to obtain a reasonable approximation. On the other hand, the solution can be expressed with a single mode (1) when allowing nonlinear dependence on time-dependent parameters.

Here, we develop a unified framework for evolving the reduced-order solutions $\hat{u}(\mathbf{x}, \mathbf{q}(t))$ which depend nonlinearly on time-dependent parameters $\mathbf{q}(t)$. For a prescribed dependence on the spatial variable \mathbf{x} , we view \hat{u} as a map from the parameter space to the function space H of the PDE. The image of this map is a manifold immersed in H . We require the evolution of the parameters $\mathbf{q}(t)$ to minimize the instantaneous discrepancy between the reduced-order dynamics and full dynamics of the PDE. This leads to a set of nonlinear ordinary differential equations (ODEs) for $\mathbf{q}(t)$, defined on the tangent bundle of the manifold. In our framework, any number of conserved quantities of the PDE can be readily enforced. Since it does not rely on the variational formulation of the PDE, our method is broadly applicable. We also show that when \hat{u} depends linearly on the parameter \mathbf{q} , our reduced-order equations coincide with the standard Galerkin projection models. We refer to the proposed method as RONS: Reduced-Order Nonlinear Solutions.

We point out that the method presented here should not be confused with *parametric model reduction* [10] where the PDE itself depends on some parameters. In contrast, in RONS, the parametric dependence appears in the reduced-order solution. Furthermore, the term *nonlinear model reduction* is often used for reduced-order modeling of nonlinear PDEs even when the reduced

solution is a linear combination of modes. In this paper, in addition to the PDE being nonlinear, the reduced solution is also a nonlinear function of the reduced variables.

1.1 Related work

This paper introduces a novel and unified framework for evolving time-dependent parameters of reduced-order nonlinear solutions. However, as mentioned earlier, the idea of using such nonlinear ansatz is not new. Here, we review some of the work which use nonlinear ansatz and the respective ad hoc methods for evolving their parameters.

The propagation of optical beams in nonlinear dispersive media has been studied by considering the evolution of a localized Gaussian ansatz with time-dependent amplitude, length scale, and phase [11, 20, 30]. A similar approach has also been taken in the context of nonlinear water waves by Ruban [33, 34] and Adcock et al. [1, 2]. Following [20], Ruban [33, 34] uses the Lagrangian associated with the nonlinear Schrödinger equation (NLSE) to derive reduced-order equations for the parameters of the ansatz. Adcock et al. [1, 2], on the other hand, use the conserved quantities of NLSE to evolve the parameters. Cousins and Sapsis [16] consider the modified NLSE which does not have a known Lagrangian structure or as many conserved quantities as NLSE. They use a hyperbolic secant as the ansatz, take a second time derivative of the modified NLSE, and project the resulting equation on the subspace of the ansatz to derive their reduced-order equations.

In fluid dynamics, vortex methods decompose the flow field into a combination of vortices, with a prescribed smooth profile, whose centers evolve over time [14, 8]. The motion of each vortex center is determined by computing the induced velocity by the other vortices.

Self-similar solutions of PDE are also obtained by varying the parameters of a prescribed ansatz [7]. Rowley et al. [32] use symmetry reduction from geometric mechanics to derive the symmetry reduced governing equations. Although rigorous, this method is only applicable for reducing continuous symmetries as opposed to RONS which is a more general reduced-order modeling framework.

Finally, we point out that the idea of optimally time-dependent modes (OTD) has been introduced in the context of stability analysis [6]. However, OTD is only applicable to linearized PDEs. In addition, to compute each OTD mode, an auxiliary PDE must be solved which renders this method computationally expensive [5, 22]. In contrast, RONS is applicable to fully nonlinear PDEs and reduces the computational cost as it only requires solving a relatively small set of ODEs.

1.2 Outline of the paper

This paper is organized as follows. In section 2, we introduce the problem setup in which our method is applicable. Section 3 describes RONS in detail and

contains our main results. In section 4, we present numerical results on three different examples. Section 5 contains our concluding remarks.

2 Set-up and preliminaries

We consider PDEs of the form

$$\frac{\partial u}{\partial t} = F(u), \quad u(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad (2)$$

for the map $u : D \times \mathbb{R}^+ \rightarrow \mathbb{R}^p$, $(\mathbf{x}, t) \mapsto u(\mathbf{x}, t)$ where D is a subset of \mathbb{R}^k . Here, F is a potentially nonlinear differential operator. For simplicity, we restrict our attention to the case $p = 1$, where u is a scalar function of the spatial variable $\mathbf{x} \in D$ and time $t \geq 0$. Our results generalize to the case $p > 1$ in a straightforward fashion. We assume that for any time t , $u(\cdot, t)$ belongs to a Hilbert function space H with the inner product $\langle \cdot, \cdot \rangle_H$ and the corresponding norm $\| \cdot \|_H$. The appropriate boundary conditions for the PDE are encoded in the Hilbert space H .

We seek approximate solutions of the form $\hat{u}(\mathbf{x}, \mathbf{q}(t))$ to PDE (2). The dependence of \hat{u} on the spatial variable \mathbf{x} is prescribed based on the type of the PDE and requires familiarity with its solutions. As such, we refer to the approximation solution $\hat{u}(\mathbf{x}, \mathbf{q}(t))$ as the *ansatz*. For instance, the heat kernel discussed in Section 1 has a Gaussian shape with time-dependent amplitude $A(t)$ and length scale $L(t)$ so that $\mathbf{q}(t) = (A(t), L(t))$. Given the prescribed spatial shape of the ansatz \hat{u} , our objective is to determine the evolution of the parameters $\mathbf{q}(t)$, so that the ansatz $\hat{u}(\mathbf{x}, \mathbf{q}(t))$ best approximates a true solution $u(\mathbf{x}, t)$ of the PDE. Since the ansatz solutions are typically smooth in the spatial variable \mathbf{x} , we only consider strong solutions of the PDE, assuming their existence and uniqueness.

We view the ansatz $\hat{u}(\mathbf{x}, \mathbf{q})$ as a map from the parameters $\mathbf{q} \in \Omega \subseteq \mathbb{R}^n$ to the function space H ,

$$\begin{aligned} \hat{u} : \Omega &\rightarrow H \\ \mathbf{q} &\mapsto \hat{u}(\cdot, \mathbf{q}), \end{aligned} \quad (3)$$

where Ω is a simply connected open subset of \mathbb{R}^n over which the ansatz \hat{u} is well-defined. For instance, for the heat kernel discussed in Section 1, Ω is the positive quadrant of \mathbb{R}^2 where $A > 0$ and $L > 0$. We note that the domain of the map \hat{u} is over the finite-dimensional parameters $\mathbf{q} = (q_1, q_2, \dots, q_n)$, whereas, for a given parameter \mathbf{q} , its image is a function in the Hilbert space H (see figure 1 for an illustration). We refer to the image of the map \hat{u} as the ansatz manifold, motivating the following definition.

Definition 1 (Ansatz Manifold). We refer to the set

$$\mathcal{M} := \{u \in H : \exists \mathbf{q} \in \Omega \text{ with } u = \hat{u}(\cdot, \mathbf{q})\} \subset H, \quad (4)$$

which is the image of the map (3), as the ansatz manifold.

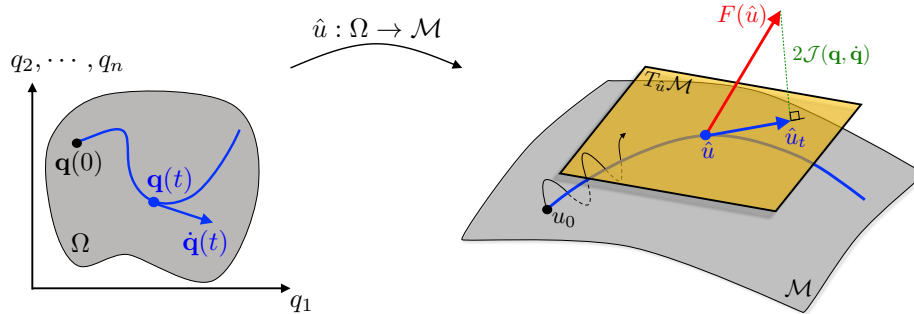


Figure 1: Geometric illustration of RONS. The reduced solution evolves on the ansatz manifold $\mathcal{M} \subset H$, which is the image of the ansatz map \hat{u} . The blue curve illustrates the evolution of the reduced-order solution. The black curve marks the true trajectory of the PDE which can leave \mathcal{M} since the manifold is not invariant under the full dynamics.

For well-posedness of the reduced-order equations governing the evolution of the parameters \mathbf{q} , a number of mild assumptions regarding the ansatz map \hat{u} are required. We list these assumptions below and explain their relevance.

Assumption 1. We assume that the map $\hat{u} : \Omega \rightarrow H$ has the following properties.

- (H1) It is injective and at least once continuously differentiable with respect to the parameters \mathbf{q} .
- (H2) For every parameter $\mathbf{q} \in \Omega$, the ansatz $\hat{u}(\mathbf{x}, \mathbf{q})$ is ℓ -times continuously differentiable with respect to the spatial variable \mathbf{x} . Here, $\ell \in \mathbb{N}$ is the highest order of spatial derivatives appearing in the PDE (2).
- (H3) The map $\hat{u} : \Omega \rightarrow H$ is an immersion.

Our reduced-order equations involve partial derivatives of the ansatz with respect to the parameters. Assumption (H1) ensures that these derivatives exist. Assumption (H2) ensures that the ansatz would in fact approximate strong solutions of the PDE. Finally, assumption (H3) requires the map \hat{u} to be an immersion, which essentially means that the tangent space of the ansatz manifold at every base point $\hat{u} \in \mathcal{M}$ is full rank. More precisely, the tangent space $T_{\hat{u}}\mathcal{M}$ at a point $\hat{u} \in \mathcal{M}$ is the linear vector space

$$T_{\hat{u}}\mathcal{M} := \text{span} \left\{ \frac{\partial \hat{u}}{\partial q_1}, \frac{\partial \hat{u}}{\partial q_2}, \dots, \frac{\partial \hat{u}}{\partial q_n} \right\}. \quad (5)$$

The map \hat{u} being an immersion implies that the partial derivatives $\partial \hat{u} / \partial q_i$ are linearly independent functions and therefore $\dim(T_{\hat{u}}\mathcal{M}) = n$ for all $\hat{u} \in \mathcal{M}$. As such, the set \mathcal{M} is an n -dimensional manifold immersed in the function space

H . Most importantly, assumption (H3) ensures that the metric tensor, to be defined in Section 3, is invertible. It should be clear that, as a special case, the ansatz map includes linear superposition of potentially parameter-dependent modes,

$$\hat{u}(\mathbf{x}, \mathbf{q}) = \sum_i a_i u_i(\mathbf{x}, \boldsymbol{\alpha}_i), \quad (6)$$

where \mathbf{q} consists of the parameters a_i and $\boldsymbol{\alpha}_i$.

As reviewed in the Introduction, earlier studies which use nonlinear reduced-order solutions are confronted with two main tasks,

1. Choosing an ansatz manifold, i.e., choosing the shape of the ansatz $\hat{u}(\mathbf{x}, \mathbf{q}(t))$,
2. Choosing a strategy for evolving the variables $\mathbf{q}(t)$ in time.

The first task relies on domain expertise and familiarity with the solutions of the PDE. The second task has thus far been carried out based on ad hoc methods. The present paper introduces a rigorous and unified approach for addressing the second task. Choosing an appropriate ansatz manifold remains a challenging problem and still relies on domain expertise.

3 Evolution of reduced-order solutions

This section contains our main results. First, in Section 3.1, we discuss the minimization principle from which the reduced-order equations are derived. For clarity, we do not assume any conserved quantities for the PDE in this section. Next, we show that our reduced-order equations are equivalent to the standard Galerkin projection if the ansatz is a linear function of the parameters (Section 3.2). Finally, for PDEs with conserved quantities, in Section 3.3 we derived reduced-order equations that respect those conservation laws.

3.1 Reduced-order equations without conserved quantities

Given an ansatz $\hat{u}(\mathbf{x}, \mathbf{q}(t))$, we would like to determine an evolution equation for the parameters $\mathbf{q}(t)$ such that the ansatz closely approximates a true solution $u(\mathbf{x}, t)$ of the PDE (2). Since the true solutions are a priori unknown, this cannot be achieved by minimizing some distance metric between the ansatz and the true solution. Therefore, an alternative metric needs to be used.

Here, we evolve the ansatz such that its instantaneous dynamics best match the true dynamics of the PDE. To this end, we consider the *instantaneous error*,

$$\mathcal{J}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \|\hat{u}_t - F(\hat{u})\|_H^2, \quad (7)$$

where \hat{u}_t is shorthand for the derivative of the ansatz $\hat{u}(\mathbf{x}, \mathbf{q}(t))$ with respect to time, i.e.,

$$\hat{u}_t(\mathbf{x}, \mathbf{q}(t)) = \frac{\partial \hat{u}}{\partial q_i}(\mathbf{x}, \mathbf{q}(t)) \dot{q}_i(t), \quad (8)$$

where summation over repeated indices is implied. We note that \hat{u}_t belongs to the tangent space $T_{\hat{u}}\mathcal{M}$ of the ansatz manifold \mathcal{M} at the point \hat{u} .

The cost function (7) measures the instantaneous discrepancy between the dynamics of the ansatz, i.e. \hat{u}_t , and the dynamics dictated by the PDE, i.e. $F(\hat{u})$. Given a set of parameters $\mathbf{q}(t)$, at any time $t \geq 0$, we seek $\dot{\mathbf{q}}(t)$ such that the instantaneous error $\mathcal{J}(\mathbf{q}, \dot{\mathbf{q}})$ is minimized. Figure 1 shows a geometric illustration of the instantaneous error \mathcal{J} .

Remark 1. As an alternative to the instantaneous error, one might be tempted to minimize the *finite-time error*,

$$S[\mathbf{q}] = \int_0^T \mathcal{J}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) dt, \quad (9)$$

over a time interval $t \in [0, T]$. At first sight, this functional seems more suitable, compared to its instantaneous counterpart (7), since it measures the accumulated error between the true dynamics and the ansatz dynamics. However, as we show in Appendix A, it generally leads to unstable reduced-order models for the evolution of the parameters. Therefore, in this paper we choose to minimize the instantaneous error.

As mentioned earlier, given the parameter $\mathbf{q}(t)$ at time t , we determine its dynamics by minimizing the instantaneous error (7) over all possible $\dot{\mathbf{q}}(t)$. More precisely, we solve the minimization problem

$$\min_{\dot{\mathbf{q}} \in \mathbb{R}^n} \mathcal{J}(\mathbf{q}, \dot{\mathbf{q}}). \quad (10)$$

We prove that this problem has a unique minimizer which satisfies a first-order ODE for the parameters \mathbf{q} . By solving this ODE, the optimal evolution of the parameters $\mathbf{q}(t)$ can be obtained. To prove these results, we need the following lemma.

Lemma 1 (Metric Tensor). Let Assumption 1 hold. Then the metric tensor M defined by

$$M_{ij} = \left\langle \frac{\partial \hat{u}}{\partial q_i}, \frac{\partial \hat{u}}{\partial q_j} \right\rangle_H, \quad i, j \in \{1, 2, \dots, n\}, \quad (11)$$

is a symmetric positive-definite matrix for all $\mathbf{q} \in \Omega$.

Proof. It is clear from its definition that M is symmetric. We show that the matrix is positive definite by proving that $\langle \boldsymbol{\xi}, M\boldsymbol{\xi} \rangle > 0$ for all nonzero $\boldsymbol{\xi} \in \mathbb{R}^n$, where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product. We first note that

$$\langle \boldsymbol{\xi}, M\boldsymbol{\xi} \rangle = \left\langle \frac{\partial \hat{u}}{\partial q_i} \xi_i, \frac{\partial \hat{u}}{\partial q_j} \xi_j \right\rangle_H = \left\| \frac{\partial \hat{u}}{\partial q_i} \xi_i \right\|_H^2 \geq 0, \quad (12)$$

where summation over repeated indices is implied. Next we show that $\langle \boldsymbol{\xi}, M\boldsymbol{\xi} \rangle$ is in fact strictly positive.

Assume that there exists $\boldsymbol{\xi} \neq \mathbf{0}$ such that $\langle \boldsymbol{\xi}, M\boldsymbol{\xi} \rangle = 0$. This, together with equation (12), implies that

$$\frac{\partial \hat{u}}{\partial q_i} \xi_i = 0.$$

Since $\boldsymbol{\xi} \neq \mathbf{0}$, this implies that $\partial \hat{u} / \partial q_i$, $i = 1, 2, \dots, n$, are linearly dependent. However, this violates assumption (H3) that the map \hat{u} is an immersion. Therefore, we must have $\langle \boldsymbol{\xi}, M\boldsymbol{\xi} \rangle > 0$ which completes the proof. \square

The fact that the metric tensor M is symmetric positive-definite, and hence invertible, plays an important role in deriving our reduced-order equations. Now we state the main result of this section which establishes that there exists a unique solution to the minimization problem (10), and that the minimizer satisfies an explicit ODE.

Theorem 1. Let Assumption 1 hold. Then there exists a unique solution to the minimization problem (10). Furthermore, the minimizer satisfies

$$\dot{\mathbf{q}} = M^{-1}(\mathbf{q})\mathbf{f}(\mathbf{q}) \quad (13)$$

where M is the metric tensor defined in Lemma 1 and $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a vector field defined by

$$f_i = \left\langle \frac{\partial \hat{u}}{\partial q_i}, F(\hat{u}) \right\rangle_H, \quad i = 1, 2, \dots, n.$$

Proof. First note that the instantaneous error (7) can be written more explicitly as

$$\begin{aligned} \mathcal{J}(\mathbf{q}, \dot{\mathbf{q}}) &= \frac{1}{2} [\langle \hat{u}_t, \hat{u}_t \rangle_H - 2\langle \hat{u}_t, F(\hat{u}) \rangle_H + \langle F(\hat{u}), F(\hat{u}) \rangle_H] \\ &= \frac{1}{2} \dot{q}_i \dot{q}_j \left\langle \frac{\partial \hat{u}}{\partial q_i}, \frac{\partial \hat{u}}{\partial q_j} \right\rangle_H - \dot{q}_i \left\langle \frac{\partial \hat{u}}{\partial q_i}, F(\hat{u}) \right\rangle_H + \frac{1}{2} \langle F(\hat{u}), F(\hat{u}) \rangle_H \\ &= \frac{1}{2} \langle \dot{\mathbf{q}}, M(\mathbf{q})\dot{\mathbf{q}} \rangle - \langle \dot{\mathbf{q}}, \mathbf{f}(\mathbf{q}) \rangle + \frac{1}{2} \langle F(\hat{u}), F(\hat{u}) \rangle_H. \end{aligned}$$

Therefore, \mathcal{J} is a smooth quadratic function of $\dot{\mathbf{q}}$. Furthermore, since the metric tensor $M(\mathbf{q})$ is symmetric positive-definite, \mathcal{J} is a strictly convex function of $\dot{\mathbf{q}}$. As such, it has a unique minimizer satisfying $\nabla_{\dot{\mathbf{q}}} \mathcal{J}(\mathbf{q}, \dot{\mathbf{q}}) = 0$, or equivalently $M(\mathbf{q})\dot{\mathbf{q}} - \mathbf{f}(\mathbf{q}) = 0$. Since the metric tensor $M(\mathbf{q})$ is invertible, we obtain equation (13) for $\dot{\mathbf{q}}$. \square

Figure 1 depicts the geometric interpretation of Theorem 1. Consider an arbitrary, but smooth, evolution of the parameters $\mathbf{q}(t)$. This can be viewed as a curve $\mathbf{q} : [0, \infty) \rightarrow \Omega$, $t \mapsto \mathbf{q}(t)$ parametrized by time t . The tangent vector to the curve at a point $\mathbf{q}(t)$ is given by $\dot{\mathbf{q}}(t)$. The ansatz \hat{u} maps this curve onto a curve on the ansatz manifold \mathcal{M} in the function space H . The tangent vector to this second curve is \hat{u}_t (see equation (8)) which is bound to belong to the tangent space $T_{\hat{u}}\mathcal{M}$ of the ansatz manifold, i.e., $\hat{u}_t \in T_{\hat{u}}\mathcal{M}$.

On the other hand, $F(\hat{u})$ does not necessarily belong to $T_{\hat{u}}\mathcal{M}$ since the ansatz manifold \mathcal{M} is not invariant under the dynamics of the PDE (2). Therefore, typically there is no evolution of the parameters $\mathbf{q}(t)$ such that the ansatz $\hat{u}(\cdot, \mathbf{q}(t))$ would solve the PDE. If the parameters $\mathbf{q}(t)$ evolve according to equation (13), Theorem 1 ensures that the deviation of the ansatz dynamics \hat{u}_t from the true dynamics $F(\hat{u})$ is instantaneously minimized. In fact, \hat{u}_t is the orthogonal projection of $F(\hat{u})$ onto the tangent space of the manifold at point \hat{u} .

We close this section with a remark about the initial condition of the parameters. In order to numerically integrate the reduced-order equation (13), we need to supply an initial condition $\mathbf{q}(0) = \mathbf{q}_0$. If the initial condition u_0 of the PDE belongs to the ansatz manifold \mathcal{M} , there exists a unique $\mathbf{q}_0 \in \Omega$ such that $u_0(\mathbf{x}) = \hat{u}(\mathbf{x}, \mathbf{q}_0)$. In this case, there is no ambiguity regarding the appropriate initial parameter values \mathbf{q}_0 . However, if $u_0 \notin \mathcal{M}$, a criterion must be devised to determine a suitable initial condition \mathbf{q}_0 . One such criterion, for instance, is to solve the optimization problem,

$$\mathbf{q}_0 = \arg \min_{\mathbf{q} \in \Omega} \|u_0 - \hat{u}(\cdot, \mathbf{q})\|_H, \quad (14)$$

which returns the closest point on the ansatz manifold to the initial condition u_0 . Optimization problem (14) can potentially be non-convex and computationally expensive to solve, but it is solved only once at the initial time.

3.2 Relation to Galerkin projection

In this section, we show that if the ansatz $\hat{u}(\mathbf{x}, \mathbf{q})$ is linear in the parameters \mathbf{q} , then the reduced-order equations (13) coincides with the standard Galerkin truncation.

Consider the ansatz

$$\hat{u}(\mathbf{x}, \mathbf{q}(t)) = \sum_{i=1}^n q_i(t) u_i(\mathbf{x}), \quad (15)$$

which is a linear superposition of the prescribed *modes* $\{u_i\}_{i=1}^n$. Without loss of generality, we assume that these modes are orthonormal with respect to the inner product on the Hilbert space H , i.e., $\langle u_i, u_j \rangle_H = \delta_{ij}$, $i, j \in \{1, 2, \dots, n\}$, where δ_{ij} denotes the Kronecker delta. These modes span an n -dimensional linear subspace V of the function space H , where $V := \text{span}\{u_i\}_{i=1}^n$. The subspace V is not generally an invariant subspace under the dynamics of the PDE (2). Therefore, the Galerkin ansatz (15) may not be an exact solution of the PDE. More precisely, although \hat{u}_t belongs to V , the right-hand side of the PDE $F(\hat{u})$ does not generally lie in V .

To obtain an approximate solution, one defines the projection operator $P : H \rightarrow V$ which is an orthogonal projection onto the subspace V , and replaces the right-hand side with $PF(\hat{u})$. Substituting the ansatz in the truncated PDE,

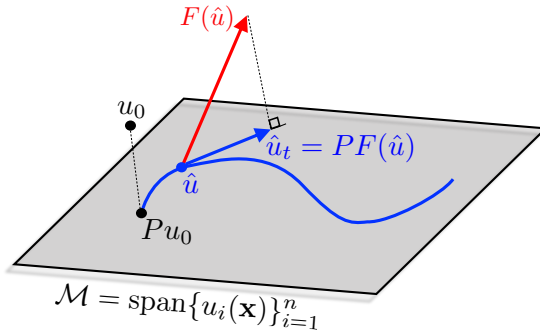


Figure 2: The geometry of the reduced-order equation for a linear ansatz. In this case, the ansatz manifold \mathcal{M} is a linear subspace of the function space H . The reduced-order equations coincide with the Galerkin truncation of the PDE to the modes u_i that span this subspace.

$\partial_t \hat{u} = PF(\hat{u})$, and taking the inner product with a mode u_k , we finally obtain the standard Galerkin projection,

$$\dot{q}_k = \langle u_k, F(\hat{u}) \rangle_H, \quad k = 1, 2, \dots, n. \quad (16)$$

Now, we show that the reduced-order equation (13) of Theorem 1 coincides with the Galerkin truncation (16). First note that, for the linear ansatz (15), we have

$$\frac{\partial \hat{u}}{\partial q_i} = u_i, \quad (17)$$

and therefore the metric tensor M is an identity matrix since $M_{ij} = \langle u_i, u_j \rangle = \delta_{ij}$. Furthermore, for a linear ansatz, the vector field $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is given by

$$f_i = \langle u_i, F(\hat{u}) \rangle_H. \quad (18)$$

Substituting these in the reduced-order equation (13), we see that it coincides exactly with the Galerkin truncation (16).

Figure 2 depicts the geometric meaning of the Galerkin projection. In the case of a linear ansatz, the ansatz manifold \mathcal{M} becomes a linear subspace of the function space H . The Galerkin method evolves the ansatz so that instantaneously the tangent vector to the path of the ansatz is the orthogonal projection of $F(\hat{u})$ onto the ansatz subspace. Therefore, our method is a generalization of the Galerkin method to the case when the ansatz manifold is a nonlinear subset of the function space. In this more general case, the evolution of the ansatz is obtained by projecting $F(\hat{u})$ onto the tangent space of the manifold \mathcal{M} at the ansatz \hat{u} (see figure 1).

3.3 Reduced-order equations with conserved quantities

Certain PDEs possess conserved quantities which are invariant along their trajectories [12]. Sometimes these conserved quantities are evident because the PDE is derived from underlying conservation laws [26]. Other conserved quantities, such as helicity in Euler’s equation for ideal fluids [28], are not evident and are only discovered by further mathematical analysis of the PDE. In either case, constructing reduced-order models that preserve these conserved quantities is attractive for two main reasons. First, conserved quantities reflect symmetries and physical properties of the system that reduced-order models should also exhibit [25]. Secondly, it is known that reduced-order models that violate the conservation laws can develop spurious finite-time blowups and other non-physical dynamics [27].

In this section, we modify the reduced-order equations developed in Section 2 to enforce conserved quantities of the PDE. Let $I_k : H \rightarrow \mathbb{R}$ with $k \in \{1, 2, \dots, m\}$ denote m conserved quantities of the PDE (2). More precisely, if $u(\cdot, t)$ denotes a solution of the PDE, then $I_k(u(\cdot, t)) = I_k(u_0)$ and therefore I_k is independent of time. We seek reduced-order equations for evolving an ansatz $\hat{u}(\mathbf{x}, \mathbf{q}(t))$ such that the conserved quantities I_k are also invariant along the trajectories of the ansatz. For notational simplicity, we write $I_k(\mathbf{q})$ instead of $I_k(\hat{u}(\cdot, \mathbf{q}))$ and view $I_k(\mathbf{q})$ as a map from the parameter space $\Omega \subseteq \mathbb{R}^n$ to the real line.

Remark 2. To be precise, $I_k(\hat{u}(\cdot, \mathbf{q}))$ is the pullback of $I_k : H \rightarrow \mathbb{R}$ under the ansatz map $\hat{u} : \Omega \rightarrow H$. In other words, we should write

$$\hat{u}^* I_k(\mathbf{q}) = I_k(\hat{u}(\cdot, \mathbf{q})), \quad \forall \mathbf{q} \in \Omega,$$

where $\hat{u}^* I_k$ denotes the pullback of I_k under the ansatz map \hat{u} . However, this unnecessarily complicates the notation. Therefore, we simply write $I_k(\mathbf{q})$ instead of $\hat{u}^* I_k(\mathbf{q})$.

To enforce the conserved quantities, we add them as constraints to the optimization problem (10), and solve

$$\begin{aligned} & \min_{\hat{\mathbf{q}} \in \mathbb{R}^n} \mathcal{J}(\mathbf{q}, \hat{\mathbf{q}}), \\ \text{s.t. } & I_k(\mathbf{q}(t)) = I_{k,0}, \quad k = 1, 2, \dots, m, \quad \forall t \geq 0, \end{aligned} \quad (19)$$

where $I_{k,0}$ are prescribed constants. In practice, these constants are determined by the initial condition so that $I_{k,0} = I_k(\mathbf{q}_0)$. To avoid degenerate solutions of the optimization problem, we must make the following assumption.

Assumption 2. We assume that, for each $\mathbf{q} \in \Omega$, the gradients $\nabla I_1(\mathbf{q}), \nabla I_2(\mathbf{q}), \dots$, and $\nabla I_m(\mathbf{q})$ are linearly independent.

In order to derive the reduced-order equations corresponding to the constrained optimization problem (19), we need the following lemma.

Lemma 2 (Constraint Matrix). Let Assumptions 1 and 2 hold and $M \in \mathbb{R}^{n \times n}$ denote the metric tensor defined in Lemma 1. Then the *constraint matrix* $C(\mathbf{q}) \in \mathbb{R}^{m \times m}$ with entries defined by

$$C_{ij} = \langle \nabla I_j, M^{-1} \nabla I_i \rangle, \quad i, j \in \{1, 2, \dots, m\}, \quad (20)$$

is symmetric positive-definite for all $\mathbf{q} \in \Omega$.

Proof. First, we define the $n \times m$ matrix $B := (\nabla I_1 | \nabla I_2 | \dots | \nabla I_m)$ and note that the constraint matrix can be written as $C = B^\top M^{-1} B$. Since the matrix tensor M is symmetric, so is the constraint matrix C .

Now we prove that C is also positive-definite. For any nonzero $\mathbf{v} \in \mathbb{R}^m$, we have

$$\langle \mathbf{v}, C\mathbf{v} \rangle = \langle \mathbf{v}, B^\top M^{-1} B\mathbf{v} \rangle = \langle B\mathbf{v}, M^{-1} B\mathbf{v} \rangle > 0, \quad (21)$$

where the last inequality is justified by the facts that M^{-1} is symmetric positive-definite and $B\mathbf{v}$ is nonzero for $\mathbf{v} \neq 0$. The last statement is the consequence of the fact that B is full-rank due to Assumption 2. Therefore, the constraint matrix C is symmetric positive-definite. \square

The following theorem gives the reduced-order equations for $\mathbf{q}(t)$ such that the quantities I_k are conserved along the trajectory of the ansatz.

Theorem 2. Let assumptions 1 and 2 hold. If a solution to the the constrained optimization problem (19) exists, it must satisfy

$$\dot{\mathbf{q}} = M^{-1}(\mathbf{q}) \left[\mathbf{f}(\mathbf{q}) - \sum_{k=1}^m \lambda_k \nabla I_k(\mathbf{q}) \right], \quad (22)$$

where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_m)^\top$ is the unique solution to the linear equation

$$C\boldsymbol{\lambda} = \mathbf{b}, \quad (23)$$

with the constrained matrix C defined in Lemma 2 and the components of the vector $\mathbf{b} = (b_1, b_2, \dots, b_m)^\top \in \mathbb{R}^m$ defined as

$$b_i = \langle \nabla I_i, M^{-1} \mathbf{f} \rangle.$$

The metric tensor M and the vector field \mathbf{f} are defined as in Theorem 1.

Proof. First, we rewrite the constraints in (19) in an equivalent form. Taking a time derivative, the constraints can be alternatively written as

$$\frac{d}{dt} I_k(\mathbf{q}(t)) = \langle \nabla I_k(\mathbf{q}), \dot{\mathbf{q}} \rangle = 0, \quad k = 1, 2, \dots, m. \quad (24)$$

Introducing the Lagrange multiplier $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m)^\top \in \mathbb{R}^m$, we define the augmented cost function,

$$\mathcal{J}_c(\mathbf{q}, \dot{\mathbf{q}}, \boldsymbol{\lambda}) := \mathcal{J}(\mathbf{q}, \dot{\mathbf{q}}) + \sum_{k=1}^m \lambda_k \langle \nabla I_k(\mathbf{q}), \dot{\mathbf{q}} \rangle. \quad (25)$$

If the minimizer of the constrained optimization problem (19) exists, the partial derivatives of \mathcal{J}_c with respect to q_i and λ_k must vanish at the minimizer. This yields

$$\nabla_{\dot{\mathbf{q}}}\mathcal{J} + \sum_{k=1}^m \lambda_k \nabla I_k = 0, \quad (26a)$$

$$\langle \nabla I_1(\mathbf{q}), \dot{\mathbf{q}} \rangle = \langle \nabla I_2(\mathbf{q}), \dot{\mathbf{q}} \rangle = \dots = \langle \nabla I_m(\mathbf{q}), \dot{\mathbf{q}} \rangle = 0. \quad (26b)$$

Since $\nabla_{\dot{\mathbf{q}}}\mathcal{J} = M(\mathbf{q})\dot{\mathbf{q}} - \mathbf{f}(\mathbf{q})$, equation (26a) yields

$$\dot{\mathbf{q}} = M^{-1}(\mathbf{q}) \left[\mathbf{f}(\mathbf{q}) - \sum_{k=1}^m \lambda_k \nabla I_k(\mathbf{q}) \right]. \quad (27)$$

Substituting this expression into (26b), yields m equations

$$\sum_{k=1}^m \lambda_k \langle \nabla I_i, M^{-1} \nabla I_k \rangle = \langle \nabla I_i, M^{-1} \mathbf{f} \rangle, \quad i = 1, 2, \dots, m. \quad (28)$$

Equation (28) can be written as the linear system $C\boldsymbol{\lambda} = \mathbf{b}$, where C is the constraint matrix (20). Lemma 2 guarantees that C is invertible and therefore there exists a unique solution $\boldsymbol{\lambda}$ to the linear system (23). Thus, $\dot{\mathbf{q}}$ must satisfy equation (27) with the Lagrange multipliers λ_k solving the linear system (23). This completes the proof. \square

We note that unlike the unconstrained problem (Theorem 1), the constrained optimization problem (19) is not guaranteed to have a solution. For a minimizer to exist, the level sets $\{\mathbf{q} \in \Omega : I_k(\mathbf{q}) = I_{k,0}\}$, with $k = 1, 2, \dots, m$, must have a nonempty intersection. Otherwise, the problem is over-constrained and a minimizer would not exist. Apart from these degenerate situations, the constrained optimization problem has a solution and the ansatz evolves according to the reduced-order equations (22). In our numerical experiments, presented in Section 4, we never encountered a degenerate case where the solution does not exist.

4 Numerical examples

In this section, we present three numerical examples. We begin with a proof-of-concept example where the ansatz captures the exact solution to a linear advection-diffusion equation. The other two examples deal with nonlinear PDEs, where the ansatz is not an exact solution, but the reduced-order equations nonetheless capture the important features of the system dynamics.

4.1 Advection-diffusion equation

As a proof-of-concept example we first consider a linear advection-diffusion equation and an ansatz solution which is an exact solution of the PDE for an appropriate choice of the time-dependent ansatz parameters. We show that the reduced-order equations (Theorem 1) reproduce this exact solution.

Consider the linear advection-diffusion equation,

$$\frac{\partial u}{\partial t} = -c \frac{\partial u}{\partial x} + \nu \frac{\partial^2 u}{\partial x^2}, \quad u(x, 0) = A_0 \sin\left(\frac{x}{L_0}\right), \quad (29)$$

where $c \in \mathbb{R}$ and A_0, L_0 and ν are positive constants. Equation (29) admits the exact solution

$$u(x, t) = A_0 \exp\left[-\frac{\nu}{L_0^2}t\right] \sin\left(\frac{x - ct}{L_0}\right), \quad (30)$$

which is a traveling sine wave with a decaying amplitude. We define the ansatz

$$\hat{u}(x, \mathbf{q}(t)) = A(t) \sin\left(\frac{x}{L(t)} + \phi(t)\right), \quad (31)$$

with the time-dependent parameters $\mathbf{q}(t) = (A(t), L(t), \phi(t))^\top$. Although the ansatz is linear in the amplitude A , it is a nonlinear function of the parameters L and ϕ .

We choose the initial parameter values $A(0) = A_0$, $L(0) = L_0$, and $\phi(0) = 0$, so that the ansatz coincides with the initial condition of the PDE, i.e., $u(x, 0) = \hat{u}(x, \mathbf{q}(0))$. Then the ansatz is an exact solution of (29) if the parameter values evolve according to

$$A(t) = A_0 \exp\left[-\frac{\nu}{L_0^2}t\right], \quad L(t) = L_0, \quad \phi(t) = -\frac{ct}{L_0}. \quad (32)$$

Now we show that the reduced-order equations (13) exactly reproduce the parameter evolution (32). The cost function (7) for the linear advection-diffusion equation reads

$$\mathcal{J}(\hat{u}) = \frac{1}{2} \int_0^{2\pi L_0} |\hat{u}_t + c\hat{u}_x - \nu\hat{u}_{xx}|^2 dx, \quad (33)$$

where the integral is taken over one period of the initial condition. The appropriate function space for this problem is the Hilbert space of periodic square integrable functions $L_{per}^2(0, 2\pi L_0)$. In this section, we do not enforce any conserved quantities. After a straightforward calculation, the reduced-order equations (13) read

$$\dot{A} = -\frac{\nu}{L^2}A, \quad \dot{L} = 0, \quad \dot{\phi} = -\frac{c}{L}, \quad (34)$$

which have the exact solution (32). In other words, our reduced-order equations applied to the ansatz (31) reproduce the exact solution (30).

4.2 Nonlinear Schrödinger equation

In this section, we derive reduced-order equations approximating the solutions to the nonlinear Schrödinger equation (NLSE). NLSE is a perturbative model for optical waves [11, 3] and surface water wave [39, 21]. NLSE has been intensely

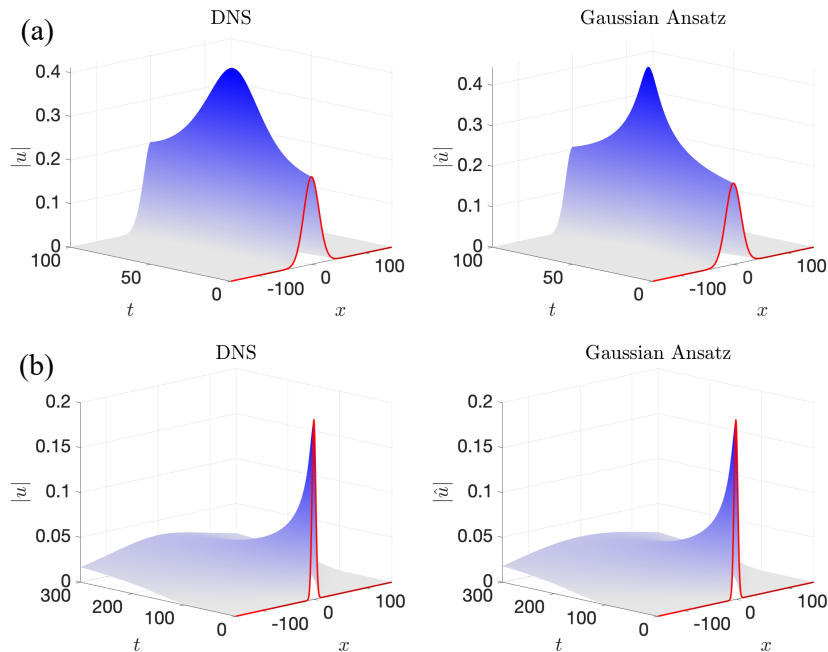


Figure 3: Comparing NLSE solutions using direct numerical simulations (DNS) and the Gaussian ansatz. (a) Focusing wave with initial parameters $A_0 = 0.2$, $L_0 = 20$, $V_0 = -0.05$, $\phi_0 = 0$. (b) Defocusing wave with initial parameters $A_0 = 0.2$, $L_0 = 5$, $V_0 = \phi_0 = 0$.

studied since, through modulational instability [9], it can reproduce the self-focusing of optical and water waves which leads to the formation of waves of extreme amplitude, often referred to as rogue waves [37, 13].

The nonlinear Schrödinger equation for unidirectional deep water waves is given by

$$\frac{\partial \tilde{u}}{\partial t} = -i \frac{\omega_0}{8k_0^2} \frac{\partial^2 \tilde{u}}{\partial \tilde{x}^2} - i \frac{\omega_0 k_0^2}{2} |\tilde{u}|^2 \tilde{u}, \quad (35)$$

where $\tilde{u}(\tilde{x}, \tilde{t})$ is the complex wave envelope. The prescribed constants ω_0 and k_0 denote the frequency and wave number of the carrier wave, respectively. The wave surface elevation is then given by $\eta(\tilde{x}, \tilde{t}) = \text{Re} [\tilde{u}(\tilde{x}, \tilde{t}) \exp(k_0 \tilde{x} - \omega_0 \tilde{t})]$. Introducing the non-dimensional variables $x = 2\sqrt{2}k_0 \tilde{x}$, $t = -\omega_0 \tilde{t}$, and $u = (k_0/\sqrt{2})\tilde{u}$, equation (35) becomes

$$\frac{\partial u}{\partial t} = i \frac{\partial^2 u}{\partial x^2} + i|u|^2 u. \quad (36)$$

In the following, we work with this non-dimensionalized NLSE.

It is known that wave groups with certain combinations of initial amplitude and length scale evolve under NLSE to grow in amplitude while their width

shrinks. This phenomena is referred to as focusing of wave groups. In contrast, defocusing wave groups exhibit decaying amplitude and widening width (see figure 3).

This observation has spurred a plethora of reduced-order methods for modeling and prediction of rogue waves [2, 1, 33, 34, 17, 22, 23]. These methods invariably assume a prescribed spatial shape for a wave group with time-dependent parameters, such as length scale, amplitude, velocity, and phase. The evolution of these parameters is then determined by ad hoc methods. Adcock et al. [2, 1], for instance, use conserved quantities of NLSE to determine the temporal evolution. Ruban [33, 34] leverages the variational formulation of NLSE to derive a set of ODEs for the parameters. Cousins and Sapsis [16] take an additional time derivative of NLSE and project the resulting equation onto their ansatz.

Here, we use the rigorous reduced-order equations (22) to evolve the time-dependent parameters and show that they correctly capture focusing and defocusing of the wave groups. Following earlier work [30, 33, 34], we use the Gaussian ansatz,

$$\hat{u}(x, \mathbf{q}(t)) = A(t) \exp \left[-\frac{x^2}{L^2(t)} + i\frac{x^2 V(t)}{L(t)} + i\phi(t) \right], \quad (37)$$

with the time-dependent parameters $\mathbf{q}(t) = (A, L, V, \phi)$, where A determines the wave amplitude, L is a length scale controlling how quickly the wave group decays away from its center, V is the wave velocity, and ϕ is the wave phase. The motivation for choosing this ansatz is that, for a special choice of the parameters $\mathbf{q}(t)$, it is an exact solution to the linear part of NLSE [34]. However, ansatz (37) is not an exact solution of NLSE except for the trivial case $A(t) = 0$. We note that \hat{u} is complex-valued, but its parameters $\mathbf{q} = (A, L, V, \phi)$ are real-valued with the additional restrictions that $A(t) > 0$ and $L(t) > 0$.

The cost function (7) for NLSE reads

$$\mathcal{J}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \int_{\mathbb{R}} |\hat{u}_t - i\hat{u}_{xx} - i|\hat{u}|^2 \hat{u}|^2 dx, \quad (38)$$

where $|\cdot|$ denotes the modulus of a complex number. The Hilbert space H here is the square integrable complex functions over \mathbb{R} . We note that, for simplicity, we stated our results in Section 3 for real-valued functions. Their generalization to complex-valued functions is straightforward.

NLSE has several conserved quantities [2, 1] which can be enforced using the method described in Section 3.3 (Theorem 2). Here, we only enforce the conservation of the two most relevant quantities,

$$I_1(u) = \int_{\mathbb{R}} |u|^2 dx, \quad I_2(u) = \int_{\mathbb{R}} |u_x|^2 dx - \frac{1}{4} \int_{\mathbb{R}} |u|^4 dx, \quad (39)$$

which are the total mass and the total energy, respectively. To obtain the reduced-order equations (22), we solve the constrained optimization problem (19) with $m = 2$ and

$$I_{1,0} = \sqrt{\frac{\pi}{2}} A_0^2 L_0, \quad I_{2,0} = \frac{\sqrt{\pi} A_0^2 (2\sqrt{2} (L_0^2 V_0^2 + 1) - A_0^2 L_0^2)}{8L_0},$$

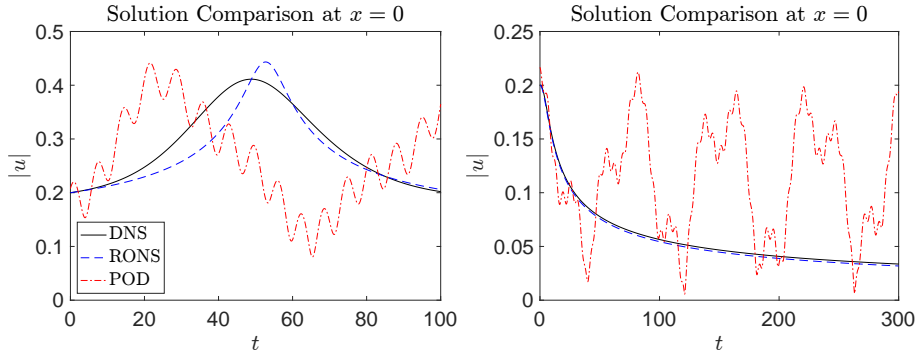


Figure 4: Evolution of NLSE solutions evaluated at $x = 0$ using direct numerical simulations (DNS), RONS with the Gaussian ansatz, and Galerkin projection onto four POD modes. Left: Focusing wave with initial parameters $A_0 = 0.2$, $L_0 = 20$, $V_0 = -0.05$, $\phi_0 = 0$. Right: Defocusing wave with initial parameters $A_0 = 0.2$, $L_0 = 5$, $V_0 = \phi_0 = 0$.

where (A_0, L_0, V_0, ϕ_0) are the initial values of the parameters of the ansatz (37). The resulting reduced-order ODEs read

$$\dot{A} = -\frac{2AV}{L}, \quad \dot{L} = 4V, \quad \dot{V} = \frac{4}{L^3} - \frac{A^2}{\sqrt{2L}}, \quad \dot{\phi} = \frac{5A^2}{4\sqrt{2}} - \frac{2}{L^2}. \quad (40)$$

For comparison, we use direct numerical simulations (DNS) of the NLSE as the ground truth. As in Ref. [17], we solve NLSE using a Fourier pseudospectral method in space and a fourth-order Runge–Kutta exponential time differencing scheme [18]. Boundary conditions are assumed to be periodic for ease of implementation and should not affect the results so long as the domain is chosen to be large enough so that u is small near the boundaries. We use the domain size $256\sqrt{2}\pi$ (64 wave periods) which we have found large enough to ensure u is small near the boundary for the duration of the computation. For all runs we use 2^{10} Fourier modes and a time step of 0.025. Using more Fourier modes, smaller time steps, or larger domain size did not significantly affect the results of our simulations. For example, doubling the number of Fourier modes and halving the time step caused a relative change of $|u|$ on the order of 10^{-7} in the L^2 -norm. The DNS results are initialized with $u_0(x) = \hat{u}(x, \mathbf{q}(0))$, so that the DNS solution and the ansatz solution coincide at the initial time.

In figure 3, we plot the evolution of the DNS results $|u|$ and the ansatz $|\hat{u}|$ for two sets of initial parameters. One set of parameters leads to a focusing wave (growing amplitude) and the other leads to a defocusing wave (decaying amplitude). In both cases, the ansatz correctly predicts the qualitative focusing or defocusing behavior of the wave group. We have repeated these comparisons for a range of parameter values \mathbf{q}_0 and in every case, the ansatz correctly predicts the focusing or defocusing behavior of the wave group.

In figure 4, we plot the evolution of the amplitudes at $x = 0$ for the same set of parameter values as in figure 3. For the case of the defocusing wave, the reduced-order solution produced by RONS is in excellent agreement with the DNS results. For the focusing wave the RONS reduced-order solution provides a reasonable approximation compared to the DNS, roughly capturing both the peak amplitude and the time that it occurs. However, the Gaussian ansatz overestimates the peak height and also produces a peak which is thinner than that of the DNS. This behavior, i.e. excellent agreement for defocusing waves and overshooting for focusing waves, is systematically observed for other parameter values \mathbf{q}_0 (not shown here).

To compare RONS results with a standard technique in model reduction, figure 4 also shows the results obtained by Galerkin projection with modes found through proper orthogonal decomposition (POD) as described in [36]. We use four POD modes as a benchmark which is on par with the four parameters (A, L, V, ϕ) involved in the Gaussian RONS ansatz (37). We point out that the POD modes are complex valued, so that the reduced-order equations involve four complex-valued ODEs or equivalently eight real-valued ODEs. In other words, the POD-reduced equations involve twice as many ODEs as the RONS reduced-order equations (40). Nonetheless, as shown in figure 4, the POD reduced-order solutions provide poor approximations for both focusing and defocusing waves, demonstrating the superiority of RONS. To obtain an accurate POD-reduced model which is capable of accurately approximating the focusing and defocusing waves, we have observed that at least 16 complex-valued POD modes are required.

4.3 Two-dimensional fluid flow

For our last example, we consider the flow of a two-dimensional, incompressible, and Newtonian fluid. The vorticity equation for such a fluid is given by

$$\frac{\partial \omega}{\partial t} + \mathbf{u} \cdot \nabla \omega = \nu \Delta \omega, \quad (41)$$

where ν is the kinematic viscosity, $\mathbf{u}(\mathbf{x}, t)$ denotes the velocity field of the fluid, and $\omega(\mathbf{x}, t)$ is the component of the vorticity orthogonal to the plane of motion. We consider the flow on the unbounded two-dimensional domain, $\mathbf{x} = (x, y) \in \mathbb{R}^2$. A two-dimensional incompressible fluid admits a stream function $\psi(x, y, t)$ which satisfies

$$\mathbf{u}(\mathbf{x}, t) = (\psi_y, -\psi_x)^\top, \quad \omega = -\Delta \psi. \quad (42)$$

Therefore, prescribing the stream function ψ is sufficient to determine the fluid velocity \mathbf{u} and the vorticity ω . As a result, we define the ansatz in terms of the stream function,

$$\hat{\psi}(\mathbf{x}, \mathbf{q}(t)) = \sum_{i=1}^N A_i(t) \exp \left[-\frac{|\mathbf{x} - \mathbf{x}_i(t)|^2}{L_i^2(t)} \right], \quad (43)$$

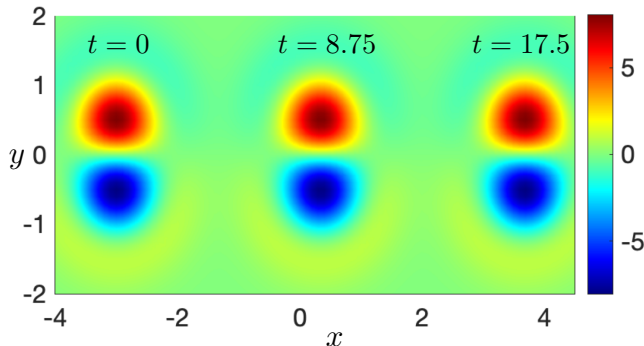


Figure 5: Reduced dynamics of a vortex dipole. The initial parameter so the ansatz are $A_1(0) = 1$, $A_2(0) = -1$, $L_1(0) = L_2(0) = 0.75$, $\mathbf{x}_1(0) = (-3, 0.5)$, $\mathbf{x}_2(0) = (-3, -0.5)$. See the supplementary material for a movie.

which is the superposition of N axisymmetric vortices, where $A_i(t)$ denotes the amplitude of the i -th vortex, $L_i(t)$ is its length scale, and $\mathbf{x}_i(t) = (x_i(t), y_i(t))$ is the vortex center. Therefore, each term of the ansatz has four time-dependent parameters, (A_i, L_i, x_i, y_i) , resulting in a total of $n = 4N$ parameters for N vortices. We use this ansatz because it belongs to the class of smooth vortex methods [14, 15] and, more importantly, vortices with the Gaussian stream function (43) have been observed in laboratory experiments [38].

After defining the stream function, the corresponding fluid velocity $\hat{\mathbf{u}}$ and vorticity $\hat{\omega}$ are computed to define the cost functional,

$$\mathcal{J}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \int_{\mathbb{R}^2} |\hat{\omega}_t + \hat{\mathbf{u}} \cdot \nabla \hat{\omega} - \nu \Delta \hat{\omega}|^2 dx. \quad (44)$$

For brevity, here we only study the inviscid case, $\nu = 0$, where the vortex dynamics has a rich mathematical structure. In particular, the inviscid vorticity equation admits a class of exact weak solutions called *point vortices* [4]. We show that the reduced equations governing the parameters of the ansatz (43) reproduce the expected point vortex dynamics, although the ansatz is not an exact solution of the vorticity equation.

Point vortex solutions of the inviscid vorticity equation consist of the superposition of Dirac delta functions,

$$\omega(\mathbf{x}, t) = \sum_{i=1}^N \Gamma_i \delta(\mathbf{x} - \mathbf{x}_i(t)), \quad (45)$$

where Γ_i are constant vortex amplitudes and $\mathbf{x}_i(t) = (x_i(t), y_i(t))$ are the time-dependent vortex centers. Using the Green's function for the Poisson equation $\Delta \psi = -\omega$, the stream function for point vortex solution (45) can be obtained as

$$\psi(\mathbf{x}, t) = -\frac{1}{2\pi} \sum_{i=1}^N \Gamma_i \log |\mathbf{x} - \mathbf{x}_i(t)|. \quad (46)$$

For the point vortex to be a weak solution of the vorticity equation, the vortex centers $\mathbf{x}_i(t)$ must satisfy the set of ordinary differential equations [29],

$$\Gamma_i \dot{x}_i = \frac{\partial \mathcal{H}}{\partial y_i}, \quad \Gamma_i \dot{y}_i = -\frac{\partial \mathcal{H}}{\partial x_i}, \quad (47)$$

where $\mathcal{H} = -\sum_{i \neq j} \Gamma_i \Gamma_j \log |\mathbf{x}_i - \mathbf{x}_j| / 4\pi$ is the Hamiltonian corresponding to the N vortex problem.

Dynamics of point vortices depends on the number of vortices N and their initial configuration. We refer the reader to Refs [4, 29] for a complete accounting. Here, we consider three distinct configurations,

1. Vortex dipole: Two vortices of equal strengths but with opposite signs,
2. Vortex pair: Two vortices of equal strengths and signs,
3. Leapfrogging: Two vortex pairs with opposite signs.

For each configuration, we show that the reduced-order equations (22) applied to the ansatz (43) reproduce the expected vortex dynamics predicted by the point vortex solution. We emphasize that, unlike the point vortex (46), our ansatz has a smooth profile and is not an exact solution of the vorticity equation. In addition, we allow the amplitude A_i and the length scale L_i of each vortex in the ansatz to vary with time.

The inviscid vorticity equation admits several conserved quantities [29]. Here, we enforce two of these conserved quantities,

$$I_1(\mathbf{u}) = \frac{1}{2} \int_{\mathbb{R}^2} |\mathbf{u}|^2 \, d\mathbf{x}, \quad I_2(\mathbf{u}) = \frac{1}{2} \int_{\mathbb{R}^2} |\nabla \times \mathbf{u}|^2 \, d\mathbf{x} = \frac{1}{2} \int_{\mathbb{R}^2} |\omega|^2 \, d\mathbf{x}, \quad (48)$$

which denote the kinetic energy and enstrophy, respectively. To derive the reduced-order equations, we minimize the cost function (44) with the constraints $I_1 = I_{1,0}$ and $I_2 = I_{2,0}$ (see Theorem 2).

First, we present the results for a vortex dipole. Point vortex dynamics predict that a vortex dipole, consisting of two vortices of opposite sign, travels on a straight line without changing shape. Figure 5 shows a similar scenario for our smooth Gaussian ansatz, where two vortices of opposite sign and equal amplitude are placed symmetrically along the x -axis. The vortices are evolved according to our reduced-order equation (22). The length-scales L_i and amplitudes A_i remain constant during the evolution. The vortex centers \mathbf{x}_i , however, translate horizontally at a constant speed, reproducing the expected dynamics that point vortices exhibit.

Next, we consider the case of a vortex pair, where the vortices have the same sign and amplitudes. Point vortex dynamics, in this case, predicts that the vortices rotate around their midpoint at a constant angular velocity. Figure 6 shows a similar vortex pair created by the Gaussian ansatz (43). The reduced-order dynamics shows that the system is in a relative equilibrium where the two vortices continually rotate counterclockwise around the origin $\mathbf{x} = 0$ at a

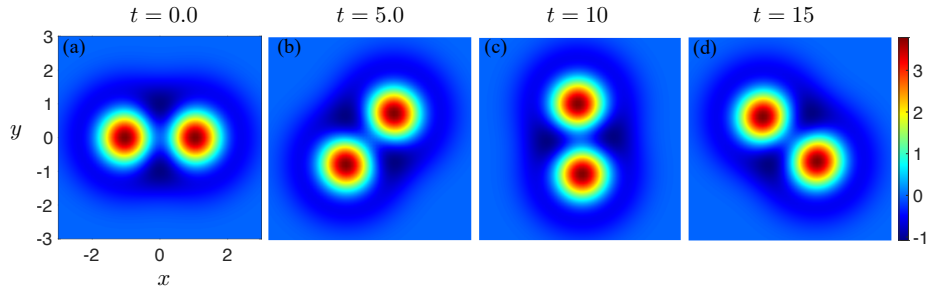


Figure 6: Reduced dynamics of a vortex pair. The initial parameter so the ansatz are $A_1(0) = A_2(0) = 1$, $L_1(0) = L_2(0) = 1$, $\mathbf{x}_1(0) = (-1, 0)$, $\mathbf{x}_2(0) = (1, 0)$. See the supplementary material for a movie.

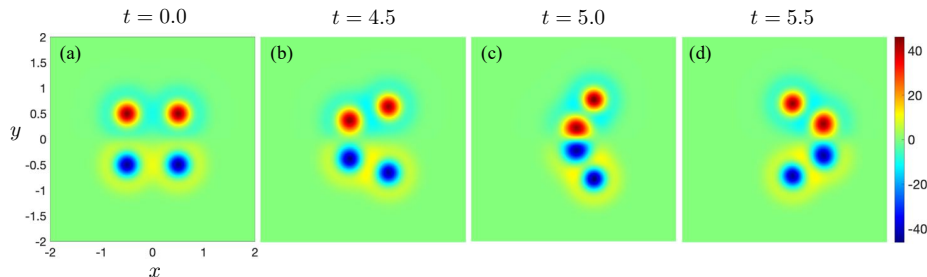


Figure 7: Reduced dynamics of leapfrogging vortices. The initial parameters values are $A_1(0) = A_3(0) = 1$, $A_2(0) = A_4(0) = -1$, $L_i(0) = 0.3$, $\mathbf{x}_1(0) = (0.5, 0.5)$, $\mathbf{x}_2(0) = (0.5, -0.5)$, $\mathbf{x}_3(0) = (-0.5, 0.5)$, $\mathbf{x}_4(0) = (-0.5, -0.5)$. See the supplementary material for a movie.

constant angular velocity. Again, the reduced-order dynamics are in excellent agreement with the expected point vortex behavior.

Finally, we consider the leapfrogging configuration, where four vortices are placed at the corners of a square (see figure 7). The vortices on the top row have positive vorticity whereas the vortices on the bottom row have negative vorticity. The point vortex dynamics predicts that the vortices to the left accelerate and zoom in between the vortices in the front. This motion repeats indefinitely, creating the so-called leapfrogging dynamics. Figure 7 shows the reduced dynamics of the Gaussian ansatz (43). One pair of vortices is placed at $(\pm 0.5, 0.5)$, and another pair of vortices with opposite sign is placed at $(\pm 0.5, -0.5)$. The RONS reduced-order equations once again reproduce the leapfrogging dynamics predicted by point vortices. The computational time to numerically integrate the RONS equations for the vortex dipole, vortex pair, and leapfrogging configurations was 0.0087, 0.0304, and 4.2188 seconds, respectively, on a 2019 Macbook Pro with a 1.7 GHz Quad-Core Intel Core i7 processor.

We chose the Gaussian stream function (43) as the ansatz since it has been observed in experiments [38]; then we showed that the RONS reduced-order

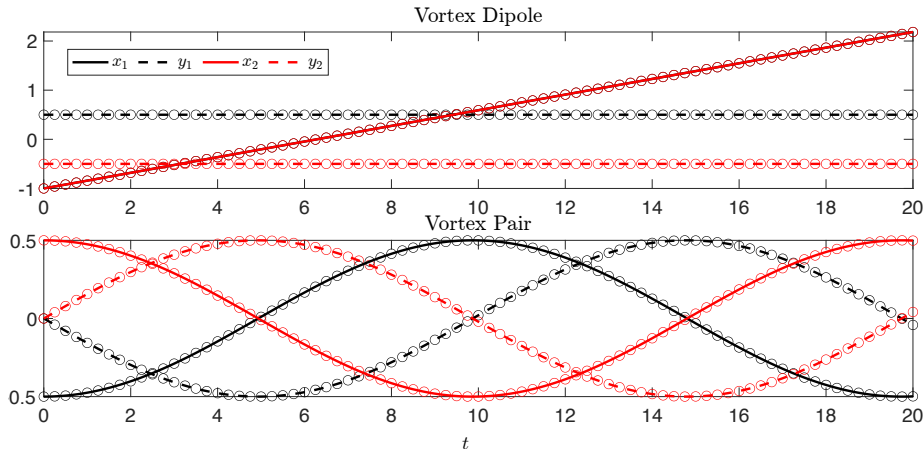


Figure 8: Comparing RONS against point vortex dynamics for a vortex dipole (top) and a vortex pair (bottom). RONS trajectories are solid and dashed lines and point vortex trajectories are marked by circles. The parameter values used for the vortex dipole are $\Gamma_1 = 1$, $\Gamma_2 = -1$, $\varepsilon(0) = 0.05$, $\mathbf{x}_1(0) = (-1, 0.5)$, $\mathbf{x}_2(0) = (-1, -0.5)$, and initial parameter values for the vortex pair are $\Gamma_1 = \Gamma_2 = 1$, $\varepsilon(0) = 0.05$, $\mathbf{x}_1(0) = (-0.5, 0)$, $\mathbf{x}_2(0) = (0.5, 0)$. Note that, in the vortex dipole plot, the time series of x_1 and x_2 overlap.

equations reproduce the expected dynamics. However, this ansatz is not convenient for a quantitative comparison with solutions produced by point vortices. To make quantitative comparisons with point vortex dynamics, we introduce the following ansatz for the vorticity,

$$\hat{\omega}(\mathbf{x}, \mathbf{q}) = \sum_{i=1}^N \frac{\Gamma_i}{2\pi\varepsilon^2} \exp\left[-\frac{|\mathbf{x} - \mathbf{x}_i|^2}{2\varepsilon^2}\right], \quad (49)$$

which is a sum of Gaussians with the amplitudes $A_i = \Gamma_i/(2\pi\varepsilon^2)$ and length-scales $L_i = \varepsilon\sqrt{2}$. Note that the ansatz (49) converges to point vortices (45) as ε tends to zero. After obtaining the corresponding fluid velocity $\hat{\mathbf{u}}$ using the Biot-Savart law [29], we derive the RONS equations by minimize the cost functional (44) while enforcing conservation of kinetic energy and enstrophy. Figure 8 shows the vortex center trajectories with $\varepsilon(0) = 0.05$ for the vortex dipole and the vortex pair. It shows that the RONS results are in excellent agreement with point vortex dynamics. We point out that, to derive the RONS reduced-order equations, we allow $\varepsilon(t)$ and $\mathbf{x}_i(t)$ to be time-dependent functions. Nonetheless, RONS predicts that ε remains constant ($\dot{\varepsilon} = 0$).

In closing, we point out that our reduced-order equations can be similarly applied to the viscous vorticity equation ($\nu > 0$) with no difficulty. In this case, the equation is dissipative and therefore there are no conserved quantities to be enforced. However, direct numerical simulations indicate that initially

axisymmetric vortices evolve under the viscous vorticity equation into an approximately elliptical shape. As a result, the axisymmetric ansatz (43) may no longer be appropriate, and a more general non-axisymmetric Gaussian ansatz might be needed. This viscous case will be explored in future work.

5 Conclusions

We proposed RONS as a general framework for evolving time-dependent nonlinear reduced-order solutions of PDEs which is applicable to a broad class of problems. The reduced-order solution has a prescribed shape in space and depends nonlinearly on a set of time-dependent parameters. Our reduced-order equations evolve the parameters such that the instantaneous error between the ansatz dynamics and the full PDE dynamics is minimized. Any number of conserved quantities can readily be enforced in our framework without requiring a Lagrangian or Hamiltonian formulation for the PDE.

We presented several numerical examples to demonstrate the effectiveness of this method: a linear advection-diffusion equation, the nonlinear Schrödinger equation, and Euler’s equation for two-dimensional ideal fluids. In every case, the reduced-order equations produce approximate solutions which capture all qualitative features of the PDE.

Our method opens a new paradigm in reduced-order modeling, with many aspects undoubtedly remaining to be explored. For instance, an upper bound should be derived to estimate the accumulated finite-time error between the reduced-order solution and the true solutions of the PDE. Here we focused on approximating strong solutions; further work is needed to derive reduced-order equations for PDEs in the weak formulation. Finally, here we resolved an important aspect of reduced-order nonlinear solutions, namely, the evolution of the ansatz variables $\mathbf{q}(t)$. However, the choice of the shape of the ansatz $\hat{u}(\mathbf{x}, \mathbf{q}(t))$ still relies on domain expertise and familiarity with the PDE. A systematic method for determining an appropriate ansatz manifold remains an open problem.

A Instability of the finite-time formulation

In Remark 1, we advised against using the finite-time error,

$$S[\mathbf{q}] = \int_0^T \mathcal{J}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) dt, \quad (50)$$

since it generally leads to unstable reduced-order models. In this section, we demonstrate this with a simple example. Consider the linear PDE,

$$\frac{\partial u}{\partial t} = \mathcal{L}u, \quad (51)$$

where \mathcal{L} is a self-adjoint negative-definite operator with orthonormal eigenfunctions $u_i(\mathbf{x})$ and corresponding eigenvalues $-\lambda_i$, where $\lambda_i > 0$.

Approximating a solution of the PDE as the linear combination of the eigen functions u_i , we consider the linear ansatz,

$$\hat{u}(x, \mathbf{q}(t)) = \sum_{i=1}^N q_i(t) u_i(x). \quad (52)$$

The standard Galerkin projection leads to the uncoupled ODEs,

$$\dot{q}_k = -\lambda_k q_k, \quad k = 1, 2, \dots, N, \quad (53)$$

which admit the exact solution $q_k(t) = q_k(0) \exp(-\lambda_k t)$. Therefore, the Galerkin method correctly predicts that the solutions decay to zero exponentially fast.

Next we derive the reduced-order equations by extremizing the action (50). Substituting the linear ansatz (52) in the Lagrangian \mathcal{J} , we obtain

$$\mathcal{J}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \int_{\mathbb{R}} |\hat{u}_t - \mathcal{L}\hat{u}|^2 dx = \frac{1}{2} \sum_{i=1}^N (\dot{q}_i^2 + 2\lambda_i \dot{q}_i q_i + \lambda_i^2 q_i^2). \quad (54)$$

We take the first variation of the functional $S[\mathbf{q}]$ with respect to perturbations $\delta\mathbf{q}(t)$, $t \in [0, T]$. The variations at the endpoints, $\delta\mathbf{q}(0)$ and $\delta\mathbf{q}(T)$, need to be prescribed. Since, in general, the true solution of the PDE is a priori unknown, the endpoint variations cannot be prescribed by taking the different between the ansatz solution and the true solution. Therefore, we adopt the usual assumption from classical mechanics that the variations vanish at the end points. Requiring the first variation of S to vanish leads to the Euler-Lagrange equations,

$$\frac{d}{dt} \frac{\partial \mathcal{J}}{\partial \dot{q}_k} - \frac{\partial \mathcal{J}}{\partial q_k} = \ddot{q}_k - \lambda_k^2 q_k = 0, \quad (55)$$

which admit the exact solution,

$$q_k(t) = \left[\frac{1}{2} q_k(0) + \frac{1}{2\lambda_k} \dot{q}_k(0) \right] e^{\lambda_k t} + \left[\frac{1}{2} q_k(0) - \frac{1}{2\lambda_k} \dot{q}_k(0) \right] e^{-\lambda_k t}. \quad (56)$$

Except for the special initial conditions $\dot{q}_k(0) = -\lambda_k q_k(0)$, the solutions grow exponentially fast. However, we know that the true solutions of the PDE must decay to zero. Even if the special initial condition $\dot{q}_k(0) = -\lambda_k q_k(0)$ is specified, numerical round-off errors will grow exponentially in time, rendering the solution unstable in practice. In the case of the nonlinear Schrödinger equation, we have observed (not presented here) a similar instability for reduced-order equations obtained from the finite-time error functional (50).

References

- [1] T. A. A. Adcock, R. H. Gibbs, and P. H. Taylor. The nonlinear evolution and approximate scaling of directionally spread wave groups on deep water. *Proc. R. Soci. A*, 468(2145):2704–2721, 2012.

- [2] T. A. A. Adcock and P. H. Taylor. Focusing of unidirectional wave groups on deep water: an approximate nonlinear Schrödinger equation-based model. *Proc. R. Soci. A*, 465(2110):3083–3102, 2009.
- [3] G. P. Agrawal. *Nonlinear fiber optics*. Springer, 5th edition, 2013.
- [4] H. Aref. Integrable, chaotic, and turbulent vortex motion in two-dimensional flows. *Ann. Rev. Fluid Mech.*, 15(1):345–389, 1983.
- [5] H. Babaei, M. Farazmand, G. Haller, and T. P. Sapsis. Reduced-order description of transient instabilities and computation of finite-time Lyapunov exponents. *Chaos*, 27(6):063103, 2017.
- [6] H. Babaei and T. P. Sapsis. A minimization principle for the description of modes associated with finite-time instabilities. *Proc. R. Soc. A*, 472(2186), 2016.
- [7] G. I. Barenblatt. *Scaling, Self-similarity, and Intermediate Asymptotics: Dimensional Analysis and Intermediate Asymptotics*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 1996.
- [8] J. T. Beale and A. Majda. High order accurate vortex methods with explicit velocity kernels. *J. Comput. Phys.*, 58(2):188–208, 1985.
- [9] T. B. Benjamin and J. E. Feir. The disintegration of wave trains on deep water part 1. theory. *J. Fluid Mech.*, 27(03):417–430, 1967.
- [10] P. Benner, S. Gugercin, and K. Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Review*, 57(4):483–531, 2015.
- [11] L. Bergé. Transient regimes of anisotropic light beam self-focusing in nonlinear dispersive media. *Physics Letters A*, 189(4):290–298, 1994.
- [12] K. Carlberg, R. Tuminaro, and P. Boggs. Preserving Lagrangian structure in nonlinear model reduction with application to structural dynamics. *SIAM Journal on Scientific Computing*, 37(2):B153–B184, 2015.
- [13] A. Chabchoub, N. P. Hoffmann, and N. Akhmediev. Rogue wave observation in a water wave tank. *Phys. Rev. Lett.*, 106(20):204502, 2011.
- [14] A. J. Chorin. Numerical study of slightly viscous flow. *Journal of Fluid Mechanics*, 57(4):785–796, 1973.
- [15] G.-H. Cottet and P. D. Koumoutsakos. *Vortex Methods: Theory and Practice*. Cambridge University Press, 2000.
- [16] W. Cousins and T. P. Sapsis. Unsteady evolution of localized unidirectional deep-water wave groups. *Phys. Rev. E*, 91(6):063204, 2015.

- [17] W. Cousins and T. P. Sapsis. Reduced-order precursors of rare events in unidirectional nonlinear water waves. *J. Fluid Mech*, 790:368–388, 3 2016.
- [18] S. M. Cox and P. C. Matthews. Exponential time differencing for stiff systems. *Journal of Computational Physics*, 176(2):430–455, 2002.
- [19] D. J. Dalle, M. L. Fotia, and J. F. Driscoll. Reduced-order modeling of two-dimensional supersonic flows with applications to scramjet inlets. *Journal of Propulsion and Power*, 26(3):545–555, 2010.
- [20] M. Desaix, D. Anderson, and M. Lisak. Variational approach to collapse of optical pulses. *J. Opt. Soc. Am. B*, 8(10):2082–2086, 1991.
- [21] K. Dysthe, H. E. Krogstad, and P. Müller. Oceanic rogue waves. *Annu. Rev. Fluid Mech.*, 40:287–310, 2008.
- [22] M. Farazmand and T. P. Sapsis. Dynamical indicators for the prediction of bursting phenomena in high-dimensional systems. *Phys. Rev. E*, 94:032212, 2016.
- [23] M. Farazmand and T. P. Sapsis. Reduced-order prediction of rogue waves in two-dimensional deep-water waves. *Journal of Computational Physics*, 340:418–434, 2017.
- [24] X. He, J. Li, C. A. Mader, A. Yildirim, and J. Martins. Robust aerodynamic shape optimization—from a circle to an airfoil. *Aerospace Science and Technology*, 87:48–61, 2019.
- [25] P. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley. *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge university press, 2012.
- [26] P. D. Lax. *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*. SIAM, 1973.
- [27] A. J. Majda and Y. Yuan. Fundamental limitations of ad hoc linear and quadratic multi-level regression models for physical systems. *Discrete & Continuous Dynamical Systems - B*, 17(1531-3492_2012_4_1333):1333, 2012.
- [28] H. K. Moffatt. The degree of knottedness of tangled vortex lines. *Journal of Fluid Mechanics*, 35(1):117–129, 1969.
- [29] P. K. Newton. *The N-vortex problem: analytical techniques*, volume 145 of *Applied Mathematical Sciences*. Springer, 2001.
- [30] V. M. Pérez-García, H. Michinel, J. I. Cirac, M. Lewenstein, and P. Zoller. Low energy excitations of a bose-einstein condensate: A time-dependent variational analysis. *Phys. Rev. Lett.*, 77:5320–5323, Dec 1996.

- [31] C. W. Rowley and S. T. M. Dawson. Model reduction for flow analysis and control. *Annual Review of Fluid Mechanics*, 49(1):387–417, 2017.
- [32] C. W. Rowley, I. G. Kevrekidis, J. E. Marsden, and K. Lust. Reduction and reconstruction for self-similar dynamical systems. *Nonlinearity*, 16(4):1257–1275, 2003.
- [33] V. P. Ruban. Anomalous wave as a result of the collision of two wave groups on the sea surface. *JETP Letters*, 102(10):650–654, 2015.
- [34] V. P. Ruban. Gaussian variational ansatz in the problem of anomalous sea waves: Comparison with direct numerical simulation. *Journal of Experimental and Theoretical Physics*, 120(5):925–932, 2015.
- [35] J. A. Samareh. Survey of shape parameterization techniques for high-fidelity multidisciplinary shape optimization. *AIAA Journal*, 39(5):877–884, 2001.
- [36] E. Shlizerman, E. Ding, M. O. Williams, and J. N. Kutz. The proper orthogonal decomposition for dimensionality reduction in mode-locked lasers and optical systems. *International Journal of Optics*, 2012:831604, 2012.
- [37] D. R Solli, C. Ropers, P. Koonath, and B. Jalali. Optical rogue waves. *Nature*, 450(7172):1054–1057, 2007.
- [38] R. R. Trieling, M. Becerks, and G. J. F. Van Heijst. Dynamics of monopolar vortices in a strain flow. *Journal of Fluid Mechanics*, 345:165–201, 1997.
- [39] V. E. Zakharov. Stability of periodic waves of finite amplitude on the surface of a deep fluid. *J. Appl. Mech. Tech. Phys.*, 9(2):190–194, 1968.