Model reduction methods for nuclear emulators

View the article online for updates and enhancements.

Guide

Model reduction methods for nuclear emulators

J A Melendez[1] , C Drischler[2] , R J Furnstahl[1],* , A J Garcia[1] and Xilin Zhang[2]

[1] Department of Physics, The Ohio State University, Columbus, OH 43210, United States of America
[2] Facility for Rare Isotope Beams, Michigan State University, MI 48824, United States of America

E-mail: melendez.27@osu.edu, drischler@frib.msu.edu, furnstahl.1@osu.edu, garcia.823@osu.edu and zhangx@frib.msu.edu

Abstract

The field of model order reduction (MOR) is growing in importance due to its ability to extract the key insights from complex simulations while discarding computationally burdensome and superfluous information. We provide an overview of MOR methods for the creation of fast & accurate emulators of memory- and compute-intensive nuclear systems, focusing on eigen-emulators and variational emulators. As an example, we describe how 'eigenvector continuation' is a special case of a much more general and well-studied MOR formalism for parameterized systems. We continue with an introduction to the Ritz and Galerkin projection methods that underpin many such emulators, while pointing to the relevant MOR theory and its successful applications along the way. We believe that this guide will open the door to broader applications in nuclear physics and facilitate communication with practitioners in other fields.

Keywords: model reduction, reduced-order models, emulators, Bayesian statistics, Galerkin projection, variational principles, complex simulations, Ritz projection, Eigenvector Continuation, nuclear physics

1. Introduction

Nuclear physics calculations often need to be repeated many times for different values of some model parameters, for example when sampling the model space for Bayesian uncertainty quantification [1–9] and experimental design [10–12]. The computational burden can be alleviated by using *emulators*, or surrogate models, which accurately approximate the response of the original (i.e., high-fidelity) model but are much cheaper to evaluate. The development

---

*Author to whom any correspondence should be addressed.

of emulators by nuclear physicists is a welcome addition to the community's toolkit, but such applications are neither the first nor the most sophisticated in the long history of *model reduction* for complex simulations.

The naïve implementation of realistic computer models in physics, mathematics, or engineering generally can demand an ever-increasing computational burden, but it has been known for decades that much of the information contributing to this burden is superfluous and can be *compressed* into a much more efficient form while retaining high accuracy. The broad and relatively mature field of model order reduction (MOR) has focused on exactly this problem of extracting the dominant information while excluding the costly, redundant information from simulations; the resulting emulator is known as the *reduced-order model*. Initially, the systems under consideration contained only fixed sets of model parameters, but later on the study of parametric MOR (PMOR) addressed the need to emulate systems as the parameter values changed.

The problems already addressed by PMOR are wide-ranging; for example, a sampling of applications in engineering, computational physics, and computer science is presented in reference [13]. The problems include large-scale systems of ordinary and partial differential equations with time dependence and nonlinearities, eigenvalue problems, and much more. Although these particular examples are classical, the nuclear physicist is likely to find problems in the MOR literature with analogous mathematical structure and goals to their own problems. In later sections we point to examples in the quantum realm such as bound-state and scattering problems and density functional theory (DFT).

Reduction schemes in the MOR literature can be classified as data-driven or model-driven, although hybrid approaches are also possible. Data-driven approaches typically need little to no understanding of the system under consideration, but rather rely on interpolating the output of the high-fidelity model. These are hence classified as *non-intrusive*. Examples include Gaussian processes [14] and dynamic mode decomposition (DMD) [15, 16], and there are further exciting developments along these lines coming from the machine learning literature (e.g., references [17–19]).

Alternatively, the model-driven approaches take the high-fidelity system of equations as given, from which they derive the reduced-order equations. Emulation via model-driven methods continue to be physics-based, respecting the underlying structure of the system, and are likely to extrapolate more effectively than data-driven emulators. Many model-based methods employ the concept of projection, where the high-dimensional system is projected onto a well-chosen low-dimensional manifold. The challenge of many model-driven approaches is related to their *intrusive* nature, i.e., they require writing new codes for projecting the system of interest into a reduced space.

Despite the mature literature on MOR, many of its aspects have been reinvented across multiple different disciplines. Indeed, the nuclear physics community has recently been exploiting an already-established model-driven, projection-based PMOR technique in the study of the eigenvalue problem. Introduced to the nuclear community as eigenvector continuation (EC) [20, 21], this method has been demonstrated to be highly effective for nuclear bound-state [3, 5, 22, 23] and scattering calculations [24–27]. EC has not been recognized in its broader context (see section 6.4) because the projection techniques common in the model reduction literature are not widely known in the nuclear physics community. Furthermore, many reviews of the subject presume a level of maturity with these tools that might make them difficult to digest.

We therefore provide a guide to projection-based PMOR for the nuclear physics community. Our goal in the present work is not to introduce specific new nuclear physics applications, but to make the existing literature more accessible. In particular, we explain how one can obtain

emulators by projecting generic differential equations and eigenvalue problems into a low-dimensional subspace, all while relating key ideas back to the broader MOR literature. By doing so, we seek to open the door to broader applications, allow the nuclear community to take advantage of the vast literature, and facilitate communication with practitioners in other fields.

In section 2 we begin by introducing important concepts and notation. We then provide an intuitive example of PMOR as applied to the eigenvalue problem in section 3 and place existing work in the nuclear physics literature into a broader context. Next, sections 4 and 5 provide two distinct methods for projecting differential equations onto effective low-dimensional manifolds via variational principles (or, the Ritz method) and the Galerkin projection, respectively. Finally, we return in section 6 to the recent developments made in the MOR community to discuss the incredible opportunities available to nuclear physicists interested in applying such techniques. An outlook is given in section 7.

2. Setting the stage

We would like to solve a differential equation or an eigenvalue problem where the operators are a function of parameters $\boldsymbol{\theta}$[3]. In the case of a differential equation, the goal is to obtain the solution ψ of

$$D(\psi; \boldsymbol{\theta}) = 0 \quad \text{in } \Omega, \tag{1a}$$

$$B(\psi; \boldsymbol{\theta}) = 0 \quad \text{on } \Gamma, \tag{1b}$$

where $\{D, B\}$ are operators, and $\{\Omega, \Gamma\}$ are the domain and boundary, respectively. Generalizations to systems of differential equations follows straightforwardly. For the eigenvalue problem the solutions are $\{E, |\psi\rangle\}$, which satisfy

$$H(\boldsymbol{\theta})|\psi\rangle = E|\psi\rangle \tag{2}$$

for a given Hermitian operator H. Throughout this work we switch between an abstract vector notation $|\psi\rangle$ and functions ψ with the representation dependencies suppressed. Time-dependence is permitted in these systems but is not explicitly considered here—see section 6 for pointers on how to handle these cases.

Here we consider systems where obtaining $|\psi\rangle$ (and E) will require a non-negligible amount of computing time, which becomes compounded when a range of parameter values are required, e.g., in a Monte Carlo sampler or an optimizer. The choice in PMOR is then made to spend compute resources in the *offline stage*, where the heavy lifting can be easily parallelized in many cases [28], such that emulation during the *online phase* can be performed efficiently. A critical component of the online–offline paradigm is the removal of all size-ψ operations during the online phase, a property available for operators with an affine parameter dependence. That is,

$$H(\boldsymbol{\theta}) = \sum_n h_n(\boldsymbol{\theta}) H_n, \tag{3}$$

or likewise for $D(\psi; \boldsymbol{\theta})$ or $B(\psi; \boldsymbol{\theta})$, where the operators can be written as a sum of products of parameter-dependent functions $h_n(\boldsymbol{\theta})$ and parameter-independent operators H_n. This factorization permits the operators to be projected *once* up front, rather than for every value of

---

[3] Note that the literature may use a different notation, including μ for the set of parameters.

$\boldsymbol{\theta}$ we would like to emulate, and will be discussed in more detail in the following sections. For non-linear systems and cases with non-affine parameters, various *hyperreduction* methods have been developed to augment the model reduction techniques (see sections 4.5 and 6.3).

The projection-based emulation approaches described here rely on (i) choosing an effective low-dimensional representation of ψ and (ii) writing equations (1) and (2) in integral form. For the first step one proposes a trial function comprised as a linear combination of a set of $N_{\rm b}$ known basis functions $\{\psi_i\}$

$$\widetilde{\psi} \equiv \sum_{i=1}^{N_{\rm b}} \beta_i \psi_i = X\vec{\beta}, \tag{4}$$

$$X \equiv \begin{bmatrix} \psi_1 & \psi_2 & \dots & \psi_{N_{\rm b}} \end{bmatrix}. \tag{5}$$

While the coefficients $\vec{\beta}$ are unspecified at this point, their values will become fixed after imposing conditions specific to the integral forms described below. We focus on bases X that are constructed out of *snapshots* [29–32], high-fidelity solutions $\psi_i = \psi(\boldsymbol{\theta}_i)$, at a set of parameter values $\{\boldsymbol{\theta}_i\}$. This requires that a high-fidelity solver for ψ exists, though the details of such a solver are irrelevant for our discussion. The snapshot approach has been found to construct highly efficient and system-specific trial bases across a wide range of cases. We will return to exactly how these snapshots can be chosen in section 6. This form of the trial function (4) will be used throughout the rest of this work. Colloquially, we will describe the space of $N_{\rm b}$ basis functions as the 'small space' and the space of ψ and its corresponding operators (D, B, and H) as the 'large space'.

One familiar integral form available to many differential equations comes from variational principles. Variational principles begin with the definition of a scalar functional $\mathcal{S}$ (the action) that can be written as

$$\mathcal{S}[\psi] = \int_\Omega \mathrm{d}\Omega \, F(\psi) + \int_\Gamma \mathrm{d}\Gamma \, G(\psi), \tag{6}$$

where F and G are known differential operators. The unknown function ψ is determined as the one that makes $\mathcal{S}$ *stationary*, i.e., $\delta\mathcal{S} = 0$, under arbitrary variations $\delta\psi$. This will form the basis of variational emulators discussed in section 4.

It is not always the case that equation (1) can be cast in the form of a variational principle as in equation (6). In this case we can instead turn to the *weak form* of equation (1) whose errors or residuals we aim to minimize [33]. This will form the basis of the more general Galerkin emulators discussed in section 5.

However, before exploring the more abstract variational and Galerkin approaches, we begin in section 3 with a concrete and (to physicists) familiar example of both: the eigenvalue problem. This contains many of the ideas fundamental to these more general cases but is likely a more accessible place to start.

3. Eigen-emulators

3.1. Derivation

The eigenvalue problem frequently appears in nuclear physics when solving the few- or many-body Schrödinger equation. We are often particularly interested in the ground state with energy $E_{\rm min}$ and the associated wave function, but the full spectrum of eigen-energies and wave functions may also be of interest. More generally, eigenvalues appear across the physical sciences

and engineering due to their appearance in the solution of differential equations, and come with a broad range of associated boundary conditions. Here we focus on converting the simple case to a form amenable to reduced-order modeling, and we point to the literature on variational forms of eigenvalue problems for more specific cases [34].

We can write down the following functional whose stationary solution approximates the eigen-energy, that is $\mathcal{S} = \mathcal{E}$, where

$$\mathcal{E}[\psi] = \langle\psi|H|\psi\rangle - \widetilde{E}(\langle\psi|\psi\rangle - 1) \tag{7}$$

and the normalization of the wave function has been imposed with a Lagrange multiplier $\widetilde{E}$. Under arbitrary variations $\delta\psi$, the change in $\mathcal{E}$ for Hermitian H can be written as

$$\delta\mathcal{E}[\psi] = 2\langle\delta\psi|[H - \widetilde{E}]|\psi\rangle - \delta\widetilde{E}(\langle\psi|\psi\rangle - 1). \tag{8}$$

Assuming $|\widetilde{\psi}\rangle = X\vec{\beta}$ as in equation (4), we then find that the values of $\vec{\beta}$ satisfying the stationary condition $\delta\mathcal{E}[\widetilde{\psi}_\star] \equiv 0$, denoted $\vec{\beta}_\star$, are those obtained from the following generalized eigenvalue problem

$$\widetilde{H}\vec{\beta}_\star = \widetilde{E}\widetilde{N}\vec{\beta}_\star, \tag{9}$$

$$\vec{\beta}_\star^\dagger\widetilde{N}\vec{\beta}_\star = 1, \tag{10}$$

where $\widetilde{H} \equiv X^\dagger H X$ is the Hamiltonian projected into the subspace spanned by X, and $\widetilde{N} \equiv X^\dagger X$ is the norm matrix. The meaning of $\widetilde{E}$ can be understood by substituting these relationships back into the variational form (7):

$$\mathcal{E}[\widetilde{\psi}_\star] = \widetilde{E}\vec{\beta}_\star^\dagger\widetilde{N}\vec{\beta}_\star = \widetilde{E}. \tag{11}$$

Thus $\widetilde{E}$ is an approximation to the energy. Equations (9) and (10), combined with equation (4), therefore constitute the reduced-order model for E and $|\psi\rangle$ projected to the $N_b \times N_b$ small space. In the context of the Rayleigh–Ritz method[4] $\widetilde{E}$ is known as the *Ritz value* while $\widetilde{\psi}_\star = X\vec{\beta}_\star$ is known as the *Ritz vector*.

We have seen how a variational principle for an eigenvalue problem can lead to a reduced-order model for both E and ψ. But there is a more general approach to arrive at exactly equations (9) and (10) that is often described as the Galerkin method. The Galerkin approach begins with the *weak* form of the eigenvalue problem (2):

$$\langle\phi|H - E|\psi\rangle = 0, \tag{12}$$

where ϕ is an arbitrary *test* function. It can be shown that if equation (12) holds for all ϕ, then equation (2) must hold as well: if $(H - E)|\psi\rangle$ has any non-zero elements, then one can immediately find a ϕ such that equation (12) does not hold.

To obtain a reduced-order model from equation (12) we would ideally like to map $\psi \to \widetilde{\psi}$ and find the $\vec{\beta}$ that satisfy equation (12) for all ϕ, hence ensuring the satisfaction of equation (2). Unfortunately this would result in an over-determined system because an arbitrary ϕ has many more degrees of freedom than $\widetilde{\psi}$. Alternatively, we can derive the Ritz subspace method by imposing that the error made by the trial eigenvector is orthogonal to the N_b-dimensional subspace $\mathcal{X}$ spanned by X:

$$H|\widetilde{\psi}\rangle - \widetilde{E}|\widetilde{\psi}\rangle \perp \mathcal{X}, \tag{13}$$

---

[4] See references [35, 36] for commentary on the history of the method name.

or likewise

$$\langle \phi | H - \widetilde{E} | \widetilde{\psi} \rangle = 0, \quad \forall \, \phi \in \mathcal{X}. \tag{14}$$

Note here that, due to a peculiarity of the eigenvalue problem, we had to make the replacement $E \to \widetilde{E}$ because the eigenvalue is an *output* of the system: if the span of $\mathcal{X}$ does not exactly contain $\psi(\boldsymbol{\theta})$ then $\widetilde{E} \neq E$ in general. Equation (14) is known as the Galerkin condition, and is equivalent to imposing that $\langle \psi_i | H - \widetilde{E} | \widetilde{\psi} \rangle = 0$ must hold for $i \in [1, N_b]$. This yields a system of N_b equations and N_b unknowns $\vec{\beta}$ and, together with the normalization condition, reduces exactly to equations (9) and (10). The Lagrange multiplier was not strictly necessary in this implementation of the Galerkin method because the normalization is irrelevant for equation (14), but see section 5 for a discussion on including constraints for other Galerkin problems.

The applications of the eigensystem reduced-order models to nuclear physics have benefited from the fact that Hamiltonians derived from chiral effective field theory (EFT) have the form of equation (3) due to their affine dependence on the parameters $\boldsymbol{\theta}$ called low-energy couplings. By projecting each $H_n \to \widetilde{H}_n \equiv X^\dagger H_n X$ in the offline stage, $\widetilde{H}(\boldsymbol{\theta})$ can be efficiently reconstructed for each new $\boldsymbol{\theta}$, and equation (9) rapidly solved, in the online stage. Furthermore, if the $h_n(\boldsymbol{\theta})$ are *smooth*, as they are for chiral EFT, then we have found that the low-dimensional representations exploited by the reduced-order models are often quite well satisfied, particularly when the dimension of H is large with respect to the dimension of $\boldsymbol{\theta}$. Furthermore, downstream observables benefiting from the same affine representation can be quickly emulated via $\langle \psi | O | \psi \rangle \approx \vec{\beta}_\star^\dagger [X^\dagger O X] \vec{\beta}_\star$ where the factor in brackets is computed and stored in the offline stage. Thus, the time spent 'training' the reduced-order model (constructing the basis X and projecting H and O) in the offline stage can lead to *multiple* fast & accurate emulators for the energy and other observables.

3.2. Results from other fields

The Hermitian and symmetric eigenvalue problems have been extensively studied in the mathematical literature and thus the ability to approximate eigenvalues and eigenvectors from a subspace is well known [37, 38]. For example, the convergence properties of eigenvalues and eigenvectors, along with their error bounds, are discussed under these subspace projections in the linear algebra and applied mathematics literature. These analyses extend beyond the extremal eigenvalues, but to all eigenvalues in the spectrum of H, where the applicability of subspace approaches to excited states has been clear for decades. The sense in which the Ritz values and vectors are *optimal* approximations is well known, and, e.g., discussed in references [37, 39].

Furthermore, eigenvalue problems had been studied in the field of PMOR well before EC was introduced into the nuclear literature. For example:

- Machiels *et al* [40] applied a parametric snapshot-based reduced-order approach to quickly evaluate eigenvalue problems, and subsequently proved theorems about the error bounds of the approximate eigenvalues. Horger *et al* [41] built upon this approach for error bounds of multiple eigenvalues and discusses efficient greedy algorithms for the basis generation.
- Pau [42] used this reduced-basis approach to quickly and accurately compute eigenvalues for band structure calculations.
- Buchnan *et al* [31] constructed a reduced-order model for parametric eigensystems in reactor physics.

- Cheng [43] constructed reduced-order models of the bound-state Schrödinger equation for electron wave functions in semiconductor nanostructures.
- Gräbner *et al* [44] emulated a parameterized non-linear eigenvalue problem using a snapshot-based reduced-order model for resonant frequencies.

Eigenvalue problems in PMOR appear during the greedy sampling (see section 6) and uncertainty quantification phases of reduced-order models for partial differential equations [45, 46]. These phases require, in part, the solution to a generalized eigenvalue problem to obtain the coercivity constant $\alpha(\boldsymbol{\theta})$ relevant for the error bounds [45–48], where $\alpha(\boldsymbol{\theta})$ is the minimum eigenvalue. A fast approximation for $\alpha(\boldsymbol{\theta})$ is needed across many values of $\boldsymbol{\theta}$ and hence calculations in the large space must be avoided. References [45, 46] considered the reduced-order model (equation (9)) for $\alpha(\boldsymbol{\theta})$ but ultimately proposed instead the so-called successive constraint method (SCM) due to its ability to provide a rigorous lower bound to $\alpha(\boldsymbol{\theta})$. Nevertheless, others still encourage the use of equation (9) due to its speed [49] and simplicity [30, chapter 4] over the SCM.

Each of the examples provided in this section contain all of the key ingredients of EC—creating snapshots at parameter values, projecting large eigensystems to the span of these snapshots, evaluating rapidly in an online phase, etc—but have been known as the reduced basis (RB) method in the model reduction community.

4. Variational emulators

4.1. Theory

Variational principles are ubiquitous in physics. Many differential equations have a corresponding action $\mathcal{S}$, where the solution to the differential equation also makes $\mathcal{S}$ stationary. This yields an alternate way of solving a set of PDEs: rather than solving the Euler–Lagrange equations themselves, one can instead find the solution that makes the action stationary under variations in ψ. The use of variational principles as a means to solve otherwise difficult problems dates back to Ritz [35, 36, 50, 51]. Thus, these methods often go under the names of (Rayleigh-)Ritz [35, 36], or are simply described as variational. But as we will see in section 5, the Galerkin approach is more general and hence these are occasionally named Ritz–Galerkin methods. Here we provide a brief description of how reduced-order models can arise from variational principles; for an extensive discussion of variational methods, see references [33, 52, 53].

One can derive a set of differential equations—Euler–Lagrange equations—from a variational principle (6) by enforcing $\delta\mathcal{S} = 0$ under arbitrary variations $\delta\psi$. However, such differential equations may require a fine grid or otherwise be expensive to solve. Instead we would like to obtain a set of reduced-order models directly from the variational principle. To do so we note that variations $\delta\psi$ can no longer be completely arbitrary given our choice of trial function in equation (4). Rather than stipulate that $\delta\mathcal{S} = 0$ for any arbitrary variation $\delta\psi$, we instead extract the optimal coefficients, $\vec{\beta}_\star$, as those for which $\mathcal{S}$ is stationary under variations in $\vec{\beta}$:

$$\delta\mathcal{S}[\widetilde{\psi}] = \sum_{i=1}^{N_{\mathrm{b}}} \frac{\partial\mathcal{S}}{\partial\beta_i} \delta\beta_i = 0. \tag{15}$$

This yields a set of N_{b} equations and unknowns $\vec{\beta}$ because the factor multiplying each $\delta\beta_i$ must be identically zero.

The general case would involve a numerical search for the solution to equation (15), but if $\mathcal{S}$ is quadratic in ψ then the solution can be determined analytically. We will focus here on the case of solving a differential equation, which will result in a linear problem to be solved,

because we have already tackled the eigenvalue problem in section 3. In this case, $\mathcal{S}$ can be written as

$$\mathcal{S}[\widetilde{\psi}] = \frac{1}{2}\langle\widetilde{\psi}|A|\widetilde{\psi}\rangle + \langle b|\widetilde{\psi}\rangle + c \tag{16}$$

$$= \frac{1}{2}\vec{\beta}^{\dagger}\widetilde{A}\vec{\beta} + \vec{b}\cdot\vec{\beta} + c, \tag{17}$$

where $\widetilde{A} = X^{\dagger}AX$, $b_i = \langle b|\psi_i\rangle$, and c is a constant. The quadratic portion could be made symmetric—if it is not already—by writing it as

$$\mathcal{S} = \frac{1}{2}\vec{\beta}^{\dagger}\widetilde{A}_s\vec{\beta} + \vec{b}\cdot\vec{\beta} + c, \tag{18}$$

$$\tilde{A}_s = \frac{\tilde{A} + \tilde{A}^{\intercal}}{2}, \tag{19}$$

which can be desirable for numerical purposes. It then follows that the optimal coefficients $\vec{\beta}_{\star}$ are those that satisfy

$$\delta\mathcal{S} = \widetilde{A}_s\vec{\beta}_{\star} + \vec{b} = 0, \tag{20}$$

which can be solved with standard linear algebra methods. This linear equation is of dimension N_b, the number of basis elements $\{\psi_i\}$, rather than of the much larger dimension of ψ itself. Therefore, as long as $\{\psi_i\}$ approximately span the space that ψ traces as a function of $\boldsymbol{\theta}$, the trial function constructed by equations (4) and (20) will be a fast & accurate emulator of ψ.

This method is particularly beneficial for quickly emulating many $\boldsymbol{\theta}$ values if both A and $|b\rangle$ are *affine* in $\boldsymbol{\theta}$, that is

$$A(\boldsymbol{\theta}) = \sum_n f_n(\boldsymbol{\theta})A_n, \tag{21}$$

$$|b(\boldsymbol{\theta})\rangle = \sum_n g_n(\boldsymbol{\theta})|b_n\rangle, \tag{22}$$

which need not contain the same number of terms and, from which,

$$\widetilde{A}(\boldsymbol{\theta}) = \sum_n f_n(\boldsymbol{\theta})\widetilde{A}_n, \tag{23}$$

$$\vec{b}(\boldsymbol{\theta}) = \sum_n g_n(\boldsymbol{\theta})\vec{b}_n \tag{24}$$

can be quickly reconstructed because $\widetilde{A}_n = X^{\dagger}A_nX$ and $b_{ni} = \langle b_n|\psi_i\rangle$ need only be computed once in the offline stage. This follows similarly from the discussion in section 3, but in this case both A and $|b\rangle$ can depend on the parameters of the system, rather than simply the Hamiltonian H.

Note that the number of basis elements N_b needed for an accurate $\widetilde{\psi}$ may be much smaller than one might naïvely expect. Even if the dimension of a high-fidelity solution ψ is quite large (e.g., due to a fine grid size in the differential equation solver), the space that ψ traces out as a function of $\boldsymbol{\theta}$ is often much smaller. Thus, constructing an emulator for accurately reproducing $\psi(\boldsymbol{\theta})$ can be achievable with a well-chosen basis X (see section 6).

4.2. Constraints

The derivation in section 4.1 assumed that ψ was unconstrained, but oftentimes one has to enforce a set of constraints $C_j(\psi) = 0$ for $j = 1, \ldots, N_c$. For example, in the eigenvalue problem example in section 3, we had to enforce that the wave function is normalized to one, i.e., $C(\psi) = \langle \psi | \psi \rangle - 1 = 0$. Constraints can be straightforwardly included in a variational principle via the method of Lagrange multipliers. Here, each constraint is appended as a term in the variational form with a corresponding λ_j, i.e., $\lambda_j C_j(\psi)$. When imposing stationarity, each of these terms yields a $\delta \lambda_j C_j(\psi) + \lambda_j \delta C_j(\psi)$ contribution to δS. Specifically, equation (15) needs to be rewritten as

$$\delta S[\widetilde{\psi}] = \sum_{i=1}^{N_b} \frac{\partial S}{\partial \beta_i} \delta \beta_i + \sum_{j=1}^{N_c} \frac{\partial S}{\partial \lambda_j} \delta \lambda_j = 0, \tag{25}$$

from which follows a set of $N_b + N_c$ equations. Enforcing these constraints then yields a larger system of equations to solve when emulating ψ in the online phase, though in some systems it is possible to solve for λ_j in terms of ψ, hence reducing the problem back to its original size while still incorporating the constraints [33]. Solving for λ_j to remove it from S is beneficial if possible because (i) it will decrease the size of the linear system to be solved and (ii) it can make the system better conditioned numerically.

4.3. Concrete example

At this point it is helpful to provide a concrete example of a variational principle that leads to a reduced-order model. Here we provide a simple projection example without complicating details or abstract notation; see reference [33] for more examples.

Consider the functional, given functions g and f [54],

$$S = \int_\Omega d\Omega \left[\frac{1}{2} \nabla \psi \cdot \nabla \psi - g \psi \right] - \int_\Gamma d\Gamma \, f \psi. \tag{26}$$

Under an infinitesimal variation $\delta \psi$, the change δS is then

$$\delta S = \int_\Omega d\Omega [\nabla \delta \psi \cdot \nabla \psi - g \delta \psi] - \int_\Gamma d\Gamma \, f \delta \psi \tag{27}$$

$$= \int_\Omega d\Omega \, \delta \psi \left[-\nabla^2 \psi - g \right] + \int_\Gamma d\Gamma \, \delta \psi \left[\frac{\partial \psi}{\partial n} - f \right], \tag{28}$$

from which it follows that the stationary solution is

$$-\nabla^2 \psi = g \quad \text{in } \Omega, \tag{29}$$

$$\frac{\partial \psi}{\partial n} = f \quad \text{on } \Gamma, \tag{30}$$

which is exactly the Poisson equation with Neumann boundary conditions.

By instead starting with $S[\widetilde{\psi}]$ and imposing equation (15), then it follows that (with summed indices implied)

$$\delta S = \delta \beta_i \left[\int_\Omega d\Omega \left[(\nabla \psi_i) \cdot (\nabla \psi_j) \beta_j - g \psi_i \right] - \int_\Gamma d\Gamma \, f \psi_i \right]. \tag{31}$$

With $\widetilde{A}_{ij} \equiv \int_\Omega d\Omega (\nabla \psi_i) \cdot (\nabla \psi_j)$, $g_i \equiv \int_\Omega d\Omega \, g \psi_i$, and $f_i \equiv \int_\Gamma d\Gamma \, f \psi_i$ the reduced-order model becomes

$$\widetilde{A}\vec{\beta}_\star = \vec{g} + \vec{f}, \tag{32}$$

which is an explicit example of equation (20). This could provide a fast & accurate emulator of $\psi(\boldsymbol{\theta}) \approx X\vec{\beta}_\star(\boldsymbol{\theta})$ for systems where g or f are affine functions of the parameters $\boldsymbol{\theta}$.

4.4. Results from other fields

Many examples from sections 3.2 and 5.4 could be listed here due to the relationship between emulators constructed from variational principles and Galerkin methods. Instead we highlight an explicitly variational problem particularly relevant for nuclear systems: DFT.

Computing the ground state energies in systems of nuclei and/or electrons reduces again to a minimization problem, where the wave functions—or equivalently densities—are those that minimize the non-linear functional $\mathcal{E}$. These minimization problems appear in quantum chemistry and nuclear physics and are often approached in the DFT framework [55, 56]. The emulation of DFTs via reduced-order models has been studied in the quantum chemistry literature, e.g., in references [57, 58] (see also references [42, 59, 60]), where an empirical interpolation method (EIM) (see section 6) was used to avoid issues with the nonlinearities of $\mathcal{E}$. This interpolation method permitted the reduced-order model to project all large-space operators to the reduced space of snapshots up front, which is a critical step to retain an efficient online-offline decomposition in the emulator.

4.5. Results from nuclear physics

We have already discussed how the eigenvalue problem can be projected to a subspace by starting with a variational principle in section 3, which corresponds to computing the energy spectrum and wave functions in a bound nuclear system. Its suitability for uncertainty quantification in low-energy nuclear physics has been demonstrated, e.g., in chiral EFT applications to few- and many-body systems, where it has been described as EC [3, 5, 22]. Variational approaches for reduced-order models [24, 26, 27] have also been very successfully applied to *scattering* states via the Kohn variational principle [61], where the Schrödinger equation is no longer an eigenvalue problem (for a brief review see reference [62]). Furthermore, a new approach for emulating directly scattering K or T matrices (without trial wave functions) was proposed in reference [25] based on the Newton variational principle (NVP) [63]. Each of these methods has benefited from an affine parameterization of the Hamiltonian $H(\boldsymbol{\theta})$, which permitted fast & accurate emulation of nuclear observables.

In addition, the emulation of scattering states with non-affine parameterization was also studied in reference [27], where $\widetilde{A}(\boldsymbol{\theta})$ (defined in equation (17)) does not satisfy equation (23). In order to reduce the computing costs in the online phase, a Gaussian process (GP) [14] was used to emulate $\widetilde{A}(\boldsymbol{\theta})$ across the parameter space. The training of the GP increased the offline computing costs, but the online costs were similar to those with affine parameterizations. Since the $\boldsymbol{\theta}$ dependence for $\widetilde{A}(\boldsymbol{\theta})$ is much smoother than those for $\psi(\boldsymbol{\theta})$ and $\widetilde{\psi}(\boldsymbol{\theta})$, the number of GP training points are kept to a minimum while still achieving great accuracy.

To our knowledge, the first application of projection-based PMOR as applied to nuclear DFT was presented in reference [64] (see reference [65] for the code). Here, a trial density was proposed as a linear combination of exact densities $\{\rho_i\}$ at a set of training locations $\boldsymbol{\theta}_i$. The coefficients $\vec{\beta}$ were found as those that minimized the energy of the system given a set of parameters at which to emulate. Because this is a non-quadratic variational problem, such

coefficients were found empirically via an optimizer. The goal of this work was to use a proof of principle to advocate for the adoption of such projection-based tools by the nuclear community outside of bound-state and scattering systems connected to Hamiltonians.

5. Galerkin emulators

5.1. Theory

The Galerkin approach, also more broadly called the 'method of weighted residuals', relies on the *weak* formulation of the differential equations in equation (1). To obtain the weak form, the differential equation and boundary condition are multiplied by arbitrary test functions ϕ and $\bar{\phi}$, integrated over the domain and boundary, and their sum set equal to zero:

$$\int_\Omega d\Omega\, \phi D(\psi) + \int_\Gamma d\Gamma\, \bar{\phi} B(\psi) = 0. \tag{33}$$

If equation (33) holds for all ϕ and $\bar{\phi}$, then equation (1) must be satisfied as well. The form of equation (33) is often rewritten using integration by parts to reduce the order of derivatives and to simplify the solution. Importantly, the weak form has the integral form needed for our emulator application. The weak form and its Galerkin projection are used extensively, for example, in the finite element method; see references [33, 66, 67] for an in-depth study and list of examples. For discussion of the convergence properties of the Galerkin method, its relation to abstract variational problems, and other salient mathematical details, see references [52, 68–70]. Here we follow the introduction of Galerkin methods as provided in reference [33].

Starting with the weak form, we can begin to construct an emulator that avoids the need for an explicit variational principle. It begins by first noting that substituting our trial function (4) into $D(\psi)$ and $B(\psi)$ will not in general satisfy equation (1) regardless of the choice of $\vec{\beta}$. Therefore, there will be some *residual*, and the goal is to find $\vec{\beta}_\star$ which minimize that residual across a range of test functions ϕ and $\bar{\phi}$. This system would be over-determined in the case of truly arbitrary test functions, so instead we propose the test bases

$$\phi = \sum_{i=1}^{N_b} \delta\beta_i \phi_i, \qquad \bar{\phi} = \sum_{i=1}^{N_b} \delta\beta_i \bar{\phi}_i, \tag{34}$$

where $\delta\beta_i$ are arbitrary parameters, not related to β_i. The $\delta\beta_i$ will play the same role as those in equation (15), namely as a bookkeeping method for determining the set of equations that are equivalently zero. By enforcing that the residuals against these test functions vanish for arbitrary $\delta\beta_i$, the bracketed expression in

$$\delta\beta_i \left[\int_\Omega d\Omega\, \phi_i D(X\vec{\beta}_\star) + \int_\Gamma d\Gamma\, \bar{\phi}_i B(X\vec{\beta}_\star) \right] = 0, \tag{35}$$

is zero for all $i \in [1, N_b]$, from which the optimal $\vec{\beta}_\star$ are extracted. Because this approximately satisfies the weak formulation, we have found an approximate solution to equation (1).

In a variety of cases [33], the test function basis is chosen to coincide with the trial function basis X, i.e., $\phi_i = \bar{\phi}_i = \psi_i$. This particular choice is known as *the* Galerkin method, but it is sometimes further specified as the Ritz–Galerkin or Bubnov–Galerkin methods. However, the method of weighted residuals is more general than the variational methods described in

section 4 because the test space need not be equivalent to the trial space (i.e., $\phi_i \neq \psi_i$). In these cases, the approach is described as the Petrov–Galerkin method [33]; this can result in more efficient emulators for some differential equations [67].

Under the Ritz–Galerkin assumption for the test space we can derive the reduced-order model for the case of a linear operator: $D(\psi) = D|\psi\rangle + |b\rangle$. If we ignore the boundary condition for simplicity, it then follows from equation (35) that

$$\widetilde{D}\vec{\beta}_\star + \vec{b} = 0, \tag{36}$$

where $\widetilde{D} = X^\dagger D X$ and $b_i = \langle b|\psi_i\rangle$. Just like in section 4, we have arrived at a linear problem for the solution to $\vec{\beta}_\star$ and insofar as N_b is small compared to the size of ψ, this will yield improvements to the time it takes to obtain a solution for $\vec{\beta}_\star$. Further speedups are available if D and $|b\rangle$ are affine in the parameters $\boldsymbol{\theta}$ so that $\widetilde{D}$ and $\vec{b}$ can be efficiently recomputed in the online phase—see sections 3 and 4.

5.2. Concrete example

Here we repeat the example provided in section 4.3 but instead start from the Poisson equation (equations (29) and (30)) and then derive the weak form.

First we multiply each equation by a test function $\phi = \bar{\phi}$, integrate over the respective domains, and add the equations together:

$$\int_\Omega d\Omega \, \phi \big[-\nabla^2\psi - g \big] + \int_\Gamma d\Gamma \, \phi \left[\frac{\partial\psi}{\partial n} - f \right] = 0. \tag{37}$$

Next we use the divergence theorem to symmetrize the system and to reduce the order of the derivatives:

$$\int_\Omega d\Omega [\nabla\phi \cdot \nabla\psi - g\phi] - \int_\Gamma d\Gamma \, f\phi = 0, \tag{38}$$

which is the weak form we desire. Finally, by asserting that equation (38) holds for $\psi \to \widetilde{\psi} = X\vec{\beta}$ and $\phi = \sum_i \delta\beta_i \psi_i$ for $i \in [1, N_b]$, then we have its discretized form

$$\delta\beta_i \left[\int_\Omega d\Omega \big[\nabla\psi_i \cdot \nabla\psi_j \beta_j - g\psi_i \big] - \int_\Gamma d\Gamma \, f\psi_i \right] = 0, \tag{39}$$

which is exactly equations (31) and (32) found via the variational approach!

5.3. When Galerkin coincides with variational emulators

We have already seen that the eigen-emulators of section 3 could be derived by both variational and Galerkin procedures, and we have found that the exact same reduced-order model for the Poisson equation arises in both the variational and the Galerkin procedures. This raises the question: when is the variational approach equivalent to the Galerkin approach?

The answer becomes clear if one restricts to cases where equation (1) is the Euler–Lagrange equation derived from the action (6). By the definition of the Euler–Lagrange equations, this statement is equivalent to

$$\delta\mathcal{S}[\psi] = 0 = \int_\Omega d\Omega \, \delta\psi D(\psi) + \int_\Gamma d\Gamma \, \delta\psi B(\psi). \tag{40}$$

We can then consider changing variables from the trial stationarity condition (15)

$$\delta\mathcal{S}[\widetilde{\psi}] = 0 = \sum_{i=1}^{N_b} \frac{\partial\mathcal{S}}{\partial\beta_i}\delta\beta_i = \sum_{i=1}^{N_b}\left[\frac{\partial\mathcal{S}}{\partial\psi}\right]_{\psi=\widetilde{\psi}}\frac{\partial\widetilde{\psi}}{\partial\beta_i}\delta\beta_i, \tag{41}$$

where the first two factors on the right-hand side are equation (40) with $\delta\psi \to \partial\widetilde{\psi}/\partial\beta_i = \psi_i$ and thus

$$0 = \delta\beta_i\int_\Omega d\Omega\,\psi_i D(X\vec{\beta}_\star) + \delta\beta_i\int_\Gamma d\Gamma\,\psi_i B(X\vec{\beta}_\star), \tag{42}$$

for all $i \in [1, N_b]$. This is exactly equation (35) under the Galerkin assumption that $\phi_i = \bar{\phi}_i = \psi_i$. Note that $\mathcal{S}$ was not assumed to have a quadratic form for this derivation, and no linearity assumptions were made about D or B.

There are some final steps to proving that Galerkin is strictly more general than a variational (Ritz) approach. The first is that all variational principles have corresponding Euler–Lagrange equations. This can be shown via the fundamental lemma of the calculus of variations. Second, there is the question of how Lagrange multipliers are to be introduced in the weak form without first starting with a variational principle. In this case, one simply adds the relevant constraint terms to the weak form. That is, if the constraint $C(\psi) = 0$ would appear as $\lambda C(\psi)$ in the functional $\mathcal{S}$, then one would simply add $\delta\lambda C(\psi) + \lambda\delta C(\psi)$ to the weak form and use the fact that $\delta\widetilde{\psi} = X\delta\vec{\beta}$. Therefore, if the differential equations to be solved directly correspond to the Euler–Lagrange equations of an action, then the variational approach is identical to the Galerkin approach. The generality of the Galerkin method comes from the fact that one need not assert that $\phi_i = \bar{\phi}_i = \psi_i$.

Although we have shown that the Galerkin method is more general than the Ritz approach, there is still value to obtaining a variational principle and deriving a reduced-order model from it. First, the functional $\mathcal{S}$ is often physically meaningful in its own right, and hence the variational emulators provide a straightforward way of quickly computing its value. For example, in the NVP approach to emulating the scattering K (or T) matrix [25], both the trial function $\widetilde{K}$ and the NVP functional $\mathcal{K}[\widetilde{K}]$ at the stationary point are estimates of K, but the variational principle has better error properties: if $\widetilde{K}$ has an error of $\mathcal{O}(\delta K)$ then $\mathcal{K}$ has an error of order $\mathcal{O}(\delta K^2)$. Thus, obtaining and using the variational principle as an emulator for K is superior to simply applying the Galerkin method. Second, the variational emulators are guaranteed to provide symmetric matrices when solving for $\vec{\beta}$. This feature can provide numerical benefits both when constructing the relevant matrices and when solving for $\vec{\beta}$.

5.4. Results from other fields

We have found that examples of reduced-order models relying on a Galerkin projection are numerous and exist across a multitude of disciplines. A wide array of such examples are given in reference [13], which is an entire volume of a three-part series dedicated to the applications of MOR in diverse settings, ranging from engineering to life sciences. A comprehensive list of reduced-order models built from Galerkin methods would be impossible; we instead provide a curated list of helpful articles below.

A broad survey on parametric model reduction [28] cites multiple examples of highly successful applications of reduced-order models, some of which we highlight here. In the thermal modeling of electric motors that depends on 20 parameters, Bruns and Benner [71] use model reduction for a speedup over 300–500 times the high-fidelity model. Lassila *et al* [72] display the power of reduced-order modeling in the study of nonlinear viscous flows by reducing

high-fidelity models of dimension $>10\,000$ to emulators of dimensions 8–20, with minimal accuracy impact and speedup factors of up to 450. Convection–diffusion models arise in the investigation of contaminant transport, where Lieberman *et al* [73] exploit MOR to reduce the model order from over one million down to 800 with negligible accuracy loss and a 3000 speedup factor.

Rozza *et al* [46] provide an illuminating introduction and motivation to reduced-order models built from Galerkin projections. Its applications include heat conduction and convection–diffusion, inviscid flow, and linear elasticity systems. Beyond the specific applications listed here, reference [46] describes how to effectively choose parameter values $\boldsymbol{\theta}$ for building the basis via a greedy algorithm, explores issues of convergence and error bounds, and performs an analysis of computational costs.

Chen *et al* [47] review the state of the RB method literature, collecting many of the main ingredients necessary for building and analyzing reduced-order models. They demonstrate their claims via benchmark problems and describe generalizations to time-dependent systems, risk prediction, Bayesian inverse problems, and more. The topic of uncertainty quantification is addressed in detail, which, for example, allows for effective placement of snapshot locations $\psi(\boldsymbol{\theta}_i)$ via a greedy algorithm.

6. The model reduction framework

Projection-based PMOR consists of (1) sampling across parameters for snapshot candidates, (2) creating the snapshots and generating a basis X, and (3) using the basis to construct the reduced system [28]. sections 4 and 5 partially address step (3) by providing two closely related methods of constructing these projected systems, one based on variational principles and another based on the Galerkin method. Here we take a bird's-eye view of the parametric model reduction workflow to show where different assumptions can lead to different types of reduced-order models, and to point to extensions at the cutting edge of the model reduction literature. For in-depth reviews, see, e.g., references [13, 28–30, 74].

6.1. Sampling parameters

Many reduced-order emulators rely on the concept of snapshots, or exact solutions at particular values across a range of parameters $\boldsymbol{\theta}$. So then the question becomes how to wisely choose the snapshot locations $\{\boldsymbol{\theta}_i\}$. For parameter spaces that are not too large, one could employ a space-filling design (such as a Latin-hypercube or grid-based approach), or one could center the design close to the range of parameter values that will be emulated. Define the finite number of points sampled from the space as N_{sample}. Ultimately, we will want a set of $N_{\text{b}} \leqslant N_{\text{sample}}$ snapshots to construct the RB, but the manner in which the snapshots are generated from the N_{sample} training parameters differs based on the basis construction method.

6.2. Constructing a basis

The next step is to take the snapshots and to construct a basis X given the set of training parameters. Notably, we have mostly restricted our attention to systems without time dependence. In the static, time-independent case, the sampling need only occur in the space of parameters $\boldsymbol{\theta}$. Here we discuss two approaches for determining the 'optimal' N_{b} basis vectors: proper orthogonal decomposition (POD) and greedy algorithms.

POD (also known as principal component analysis) [75] is an explore-and-compress strategy used to extract the most important basis vectors from a set of snapshots [68]. It computes

snapshots $\psi(\boldsymbol{\theta}_i)$ at *all* N_{sample} parameter values and subsequently keeps only the N_{b} most important vectors. It performs a singular value decomposition of the N_{sample} snapshot vectors and then removes those least important to the spanning set, i.e., those with the smallest eigenvalues. Often the set of important vectors are chosen such that the percent of the remaining 'energy' (sum of the eigenvalues) is large, say, 99% relative to the total set of vectors. The orthonormalization performed during the POD step is also helpful during the online phase of the emulator because it can improve the conditioning of the system. Because POD evaluates the snapshots at all of the proposed N_{sample} parameter values before then compressing the information, this can either be wasteful of computing resources or severely limit the size of N_{sample}. Hence, we now turn to an alternative.

The greedy algorithm for basis generation is an iterative approach [46, 47, 68] that does not evaluate the high-fidelity model at all N_{sample} parameter values. At each step the next location $\boldsymbol{\theta}_i$ to take a snapshot is chosen to be that which is expected to minimize the error in the emulator. Critical to such an approach is a fast approximation to the emulator error. Uncertainty quantification for reduced-order models have been well studied and, e.g., are available for parabolic and elliptic PDEs [47, 68], and the eigenvalue problem [41, 76]. At each step, the error at the set of N_{sample} parameter values is estimated, and that with the largest expected error is then evaluated with the high-fidelity model and appended to the basis. The search stops once the desired error tolerance has been achieved, or after a given number of steps. Because the error estimate is built to be much faster than the high-fidelity model, this approach is often much more efficient than the POD method.

Now we move on to the time-dependent cases, where sampling can occur not only in parameter space, but also in time (or frequency). In the time-dependent case there are more options, such as the rational interpolation method [77, 78], whose snapshots will include samples from the frequency domain and the parameter domain. Additionally, there exist POD variants for both the time and the frequency domains [28]. Note that mixed approaches exist: for example, in a POD-greedy approach one can opt to use a greedy algorithm in an 'outer loop' in parameter search, while using a POD-based approach to evaluate each time snapshot before discarding the least important [68]. Another common technique is known as balanced truncation, which creates a set of reduced models at each parameter value, which are then interpolated to create an emulator across $\boldsymbol{\theta}$ [28].

6.3. Building the reduced-order model

Finally, the reduced model must be created from the basis. In the examples shown above, where the snapshots are collected into a single basis X and the operators are projected to this basis, this is straightforward. This can work well if the system is linear and the operators are affine with respect to the parameters $\boldsymbol{\theta}$, which critically permits the reduced-order model to be independent of the size of the high-fidelity space during the online phase. However these conditions are not always satisfied.

In the non-affine or non-linear cases, one can turn to so-called *hyperreduction* approaches to construct approximate affine representations, which trade accuracy for speed [32]. These can be classified into the 'approximate then project' or the 'project then approximate' classes. The approximate-then-project approach first approximates the non-affine or non-linear operators as a linear combination of affine operators whose coefficients are to be determined, and whose operators can then be projected via X into the small space up front. Some common approximate-then-project methods [68] include the EIM [79, 80], the discrete EIM [81, 82], and the gappy-POD method [83, 84]. Project-then-approximate approaches have been developed

more recently and attempt to interpolate the basis X or the projected operators $\widetilde{H}(\boldsymbol{\theta}) = X^\dagger H(\boldsymbol{\theta})X$ themselves; see references [30, 85–87] for examples. Outside of explicit projection-based approaches, non-intrusive methods have been proposed for dealing with non-linear and non-affine systems, including the use of machine learning tools to approximate the basis coefficients $\vec{\beta}$ or projected operators $\widetilde{H}$, e.g., see references [27, 88].

But in fact, one need not create one basis and one reduced system; rather one could partition X into multiple bases across the span of $\boldsymbol{\theta}$. This can help alleviate the curse of dimensionality in the parameters $\boldsymbol{\theta}$ and the computational costs during the offline stage. Additionally, it permits the use of variable fidelity bases in different parts of the domain. If one opts to create a set of local systems, each component must be then coupled with one another across the interfaces. For a survey on this topic, see reference [30].

6.4. Discussion

Now that we have introduced many of the common steps in model reduction, we can begin to contextualize other named methods, such as the RB method and EC. The RB method was initially introduced in reference [89] and has found widespread use in the emulation of PDEs in a reduced-order approach [68]. Due to its similarities with many of the methods discussed here and in the model reduction literature, it may be difficult to distinguish between the RB method and more general model reduction techniques.

In fact, the RB method corresponds to specific choices in the model reduction framework [28]. First, the parameter set for the RB method is often chosen using a greedy algorithm with the help of a fast error estimate, though POD approaches are sometimes adopted, particularly in time-dependent systems where a POD-greedy combination is employed [68]. Next, a single basis X is constructed out of snapshots and often orthonormalized for stability. Finally, the RB model is built from a global basis projection, i.e., the same basis is used for the entire space of $\boldsymbol{\theta}$. These are but one of many choices that can be made at each step in the construction of a reduced-order model.

Likewise, we are able to help place EC into its proper context. EC is a parametric reduced-order model for an eigenvalue problem. It uses a global basis that is constructed with a snapshot-based POD approach. The 'active learning' approach proposed in reference [76] is the inclusion of a greedy sampling algorithm to obtain the next parameter value $\boldsymbol{\theta}_i$. Each of these are common throughout the model reduction literature and have been studied for eigenvalue problems; therefore EC is a specific implementation of the RB method to construct a reduced-order model.

We conclude by noting that reduced-order emulators have pros and cons. First, these emulators work better in some systems than others—where the quantity of interest lies in on a low-dimensional manifold compared to the size of ψ, and where operations on the large space of ψ can be avoided during the online phase. As one might expect, the relative size of the low-dimensional representation depends on the specific differential equation. For example, highly non-linear equations may not permit the same low-dimensional representation. Further, how the operators depend on the parameters could be critical to the effectiveness of the reduced system in two distinct ways: (i) the relative smoothness of the parameter dependence impacts the ability of ψ to live in a low-dimensional representation, and (ii) affine parameter dependence, or at least an effective hyperreduction approach, is critical to avoiding matrix multiplication in the large space of ψ. Lastly, one should not overlook how the quantity of interest itself affects the quality of the low-dimensional representation. For example, E has better convergence properties than $|\psi\rangle$ in the eigenvalue problem; the RB literature discusses the improvements of focusing on *compliant* quantities of interest [30, 47].

17

Table 1. A sampling of recent MOR software libraries; see reference [13, section 13.3] for an extensive listing.

Library	Language	Website	Remark
pyMOR[a] [90]	Python	pymor.org	Focuses on RBMs for parameterized PDEs; integrates with external PDE solvers
libROM	C++	librom.net	Library for efficient MOR techniques and physics-constrained data-driven methods; includes POD, DMD, projection-based ROM, hyper-reduction, greedy algorithm
MORLAB [91]	MATLAB	mpi-magdeburg.mpg.de/projects/morlab	MOR of dynamical systems based on the solution of matrix equations using spectral projection methods
modred [92]	Python	modred.readthedocs.io	Library for computing modal decompositions and ROMs, including POD, DMD, and Petrov–Galerkin projection
pyROM [93]	Python	github.com/CurtinIC/pyROM	Framework that employs Python visualisation tools; includes POD and DMD
pressio [94]	C++	pressio.github.io	Minimally-intrusive interface for MOR routines, including Galerkin projections

[a]See also the website of the Model Reduction for Parametrized Systems (MoRePaS) collaboration: morepas.org.

Beyond the considerations of emulator quality, building an intrusive emulator for complex systems can be a challenge. There has been great progress in building general software tools for practitioners in reduced-order modeling [13, 48]. Table 1 provides a sampling of recent MOR software libraries. Continuing to develop and publicize such tools will permit greater acceptance of these powerful methods in nuclear physics.

7. Outlook

The present work provides for nuclear physicists accessible pointers to some of the relevant literature on projection-based model reduction and shows that solving the parametric eigenvector problem from a subspace is merely a special case of a much broader set of tools. By properly contextualizing the methods under the projection-based model reduction umbrella, we have shown how reduced-order models built from a variational principle relate to an equally vast literature on the Galerkin method, which is even more general. We have not proposed anything novel; rather our message is that much information on emulators, some partially rediscovered and more not-yet-applied, can be at our fingertips if we look more widely.

We have shown that the 'RB method' is the established name of the methods described in the nuclear physics literature as EC, and suggest its adoption. We believe that putting EC in a more general context, and particularly using a unified naming convention, will not only alleviate confusion due to a conflict of terms used in other fields, but will permit access to a much broader literature. It would surely accelerate progress in the application of emulators in the nuclear community [95] and facilitate fruitful external collaborations.

Acknowledgments

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

J A Melendez ⓘ https://orcid.org/0000-0003-1359-1594
C Drischler ⓘ https://orcid.org/0000-0003-1534-6285
R J Furnstahl ⓘ https://orcid.org/0000-0002-3483-333X
A J Garcia ⓘ https://orcid.org/0000-0003-1723-3225
Xilin Zhang ⓘ https://orcid.org/0000-0001-9278-5359

References

[1] Neufcourt L, Cao Y, Nazarewicz W, Olsen E and Viens F 2019 *Phys. Rev. Lett.* **122** 062502
[2] King G B, Lovell A E, Neufcourt L and Nunes F M 2019 *Phys. Rev. Lett.* **122** 232502

[3] Ekström A and Hagen G 2019 *Phys. Rev. Lett.* **123** 252501
[4] Catacora-Rios M, King G B, Lovell A E and Nunes F M 2021 *Phys. Rev.* C **104** 064611
[5] Wesolowski S, Svensson I, Ekström A, Forssén C, Furnstahl R J, Melendez J A and Phillips D R 2021 *Phys. Rev.* C **104** 064001
[6] Svensson I, Ekström A and Forssén C 2022 *Phys. Rev.* C **105** 014004
[7] Odell D, Brune C R and Phillips D R 2022 *Phys. Rev.* C **105** 014625
[8] Djärv T, Ekström A, Forssén C and Johansson H T 2022 *Phys. Rev.* C **105** 014005
[9] Alnamlah I K, Pérez E A C and Phillips D R 2022 arXiv:2203.01972 [nucl-th]
[10] Melendez J A, Furnstahl R J, Grießhammer H W, McGovern J A, Phillips D R and Pratola M T 2021 *Eur. Phys. J.* A **57** 81
[11] Phillips D R *et al* 2021 *J. Phys. G: Nucl. Part. Phys.* **48** 072001
[12] Farr J N, Meisel Z and Steiner A W 2021 arXiv:2111.11536 [nucl-th]
[13] Benner P, Schilders W, Grivet-Talocia S, Quarteroni A, Rozza G and Miguel Silveira L 2020 *Model Order Reduction: Volume 3: Applications* (Berlin: de Gruyter & Co)
[14] Rasmussen C E and Williams C K I 2006 *Gaussian Processes for Machine Learning* (*Adaptive Computation and Machine Learning Series*) (Cambridge, MA: University Press Group Limited)
[15] Mezić I 2013 *Annu. Rev. Fluid Mech.* **45** 357
[16] Kutz J N, Brunton S L, Brunton B W and Proctor J L 2016 *Dynamic Mode Decomposition* (*Other Titles in Applied Mathematics*) (Philadelphia, PA: SIAM)
[17] Raissi M, Perdikaris P and Karniadakis G E 2019 *J. Comput. Phys.* **378** 686
[18] Chen W, Wang Q, Hesthaven J S and Zhang C 2021 *J. Comput. Phys.* **446** 110666
[19] Fresca S and Manzoni A 2022 *Comput. Methods Appl. Mech. Eng.* **388** 114181
[20] Frame D, He R, Ipsen I, Lee D, Lee D and Rrapaj E 2018 *Phys. Rev. Lett.* **121** 032501
[21] Sarkar A and Lee D 2021 *Phys. Rev. Lett.* **126** 032501
[22] König S, Ekström A, Hebeler K, Lee D and Schwenk A 2020 *Phys. Lett.* B **810** 135814
[23] Yoshida S and Shimizu N 2021 *Progress of Theoretical and Experimental Physics* **2022** 053D02
[24] Furnstahl R J, Garcia A J, Millican P J and Zhang X 2020 *Phys. Lett.* B **809** 135719
[25] Melendez J A, Drischler C, Garcia A J, Furnstahl R J and Zhang X 2021 *Phys. Lett.* B **821** 136608
[26] Drischler C, Quinonez M, Giuliani P G, Lovell A E and Nunes F M 2021 *Phys. Lett.* B **823** 136777
[27] Zhang X and Furnstahl R J 2022 *Phys. Rev.* C **105** 064004
[28] Benner P, Gugercin S and Willcox K 2015 *SIAM Rev.* **57** 483
[29] Benner P, Cohen A, Ohlberger M and Willcox K 2017 *Model Reduction and Approximation* (Philadelphia, PA: SIAM)
[30] Benner P, Schilders W, Grivet-Talocia S, Quarteroni A, Rozza G and Miguel Silveira L 2020 *Model Order Reduction: Volume 2: Snapshot-Based Methods and Algorithms* (Berlin: de Gruyter & Co) pp 1–348
[31] Buchan A G, Pain C C, Fang F and Navon I M 2013 *Int. J. Numer. Methods Eng.* **95** 1011
[32] Quarteroni A, Manzoni A and Negri F 2016 *Reduced Basis Methods for Partial Differential Equations. An Introduction* (*La Matematica per il 3+2*) (Berlin: Springer)
[33] Zienkiewicz O, Taylor R and Zhu J 2013 *The Finite Element Method: Its Basis and Fundamentals* 7th edn (Oxford: Butterworth-Heinemann)
[34] Babuška I and Osborn J 1991 *Finite Element Methods (Part 1): Handbook of Numerical Analysis* vol 2 (Amsterdam: Elsevier) pp 641–787
[35] Leissa A W 2005 *J. Sound Vib.* **287** 961
[36] Ilanko S 2009 *J. Sound Vib.* **319** 731
[37] Parlett B 1980 *The Symmetric Eigenvalue Problem* (*Classics in Applied Mathematics*) (Philadelphia, PA: SIAM)
[38] Jia Z and Stewart G W 2001 *Math. Comput.* **70** 637
[39] Beattie C 2000 *Math. Comp.* **69** 1409
[40] Machiels L, Maday Y, Oliveira I B, Patera A T and Rovas D V 2000 *C. R. Acad. Sci., Paris* I **331** 153
[41] Horger T, Wohlmuth B and Dickopf T 2017 *ESAIM Math. Model. Numer. Anal.* **51** 443
[42] Pau G S H 2007 *Phys. Rev.* E **76** 046704
[43] Cheng M-C 2016 *AIP Adv.* **6** 095121
[44] Gräbner N, Mehrmann V, Quraishi S, Schröder C and von Wagner U 2016 *Z. Angew. Math. Mech.* **96** 1388
[45] Huynh D B P, Rozza G, Sen S and Patera A T 2007 *C. R. Math.* **345** 473
[46] Rozza G, Huynh D B P and Patera A T 2008 *Arch. Comput. Methods Eng.* **15** 229

[47] Chen P, Quarteroni A and Rozza G 2017 *SIAM/ASA J. Uncertain. Quantification* **5** 813
[48] Haasdonk B 2017 Chapter 2: reduced basis methods for parametrized PDEs—a tutorial introduction for stationary and instationary problems *Model Reduction and Approximation* (Philadelphia, PA: SIAM) pp 65–136
[49] Hess M W, Grundel S and Benner P 2015 *IEEE Trans. Microw. Theory Tech.* **63** 3549
[50] Ritz W 1909 *J. fur Reine Angew. Math.* **1909** 1
[51] Ritz W 1909 *Ann. Phys.* **333** 737
[52] Mikhlin S G 1967 *Variational Methods in Mathematical Physics* (Cambridge: Cambridge University Press) (Translated by T Boddington for Pergamon Press, 1964)
[53] Blanchard P, Brüning E and Hayes G 1992 *Variational Methods in Mathematical Physics: A Unified Approach* (*Texts and Monographs in Physics*) (Berlin: Springer)
[54] Jackson J D 1999 *Classical Electrodynamics* 3rd edn (New York: Wiley)
[55] Engel E and Dreizler R 2011 *Density Functional Theory: An Advanced Course* (*Theoretical and Mathematical Physics*) (Berlin: Springer)
[56] Schunck N (ed) 2019 *Energy Density Functional Methods for Atomic Nuclei* (Bristol: IOP Publishing) pp 2053–563
[57] Cancès E, LeBris C, Maday Y and Turinici G 2002 *J. Sci. Comput.* **17** 461
[58] Cancès É, Le Bris C, Maday Y, Nguyen N, Patera A and Pau G 2007 *High-Dimensional Partial Differential Equations in Science and Engineering* vol 41 (Providence, RI: American Mathematical Society) p 15
[59] Lin L, Lu J, Ying L and E W 2012 *J. Comput. Phys.* **231** 2140
[60] Zhang G, Lin L, Hu W, Yang C and Pask J E 2017 *J. Comput. Phys.* **335** 426
[61] Taylor J R 2006 *Scattering Theory: The Quantum Theory of Nonrelativistic Collisions* (New York: Dover)
[62] Drischler C and Zhang X 2022 ed , , Nuclear forces for precision nuclear physics - a collection of perspectives arXiv 2201.01105 [nucl-th])
[63] Newton R G 2002 *Scattering Theory of Waves and Particles* (New York: Dover)
[64] Melendez J 2021 Quantum emulator examples *GitHub repository* (http://dx.doi.org/10.5281/zenodo.6332487)
[65] Melendez J 2021 Reduced-order DFT emulators for nuclear physics *GitHub repository* (http://dx.doi.org/10.5281/zenodo.6332487)
[66] Zienkiewicz O, Taylor R and Fox D 2014 *The Finite Element Method for Solid and Structural Mechanics* 7th edn (Oxford: Butterworth-Heinemann)
[67] Zienkiewicz O, Taylor R and Nithiarasu P 2014 *The Finite Element Method for Fluid Dynamics* 7th edn (Oxford: Butterworth-Heinemann)
[68] Hesthaven J, Rozza G and Stamm B 2015 *Certified Reduced Basis Methods for Parametrized Partial Differential Equations* (*SpringerBriefs in Mathematics*) (Berlin: Springer)
[69] Evans J D 1996 *Straightforward Statistics for the Behavioral Sciences* (Pacific Grove, CA: Brooks/Cole Publishing)
[70] Brenner S C and Scott L R 2008 *The Mathematical Theory of Finite Element Methods* (*Texts in Applied Mathematics* vol 15) (Berlin: Springer)
[71] Bruns A and Benner P 2015 *Math. Comput. Model. Dyn. Syst.* **21** 103
[72] Lassila T, Manzoni A, Quarteroni A and Rozza G 2013 *Int. J. Numer. Methods Biomed. Eng.* **29** 741
[73] Lieberman C, Fidkowski K, Willcox K and van Bloemen Waanders B 2013 *Int. J. Numer. Methods Fluids* **71** 135
[74] Benner P, Schilders W, Grivet-Talocia S, Quarteroni A, Rozza G and Miguel Silveira L 2021 *System- and Data-Driven Methods and Algorithms* (Berlin: de Gruyter & Co)
[75] Gubisch M and Volkwein S 2017 Chapter 1: proper orthogonal decomposition for linear-quadratic optimal control *Model Reduction and Approximation* (Philadelphia, PA: SIAM) pp 3–63
[76] Sarkar A and Lee D 2022 Self-learning emulators and eigenvector continuation *Phys. Rev. Research* **4** 023214
[77] Baur U, Beattie C, Benner P and Gugercin S 2011 *SIAM J. Sci. Comput.* **33** 2489
[78] Baur U, Benner P, Haasdonk B, Himpe C, Martini I and Ohlberger M 2017 Chapter 9: comparison of methods for parametric model order reduction of time-dependent problems *Model Reduction and Approximation* (Philadelphia, PA: SIAM) ch 9 pp 377–407
[79] Barrault M, Maday Y, Nguyen N C and Patera A T 2004 *C. R. Math.* **339** 667
[80] Grepl M A, Maday Y, Nguyen N C and Patera A T 2007 *ESAIM Math. Model. Numer. Anal.* **41** 575

[81] Chaturantabut S and Sorensen D C 2009 *Proc. 48h IEEE Conf. Decision and Control (CDC) Held Jointly with 2009 28th Chinese Control Conf.* pp 4316–21
[82] Chaturantabut S and Sorensen D C 2010 *SIAM J. Sci. Comput.* **32** 2737
[83] Bui-Thanh T, Damodaran M and Willcox K 2003 Proper orthogonal decomposition extensions for parametric applications in compressible aerodynamics *21st AIAA Applied Aerodynamics Conf.*
[84] Carlberg K, Bou-Mosleh C and Farhat C 2011 *Int. J. Numer. Methods Eng.* **86** 155
[85] Amsallem D, Cortial J and Farhat C 2010 *AIAA J.* **48** 2029
[86] An S S, Kim T and James D L 2008 *ACM Trans. Graph.* **27** 165
[87] Farhat C, Avery P, Chapman T and Cortial J 2014 *Int. J. Numer. Methods Eng.* **98** 625
[88] Guo M and Hesthaven J S 2018 *Comput. Methods Appl. Mech. Eng.* **341** 807
[89] Noor A K and Peters J M 1980 *AIAA J.* **18** 455
[90] Milk R, Rave S and Schindler F 2016 *SIAM J. Sci. Comput.* **38** S194
[91] Benner P and Werner S W R 2021 *Model Reduction of Complex Dynamical Systems* (*International Series of Numerical Mathematics* vol 171) ed P Benner, T Breiten, H Faßbender, M Hinze, T Stykel and R Zimmermann (Basel: Birkhäuser) pp 393–415
[92] Belson B A, Tu J H and Rowley C W 2014 *ACM Trans. Math. Softw.* **40** 30
[93] Puzyrev V, Ghommem M and Meka S 2019 *J. Comput. Sci.* **30** 157
[94] Rizzi F, Blonigan P J, Parish E J and Carlberg K T 2021 arXiv:2003.07798 [cs.MS]
[95] Bonilla E, Giuliani P, Godbey K and Lee D 2022 arXiv:2203.05284 [nucl-th]
[96] BAND Collaboration 2020 Bayesian analysis of nuclear dynamics (BAND) framework project https://bandframework.github.io/