10 Nov 2020

New Proper Orthogonal Decomposition Approximation Theory for PDE Solution Data

Sarah Locke

John R. Singler
Missouri University of Science and Technology, singlerj@mst.edu

Recommended Citation

NEW PROPER ORTHOGONAL DECOMPOSITION APPROXIMATION THEORY FOR PDE SOLUTION DATA*

SARAH LOCKE† AND JOHN SINGLER†

Abstract. In our previous work [J. R. Singler, *SIAM J. Numer. Anal.*, 52 (2014), pp. 852–876], we considered the proper orthogonal decomposition (POD) of time varying PDE solution data taking values in two different Hilbert spaces. We considered various POD projections of the data and obtained new results concerning POD projection errors and error bounds for POD reduced order models of PDEs. In this work, we improve on our earlier results concerning POD projections by extending to a more general framework that allows for nonorthogonal POD projections and seminorms. We obtain new exact error formulas and convergence results for POD data approximation errors, and also prove new pointwise convergence results and error bounds for POD projections. We consider both the discrete and continuous cases of POD. We also apply our results to several example problems and show how the new results improve on previous work.

1. Introduction. Proper orthogonal decomposition (POD) is a model order reduction technique for partial differential equations (PDEs) and other mathematical models. With this method, modes are computed from simulation or experimental data and a Galerkin projection is used with these modes to reduce the model. Because POD reduced order models often have very low dimension, they can be used to efficiently simulate computationally demanding problems. Therefore, POD has been used in many fields of study including fluid dynamics and control theory. For a small selection of applications, see [4, 21, 33, 36, 45]. For more information about POD and many known results, see, e.g., [14, 17, 30].

Because of the wide use of POD in many application areas, it is of great interest to study the approximation errors in POD model order reduction procedures. Numerical analysis results for POD reduced order models of PDEs were first obtained by Kunisch and Volkwein [26, 27] and then by many others; see, e.g., [1, 5, 9, 10, 13, 15, 16, 18, 19, 20, 21, 23, 24, 29, 32, 37, 38, 40, 44, 46, 47, 48] and the references therein.

Understanding POD data approximation errors is typically important for these numerical analysis works. To see this, let w be the solution of the mathematical model, let w_r be the solution of the POD reduced order model, and let π_r be a projection onto the span of the first r POD modes. Split the error as

$$w - w_r = \rho_r + \theta_r, \quad \rho_r = w - \pi_r w, \quad \theta_r = \pi_r w - w_r.$$

Energy estimates can often be used to bound θ_r by quantities including various norms of ρ_r, the POD data approximation error for that projection.

In our previous work [40], exact error formulas and convergence results were proven for norms of ρ_r involving two Hilbert spaces, where one space is a subset of the other. In that work, we considered the continuous POD setting and proved

---

†Department of Mathematics and Statistics, Missouri University of Science and Technology, Rolla, MO 65409 USA (sld77@mst.edu, singlerj@mst.edu).

results for different combinations of POD spaces, projections, and norms. Shortly after [40], Iliescu and Wang [20] provided analogous error formulas for the discrete POD case, and many of the recent numerical analysis works mentioned above use results from [20, 40] or extensions of these results to other scenarios.

As POD is increasingly applied in a variety of situations, it becomes more useful to have error results that can be easily applied in a wide range of scenarios. Therefore, in this work we extend POD data approximation results in [20, 40] to a generalized framework that allows us to treat nonorthogonal POD projections and seminorms. We prove new error formulas and convergence results for norms of quantities involving $\rho_r = w - \pi_r w$ with various POD projections π_r. We also prove new pointwise convergence results for different POD projections. Nonorthogonal POD projections have been used in the numerical analysis for POD reduced order models [18, 37]; however, the exact POD data approximation error formulas and convergence results obtained here are new. Exact POD data approximation errors using various seminorms have been obtained in some cases (see, e.g., [10, section 3.3], [38, Lemma 3.1]); the general extension and convergence results in this work are new. Finally, some pointwise convergence results for POD projections were obtained in our earlier work [40]; we obtain new error bounds and improved convergence results here.

The POD data approximation error formulas presented in this work are exact and do not require the use of POD inverse inequalities. We consider both the discrete and continuous cases for POD and generalize the setting in [20, 40] to allow a linear mapping between two Hilbert spaces to act on the data. We require minimal assumptions on the data, the linear operator, and the Hilbert spaces; the assumptions we do require are naturally satisfied in many applications and also allow us to obtain convergence results even in the fully continuous case when the data has infinitely many positive POD eigenvalues. We note that most of the proof strategies in this work are new; some proofs do rely on techniques from [40, 41].

The rest of the paper is outlined as follows. Section 2 provides both a brief general background and POD specific background for both the discrete and continuous cases. Then section 3 provides an overview of the new results along with the notation and main assumptions needed. Properties of POD and POD projections are given in section 4. Error formulas are presented in section 5 and pointwise convergence results are given in section 6. Finally, in section 7, we consider examples and compare the results from previous work and the current work.

2. Background. In this section, we recall some functional analysis background material, and also the basic theory for discrete POD and continuous POD. For details and proofs for the basic discrete and continuous POD theory, see, e.g., [7, 14, 17, 27, 34, 43] and also Appendix A.

2.1. Functional analysis background. Let V and W be Hilbert spaces with inner products[1] $(\cdot, \cdot)_V$ and $(\cdot, \cdot)_W$ and corresponding norms $\|\cdot\|_V$ and $\|\cdot\|_W$. Throughout this work, the scalar field $\mathbb{K}$ for all spaces is either $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

Linear operators. Let $T : V \to W$ be a linear operator with domain $\mathcal{D}(T) \subset V$, range $\mathcal{R}(T) \subset W$, and null space $\ker(T) \subset V$. The *rank* of T is the dimension of $\mathcal{R}(T)$. The operator T is *bounded* if $\|Tv\|_W \leq M\|v\|_V$ for all $v \in \mathcal{D}(T)$. Throughout this paper, we only consider bounded operators $T : V \to W$ that are defined on the whole space, so $\mathcal{D}(T) = V$. For such a bounded operator $T : V \to W$, the usual

---

[1] In this paper, all inner products and sesquilinear forms are linear in the first argument and conjugate linear in the second argument.

operator norm is given by $\|T\| = \sup\{\|Tv\|_W : v \in V, \|v\|_V = 1\}$. We also consider unbounded linear operators that are not defined everywhere, so that $\mathcal{D}(T) \neq V$. The operator T is *closed* if its graph, $\mathcal{G}(T) = \{(v, w) : v \in \mathcal{D}(T), w = Tv\}$, is closed in $V \times W$. If T is bounded (and everywhere defined), then T is closed. If T is closed and invertible, then T^{-1} is closed.

Adjoint operators. The *Hilbert-adjoint operator* $T^* : W \to V$ of $T : V \to W$ satisfies $(Tv, w)_W = (v, T^*w)_V$ for all $v \in \mathcal{D}(T)$ and $w \in \mathcal{D}(T^*)$. If T is bounded, then T^* exists, is unique, and is also bounded. If T is densely defined, then T^* exists, is unique, and is closed; in addition, if T is closed, then T^* is densely defined. If $T : V \to W$ is invertible, then we let $T^{-*} : V \to W$ denote the Hilbert adjoint operator of the inverse $T^{-1} : W \to V$. We note for T^* to exist we need T bounded or densely defined, and for T^{-*} to exist we need T^{-1} bounded or densely defined. We note these assumptions when necessary.

The following basic result is important in this work.

LEMMA 2.1. *Let V and W be Hilbert spaces. If $T : V \to W$ is a bounded linear operator, then $\ker(TT^*) = \ker(T^*)$ and $\ker(T^*T) = \ker(T)$.*

Proof. We only prove the first one. Let $w \in \ker(TT^*)$. Then,

$$TT^*w = 0 \Rightarrow (TT^*w, w)_W = 0 \Rightarrow (T^*w, T^*w)_V = 0 \Rightarrow \|T^*w\|_V^2 = 0 \Rightarrow T^*w = 0.$$

Next, let $w \in \ker(T^*)$. Then $T^*w = 0 \Rightarrow TT^*w = 0$. $\square$

Projections. A bounded linear operator $\Pi : V \to V$ is a *projection* onto $U = \mathcal{R}(\Pi)$ if $\Pi^2 = \Pi$. Then we have $\Pi v \in U$ for all $v \in V$ and $\Pi u = u$ for all $u \in U$. Also, Π is an *orthogonal projection* if $u = \Pi v \in U$ minimizes $\inf_{u \in U} \|v - u\|_V$ for any $v \in V$. A nontrivial orthogonal projection Π is automatically self-adjoint, i.e., $\Pi^* = \Pi$, and satisfies $\|\Pi\| = 1$. We consider nonorthogonal projections in this work, and therefore we do not assume a projection is orthogonal or self-adjoint unless explicitly specified. Sometimes, we assume a family of projections $\{\Pi_r\}$ is uniformly bounded in operator norm, i.e., there exists a constant C such that $\|\Pi_r\| \leq C$ for all r.

The singular value decomposition of a compact operator. If $T : V \to W$ is a compact linear operator, with separable Hilbert spaces V and W, then T has a *singular value decomposition* (SVD). The positive singular values of T are defined to be the square roots of the positive eigenvalues of the self-adjoint nonnegative compact operators $TT^* : W \to W$ and $T^*T : V \to V$. Further, the nonzero eigenvalues of these operators are equal, and we consider zero a singular value of T if either operator has a zero eigenvalue. If the ordered singular values of T are given by $\mu_1 \geq \mu_2 \geq \cdots \geq 0$ (including repetitions), the orthonormal basis of eigenvectors of TT^* is given by $\{\psi_k\} \subset W$, and the orthonormal basis of eigenvectors of T^*T is given by $\{g_k\} \subset V$, then the SVD of T is the expansion given by

$$Tg = \sum_{k \geq 1} \mu_k (g, g_k)_V \psi_k$$

for all $g \in V$. If $\mu_k > 0$, then

$$Tg_k = \mu_k \psi_k \text{ and } T^*\psi_k = \mu_k g_k.$$

Also, the rank r truncated SVD $T_r : V \to W$ of T is defined for $g \in V$ by

$$T_r g := \sum_{k=1}^{r} \mu_k (g, g_k)_V \psi_k.$$

For more information, see, e.g., [11, Chapters VI–VIII], [22, section V.2.3], [28, Chapter 30], [35, sections VI.5–VI.6].

Hilbert–Schmidt operators. Let $T : V \to W$ be a linear operator, with separable Hilbert spaces V and W, and let $\{g_k\}$ be any orthonormal basis for V. Define the *Hilbert–Schmidt norm* of T as

$$(2.1) \qquad \|T\|_{\mathrm{HS}(V,W)} = \left(\sum_{k \geq 1} \|Tg_k\|_V^2 \right)^{1/2}.$$

If the sum converges we say the operator T is *Hilbert–Schmidt*. The Hilbert–Schmidt norm is independent of the choice of orthonormal basis, every Hilbert–Schmidt operator is compact, $\|T\| \leq \|T\|_{\mathrm{HS}(V,W)}$, T is Hilbert–Schmidt if and only if T^* is Hilbert–Schmidt, and T is Hilbert–Schmidt if and only if $\sum_{k \geq 1} \sigma_k^2 < \infty$, where $\{\sigma_k\}$ are the singular values (including repetitions) of T. We also have

$$\|T\|_{\mathrm{HS}(V,W)}^2 = \|T^*\|_{\mathrm{HS}(W,V)}^2 = \sum_{k \geq 1} \sigma_k^2.$$

For more, see, e.g., [11, Chapter VIII], [22, section V.2.4], [35, section VI.6].

Bochner spaces. Let $\mathcal{O}$ be an open subset of $\mathbb{R}^d$ for some $d \geq 1$. For $p \in [1, \infty)$, let $L^p(\mathcal{O}; V)$ denote the Bochner space of (equivalence classes of) Lebesgue measurable functions $v : \mathcal{O} \to V$ satisfying $\int_{\mathcal{O}} \|v(t)\|_V^p \, dt < \infty$. For $p = 2$, $L^2(\mathcal{O}; V)$ is a Hilbert space with inner product

$$(v, w)_{L^2(\mathcal{O};V)} = \int_{\mathcal{O}} (v(t), w(t))_V \, dt.$$

The following theorem (see, e.g., [8, Theorem III.6.20] and [31, Theorem 4.2.10]) allows us to bring a closed linear operator inside an integral.

THEOREM 2.2. *Suppose $T : \mathcal{D}(T) \subset V \to W$ is a closed linear operator. If $v : \mathcal{O} \to \mathcal{D}(T)$, $v \in L^1(\mathcal{O}; V)$, and $Tv \in L^1(\mathcal{O}; W)$, then*

$$\int_{\mathcal{O}} v(t) \, dt \in \mathcal{D}(T) \quad and \quad T \int_{\mathcal{O}} v(t) \, dt = \int_{\mathcal{O}} Tv(t) \, dt.$$

2.2. Discrete POD. Let X be a separable Hilbert space. For the discrete case, let s be a positive integer and assume the POD data is given by $\{w_j\}_{j=1}^s \subset X$. Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, and define $S := \mathbb{K}_\Gamma^s$ with the weighted inner product given by

$$(u, v)_S = v^* \Gamma u = \sum_{j=1}^s \gamma_j u_j \overline{v_j},$$

where $u, v \in S$, $\Gamma = \mathrm{diag}(\gamma_1, \gamma_2, \ldots, \gamma_s)$, and the values $\{\gamma_j\}_{j=1}^s$ are positive weights. Note that these weights commonly arise from integral approximations. Define the POD operator $K : S \to X$ by

$$(2.2) \qquad Kf = \sum_{j=1}^s \gamma_j f_j \, w_j, \quad f = [f_1, f_2, \ldots, f_s]^T.$$

Since K has finite dimensional range, it is a compact operator and has a SVD. Let $\{\sigma_k, f_k, \varphi_k\} \subset \mathbb{R} \times S \times X$ be the singular values and orthonormal singular vectors ordered so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$. Thus, the SVD is given by

$$(2.3) \qquad Kf = \sum_{j \geq 1} \sigma_j (f, f_j)_S \varphi_j.$$

When $\sigma_k > 0$, we have

$$Kf_k = \sigma_k \varphi_k \quad \text{and} \quad K^* \varphi_k = \sigma_k f_k,$$

where $K^* : X \to S$ is the Hilbert adjoint operator given by

$$K^* x = [(x, w_1)_X, (x, w_2)_X, \ldots, (x, w_s)_X]^T.$$

For a positive integer r, define $X_r = \text{span}\{\varphi_k\}_{k=1}^r$. Let $\Pi_r^X : X \to X$ be the orthogonal projection onto X_r, i.e., for $x \in X$ fixed, $\Pi_r^X x \in X_r$ minimizes the approximation error $\|x - x_r\|_X$ over all choices of $x_r \in X_r$. Since $\{\varphi_k\}$ is an orthonormal set in X, we have the exact representation

$$(2.4) \qquad \Pi_r^X x = \sum_{k=1}^r (x, \varphi_k)_X \varphi_k.$$

The singular vectors $\{\varphi_k\}$ are called the POD modes of the data $\{w_k\} \subset X$. The POD modes provide the best low rank approximation to the data in the following sense: we have

$$(2.5) \qquad \sum_{k=1}^s \gamma_k \left\| w_k - \Pi_r^X w_k \right\|_X^2 = \sum_{k>r} \sigma_k^2,$$

and no other choice of an orthonormal basis in (2.4) gives a smaller value for the approximation error.

DEFINITION 2.3. *We call the singular values $\{\sigma_k\}$ and singular vectors $\{\varphi_k\} \subset X$ of K the POD singular values and POD modes for the data $\{w_j\}_{j=1}^s$, respectively. We also call the eigenvalues $\{\lambda_k\}$ of the operator $KK^* : X \to X$ the POD eigenvalues for the data $\{w_j\}_{j=1}^s$. We let s_X denote the number of positive POD singular values (or positive POD eigenvalues) for the data $\{w_j\}_{j=1}^s$, i.e., $s_X = \text{rank}(K)$.*

From subsection 2.1, we know $\lambda_k = \sigma_k^2$ whenever $\lambda_k > 0$. Also, we have $s_X \leq s < \infty$. It is possible for data to have a zero POD singular value but have all positive POD eigenvalues; this can happen if $s > \dim(X)$.

2.3. Continuous POD. Similarly to the discrete case we define the POD operator $K : S \to X$ for the continuous case, where again X is a separable Hilbert space. Let d and m be positive integers and let $\mathcal{O} \subset \mathbb{R}^d$ be an open set. Then define $S := L^2(\mathcal{O}; \mathbb{K}^m)$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. We note that $L^2(\mathcal{O})$ is separable (see, e.g., [6, Theorem 2.5-4]), and therefore so is S. Assume the POD data is given by $\{w_j\}_{j=1}^m \subset L^2(\mathcal{O}; X)$.

Remark 2.4. In POD applications the set $\mathcal{O}$ is frequently a time interval; however, researchers also take $\mathcal{O}$ to be a multidimensional parameter domain as well. Note that we could also consider multiple open sets, $\mathcal{O}_j \subset \mathbb{R}^{d_j}$, and data $w_j \in L^2(\mathcal{O}_j; X)$ for $j = 1, \ldots, m$. In this case, we would define $S := L^2(\mathcal{O}_1) \times \cdots \times L^2(\mathcal{O}_m)$. All results in this paper hold for this case as well. The previous case is chosen to simplify notation.

Define the POD operator $K : S \to X$ by

$$(2.6) \qquad Kf = \sum_{j=1}^{m} \int_{\mathcal{O}} f_j(t) w_j(t) dt, \quad f \in S.$$

Since $f \in S$, note that $f = [f_1, f_2, \ldots, f_m]^T$, where each $f_j \in L^2(\mathcal{O})$. As in the discrete case, we know that K is a compact operator and has an SVD. We let $\{\sigma_k, f_k, \varphi_k\} \subset \mathbb{R} \times S \times X$ denote the singular values and orthonormal singular vectors ordered so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$. The SVD of K is given as in the discrete case (2.3). Thus, when $\sigma_k > 0$, we have

$$Kf_k = \sigma_k \varphi_k \quad \text{and} \quad K^* \varphi_k = \sigma_k f_k,$$

where $K^* : X \to S$ is the Hilbert adjoint operator defined by

$$[K^* x](t) = [(x, w_1(t))_X, (x, w_2(t))_X, \ldots, (x, w_m(t))_X]^T.$$

We define $X_r := \mathrm{span}\{\varphi_k\}_{k=1}^{r}$ and the orthogonal projection $\Pi_r^X : X \to X$ (2.4) as before. The data approximation error is given by

$$(2.7) \qquad \sum_{j=1}^{m} \int_{\mathcal{O}} \left\| w_j(t) - \Pi_r^X w_j(t) \right\|_X^2 dt = \sum_{k>r} \sigma_k^2,$$

and the error goes to zero as $r \to \infty$. As in the discrete case, no other orthonormal basis in (2.4) gives a smaller value for the error.

We define the POD singular values, POD modes, POD eigenvalues, and $s_X = \mathrm{rank}(K)$ as in Definition 2.3 for the discrete case. Again, it is possible for data to have a zero POD singular value but have all positive POD eigenvalues; an example where X is infinite dimensional can be found in [41, section 3.1, Example 3]. Also, if X is finite dimensional, then the data always has a zero POD singular value.

3. Main assumptions, notation, and new results. In this section we highlight the notation used in each case as well as the main assumptions made throughout the paper. Furthermore, we briefly present an overview of the new results and give an example to illustrate how the new results can be used.

Throughout the remainder of this paper, assume X and Y are separable Hilbert spaces and $L : \mathcal{D}(L) \subset X \to Y$ is a linear operator. We study POD error formulas and POD projections involving the data $\{w_j\}$ and the data $\{Lw_j\}$.

3.1. Discrete case: Assumptions and notation. As in subsection 2.2, we consider data $\{w_j\}_{j=1}^{s} \subset X$ and the corresponding POD operator $K : S \to X$ defined by $Kf = \sum_{j=1}^{s} \gamma_j f_j w_j$, where $S = \mathbb{K}_\Gamma^s$ and $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$. The SVD of K is given by $Kf = \sum_{k \geq 1} \sigma_k (f, f_k)_S \varphi_k$. The set X_r is the span of $\{\varphi_k\}_{k=1}^{r}$, and $\Pi_r^X : X \to X$ is the orthogonal projection onto X_r.

To consider POD projections involving the data $\{Lw_j\}$, we make the following assumption:

> **Main assumption.** For the discrete case, we assume throughout the paper that (i) $\{w_j\}_{j=1}^{s} \subset \mathcal{D}(L)$, and also (ii) $\sigma_r > 0$ whenever we consider the projection Π_r^X.

Assumption (i) has two important consequences. First, since $w_j \in \mathcal{D}(L)$ for each j, we know the range of K is contained in $\mathcal{D}(L)$. Second, assumption (i) allows us to consider the POD operator $K^Y : S \to Y$ for the data $\{Lw_j\}_{j=1}^s \subset Y$ defined by

$$(3.1) \qquad K^Y f = LKf = \sum_{j=1}^{s} \gamma_j f_j L w_j, \quad f = [f_1, f_2, \ldots, f_s]^T.$$

Note that K^Y is the result of applying L to the POD operator K for the data $\{w_j\}$, i.e., $K^Y = LK$. Since K^Y has finite rank, it is compact and has an SVD. Define $s_Y = \mathrm{rank}(K^Y)$ to be the number of positive singular values of K^Y. Note that assumption (i) is automatically satisfied if L is bounded.

For assumption (ii), note that if $\sigma_k > 0$, then assumption (i) implies the corresponding singular vector φ_k is in $\mathcal{D}(L)$ since

$$(3.2) \qquad \varphi_k = \sigma_k^{-1} K f_k \in \mathcal{D}(L).$$

Since $\sigma_r > 0$, this implies $X_r \subset \mathcal{D}(L)$ and Π_r^X maps into $\mathcal{D}(L)$.

To guarantee the boundedness of certain POD projections, in some cases of Theorem 4.3 we need to assume the POD modes $\{\varphi_k\}_{k=1}^r \subset \mathcal{D}(L)$ satisfy some additional regularity properties. These properties can be guaranteed by making additional regularity assumptions on the data.

First, the condition $\{\varphi_k\}_{k=1}^r \subset \mathcal{D}(L^{-*})$ is guaranteed to hold if we assume $\sigma_r > 0$ and $w_j \in \mathcal{D}(L^{-*})$ for each j. With this assumption, we know as above that $\mathcal{R}(K) \subset \mathcal{D}(L^{-*})$ and also $\varphi_k \in \mathcal{D}(L^{-*})$ whenever $\sigma_k > 0$. Since $\sigma_r > 0$, we can guarantee $\{\varphi_k\}_{k=1}^r \subset \mathcal{D}(L^{-*})$.

Next, a similar argument using (3.2) shows the condition $\{L\varphi_k\}_{k=1}^r \subset \mathcal{D}(L^*)$ is guaranteed to hold if we assume $\sigma_r > 0$ and $Lw_j \in \mathcal{D}(L^*)$ for each j.

3.2. Continuous case: Assumptions and notation. The continuous case requires a few more assumptions. Recall $K : S \to X$, where $S := L^2(\mathcal{O}; \mathbb{K}^m)$ and $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$. In order to define the POD operator K^Y and ensure $\{\varphi_k\}_{k=1}^r \subset \mathcal{D}(L)$, we make the following assumption:

Main assumption. For the continuous case, we assume throughout the paper that (i) $\{Lw_j\}_{j=1}^m \subset L^2(\mathcal{O}; Y)$, and for all $f \in S$ we have $Kf \in \mathcal{D}(L)$ and

$$LKf = \sum_{j=1}^{m} \int_{\mathcal{O}} f_j(t) L w_j(t) dt,$$

and also (ii) $\sigma_r > 0$ whenever we consider the projection Π_r^X.

As in the discrete case, assumption (i) gives $\mathcal{R}(K) \subset \mathcal{D}(L)$ and allows us to define the (compact) POD operator $K^Y = LK$ for the data $\{Lw_j\}_{j=1}^m \subset L^2(\mathcal{O}; Y)$. As before, we let $s_Y = \mathrm{rank}(K^Y)$ be the number of positive singular values of K^Y. Also as in the discrete case, assumptions (i) and (ii) imply $\{\varphi_k\}_{k=1}^r \subset \mathcal{D}(L)$ and Π_r^X maps into $\mathcal{D}(L)$.

Remark 3.1. There are three common conditions that guarantee assumption (i) holds.

1. If $L : X \to Y$ is bounded, the operator L can be pulled through the integral in the definition of K and assumption (i) clearly holds.

2. If each $w_j \in L^2(\mathcal{O}; X)$ takes the form

$$w_j(t) = \sum_{\ell,k=1}^{n_j} a_{jk\ell} g_{kj}(t) x_{\ell j},$$

where $a_{jk\ell}$ are constants in $\mathbb{K}$, $g_{kj}(t) \in L^2(\mathcal{O})$, and $x_{\ell j} \in \mathcal{D}(L)$, then it can be checked that assumption (i) holds. This condition is similar to the assumption made in the discrete case.

3. If $L : \mathcal{D}(L) \subset X \to Y$ is closed, $w_j \in \mathcal{D}(L)$ a.e., and $Lw_j \in L^2(\mathcal{O}; Y)$, then Theorem 2.2 implies assumption (i) holds.

Again, for certain cases of Theorem 4.3 we need to assume the POD modes $\{\varphi_k\}_{k=1}^r \subset \mathcal{D}(L)$ satisfy some additional regularity properties. As in the discrete case, we can make additional assumptions on the data to satisfy these regularity properties.

We briefly mention conditions on the data similar to Remark 3.1, item 3, that yield the needed regularity. First, if L^{-*} exists, it is closed. Therefore, $\{\varphi_k\}_{k=1}^r \subset \mathcal{D}(L^{-*})$ holds if we assume $\sigma_r > 0$, $w_j \in \mathcal{D}(L^{-*})$ a.e., and $\{L^{-*}w_j\}_{j=1}^m \in L^2(\mathcal{O}; Y)$. Second, if L^* exists, then it is closed. Therefore, $\{L\varphi_k\}_{k=1}^r \subset \mathcal{D}(L^*)$ holds if we assume $\sigma_r > 0$, $Lw_j \in \mathcal{D}(L^*)$ a.e., and $\{L^*Lw_j\}_{j=1}^m \in L^2(\mathcal{O}; X)$.

We also note that the condition in Remark 3.1, item 2, can be modified similarly to the discrete case to yield the required regularity.

3.3. An overview of the new results. Here we give an overview of the new results presented in this paper. For the overview we focus on the continuous case, but there are analogous results for the discrete case.

Recall the standard POD orthogonal projection, $\Pi_r^X : X \to X$ given by (2.4), and the known POD data approximation error given by

$$\sum_{j=1}^m \int_{\mathcal{O}} \left\| w_j(t) - \Pi_r^X w_j(t) \right\|_X^2 dt = \sum_{k>r} \sigma_k^2.$$

One of the goals of this paper is to find extensions of this error formula to other scenarios involving the linear operator $L : X \to Y$ and another sequence of projections, which need not be orthogonal.

DEFINITION 3.2. *For a positive integer r with $\sigma_r > 0$, we define $Y_r := LX_r = \text{span}\{L\varphi_k\}_{k=1}^r$ and we let $\Pi_r^Y : Y \to Y$ be a projection onto Y_r.*

Remark 3.3. First, the condition $\sigma_r > 0$ implies $X_r \subset \mathcal{D}(L)$ and so the definition makes sense. We assume throughout that $\sigma_r > 0$ whenever we consider Π_r^Y. Next, it is important to note that unless stated otherwise we do not assume the projection Π_r^Y is orthogonal. To obtain convergence results as r increases, we sometimes need to require $\{\Pi_r^Y\}$ are uniformly bounded in operator norm. If $\{\Pi_r^Y\}$ are the orthogonal projections onto Y_r, then this condition is satisfied.

Under the main assumption we have the data approximation errors

$$(3.3) \qquad \sum_{j=1}^m \int_{\mathcal{O}} \left\| Lw_j(t) - L\Pi_r^X w_j(t) \right\|_Y^2 dt = \sum_{k>r} \sigma_k^2 \|L\varphi_k\|_Y^2$$

and

$$(3.4) \qquad \sum_{j=1}^m \int_{\mathcal{O}} \left\| Lw_j(t) - \Pi_r^Y Lw_j(t) \right\|_Y^2 dt = \sum_{k>r} \sigma_k^2 \left\| L\varphi_k - \Pi_r^Y L\varphi_k \right\|_Y^2.$$

The error in (3.3) converges to zero as $r \to \infty$, and the error in (3.4) tends to zero as r increases when the projections $\{\Pi_r^Y\}$ are uniformly bounded. Also, under a basic condition on L^{-1}, we have the data approximation error

$$
(3.5) \qquad \sum_{j=1}^m \int_{\mathcal{O}} \left\| w_j(t) - L^{-1}\Pi_r^Y L w_j(t) \right\|_X^2 dt = \sum_{k>r} \sigma_k^2 \left\| \varphi_k - L^{-1}\Pi_r^Y L\varphi_k \right\|_X^2 .
$$

There are further conditions implying the error in (3.5) converges to zero as well. The details for the assumptions, theorem statements, and proofs can be found in subsection 5.2 for the continuous case and subsection 5.1 for the discrete case.

We also have pointwise convergence results in section 6 for both the discrete and continuous cases. One new result gives that if all POD eigenvalues for the data $\{Lw_j\}$ are nonzero and $\{\Pi_r^Y\}$ is uniformly bounded, then $\Pi_r^Y y \to y$ for all $y \in Y$ as r increases. We also prove error bounds for pointwise convergence of the other projections considered. Boundedness of either L or L^{-1}, along with various range conditions, also plays important roles in the pointwise convergence of these POD projections.

In the pointwise convergence result for Π_r^Y mentioned above, we required all of the POD eigenvalues for $\{Lw_j\}$ to be nonzero. This improves on a similar result from our earlier work [40], where we assumed all of the POD *singular values* are nonzero. The current result is less restrictive; see subsections 2.2 and 2.3. We also explore the boundedness of certain nonorthogonal POD projections in subsection 4.2 and the relationship between the two sets of POD singular values for the data $\{w_j\}$ and the data $\{Lw_j\}$ in subsection 4.3.

3.4. A brief example. Next, we briefly present numerical results for an example to demonstrate our new results. POD model order reduction is considered for this example in [44]; here, we focus on the POD data approximation errors. The new results are discussed in greater detail for other examples in section 7.

Consider a nerve impulse model, the FitzHugh–Nagumo system in one dimension. This model is given by

$$
\frac{\partial u(t,x)}{\partial t} = \mu \frac{\partial^2 u(t,x)}{\partial x^2} - \frac{1}{\mu} v(t,x) + \frac{1}{\mu} f(u) + \frac{c}{\mu}, \quad 0 < x < 1,
$$

$$
\frac{\partial v(t,x)}{\partial t} = bu(t,x) - \gamma v(t,x) + c, \quad 0 < x < 1,
$$

where

$$
f(u) = u(u - 0.1)(1 - u),
$$

$\mu = 0.015$, $b = 0.5$, $\gamma = 2$, and $c = 0.05$. Further, the boundary conditions are given by

$$
u_x(t,0) = -50000t^3 e^{-15t} \quad \text{and} \quad u_x(t,1) = 0,
$$

and the initial conditions are zero.

For this example, we take the Hilbert spaces $X = Y = L^2(0,1) \times L^2(0,1)$ with the usual inner product and define the operator $L : X \to Y$ by

$$
L \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \partial_x u \\ \partial_x v \end{bmatrix}.
$$

Note that here L is unbounded and closed, but not invertible. Thus, this operator satisfies the main assumption made for the continuous case. We let Π_r^Y be the orthogonal projection onto $Y_r = \text{span}\{L\varphi_k\}_{k=1}^r$, where $\{\varphi_k\} \subset X$ are the POD modes.

TABLE 1
Error comparison with $r = 4$.

POD error equation	Actual error	Error formula	Difference
Equation (2.7)	6.2755×10^{-5}	6.2792×10^{-5}	3.7584×10^{-8}
Equation (3.3)	2.1584×10^{-1}	2.1593×10^{-1}	9.1863×10^{-5}
Equation (3.4)	9.8536×10^{-3}	9.8541×10^{-3}	4.7712×10^{-7}

TABLE 2
Error comparison with $r = 12$.

POD error formula	Actual error	Error formula	Difference
Equation (2.7)	4.1453×10^{-8}	4.1487×10^{-8}	3.3661×10^{-11}
Equation (3.3)	2.2536×10^{-4}	2.2541×10^{-4}	5.2146×10^{-8}
Equation (3.4)	1.2664×10^{-5}	1.2668×10^{-5}	3.5150×10^{-9}

To approximate the solution of the PDE we used the interpolated coefficient finite element method with continuous piecewise linear basis functions from [44] and used `ode23s` from MATLAB for the time stepping scheme. We approximated the solution using 100 equally spaced finite element nodes on the time interval $\mathcal{O} = (0, 10)$. Increasing the number of finite element nodes gave similar results below.

For the POD computations, the solution values were approximated at each time step, $w(t_k)$, where $w = [u, v]^T$, and a piecewise constant function in time was formed. The constant on each interval is given by the average of the solution at the current step and the solution at the next step, i.e., $0.5(w(t_{k+1}) + w(t_k))$. Note that for this problem we can calculate the POD eigenvalues, POD modes, and data approximation errors exactly. Thus, comparisons between the actual approximation errors and the error formulas can be made.

In Tables 1 and 2 we present the errors from the relevant projections considered in this paper for $r = 4$ and $r = 12$. Note that errors for projections involving the inverse mapping L^{-1} are not included since L is not invertible for this example. In the tables, the actual error is the integral error measure and the error formula is the sum involving the POD singular values. The first line in the tables represents computations for the known error result (2.7). The second and third lines of the tables are computations for the new results (3.3)–(3.4). The second line of each table gives the values for

$$\text{actual error} \ = \int_{\mathcal{O}} \left\| Lw(t) - L\Pi_r^X w(t) \right\|_Y^2 dt, \quad \text{error formula} \ = \sum_{k>r} \sigma_k^2 \left\| L\varphi_k \right\|_Y^2 ,$$

while the third line of each table shows computational results for

$$\text{actual error} \ = \int_{\mathcal{O}} \left\| Lw(t) - \Pi_r^Y Lw(t) \right\|_Y^2 dt, \text{error formula} \ = \sum_{k>r} \sigma_k^2 \left\| L\varphi_k - \Pi_r^Y L\varphi \right\|_Y^2 .$$

The differences in the computed values are likely due to round-off errors. Note that as r increases the errors tend toward zero, as expected by the theory.

4. POD properties. In this section, we consider three topics. In subsection 4.1, we give two results concerning Hilbert–Schmidt operator norms of POD operators and approximations of POD operators. These Hilbert–Schmidt results are used throughout sections 5 and 6. In subsection 4.2, we study the boundedness of various nonorthogonal POD projections. These boundedness results are used in section 6. In

subsection 4.3, we study the relationship between POD singular values and POD eigenvalues for different data. This investigation is motivated by some results in section 6, where we assume the POD eigenvalues of different data are all nonzero.

4.1. Hilbert–Schmidt results for POD operators. Below, we give two Hilbert–Schmidt results concerning POD operators. The first result is known (see, e.g., [3, section 3.5], [2, Theorem 12.6.1], [39, Lemma 4.4]), although perhaps not exactly in this precise form. We provide a proof to be complete, and also since the result is crucial to this work.

LEMMA 4.1. *Let Z be a separable Hilbert space, and let $S = L^2(\mathcal{O}; \mathbb{K}^m)$, where $\mathcal{O}$ is an open subset of $\mathbb{R}^d$. If $K : S \to Z$ is defined by*

$$Kf = \sum_{j=1}^m \int_{\mathcal{O}} f_j(t) z_j(t)\, dt,$$

for $\{z_j\}_{j=1}^m \subset L^2(\mathcal{O}; Z)$, then K is Hilbert–Schmidt and

$$\|K\|_{\mathrm{HS}(S,Z)}^2 = \sum_{j=1}^m \|z_j\|_{L^2(\mathcal{O};Z)}^2.$$

Proof. Let $\{\chi_i\}_{i\geq 1} \subset L^2(\mathcal{O})$ and $\{\xi_n\}_{n\geq 1} \subset Z$ be orthonormal bases. Therefore, $\{\overline{\chi}_i\}_{i\geq 1}$ is also an orthonormal basis for $L^2(\mathcal{O})$, and $\{\chi_i \xi_n\}_{i,n\geq 1}$ is an orthonormal basis for $L^2(\mathcal{O}; Z)$ (see, e.g., [2, Theorem 12.6.1]).

For $\xi \in Z$, let $[K^*\xi]_j = (\xi, z_j(t))_Z$ denote the jth component of $K^*\xi \in S$. Working with the Hilbert adjoint operator K^* and using Parseval's equality gives

$$\begin{aligned}
\|K^*\|_{\mathrm{HS}(S,Z)}^2 &= \sum_{n\geq 1} \|K^*\xi_n\|_S^2 \\
&= \sum_{j=1}^m \sum_{n\geq 1} \left\| [K^*\xi_n]_j \right\|_{L^2(\mathcal{O})}^2 \\
&= \sum_{j=1}^m \sum_{n,i\geq 1} \left| \left(\overline{\chi}_i, [K^*\xi_n]_j \right)_{L^2(\mathcal{O})} \right|^2 \\
&= \sum_{j=1}^m \sum_{n,i\geq 1} \left| \int_{\mathcal{O}} \overline{\chi_i(t)}\, (z_j(t), \xi_n)_Z\, dt \right|^2 \\
&= \sum_{j=1}^m \sum_{n,i\geq 1} \left| \int_{\mathcal{O}} (z_j(t), \chi_i(t)\xi_n)_Z\, dt \right|^2 \\
&= \sum_{j=1}^m \sum_{n,i\geq 1} \left| (z_j, \chi_i\xi_n)_{L^2(\mathcal{O};Z)} \right|^2 \\
&= \sum_{j=1}^m \|z_j\|_{L^2(\mathcal{O};Z)}^2. \qquad \square
\end{aligned}$$

The next result gives three different Hilbert–Schmidt norm approximation results involving the POD operator K for the data $\{w_j\}$ and the POD operator $K^Y = LK$ for the data $\{Lw_j\}$. The result will be of particular usefulness when discussing the continuous case in subsection 5.2, but it applies to the discrete case as well. We also use this result throughout section 6.

LEMMA 4.2. *The Hilbert–Schmidt norm errors are given by*

$$(4.1) \qquad \left\| LK - L\Pi_r^X K \right\|_{\mathrm{HS}(S,Y)}^2 = \sum_{k>r} \sigma_k^2 \| L\varphi_k \|_Y^2,$$

$$(4.2) \qquad \left\| LK - \Pi_r^Y LK \right\|_{\mathrm{HS}(S,Y)}^2 = \sum_{k>r} \sigma_k^2 \left\| L\varphi_k - \Pi_r^Y L\varphi_k \right\|_Y^2,$$

and

$$(4.3) \qquad \left\| K - L^{-1}\Pi_r^Y LK \right\|_{\mathrm{HS}(S,X)}^2 = \sum_{k>r} \sigma_k^2 \left\| \varphi_k - L^{-1}\Pi_r^Y L\varphi_k \right\|_X^2.$$

In the case $s_X = \infty$, the following convergence results hold. For (4.1), *the error tends to zero as $r \to \infty$. For* (4.2), *if $\{\Pi_r^Y\}$ is uniformly bounded in operator norm, then the error goes to zero as $r \to \infty$. For* (4.3), *if L^{-1} is bounded and $\{\Pi_r^Y\}$ is uniformly bounded in operator norm, then the error tends to zero as $r \to \infty$. For* (4.3), *if $\{L^{-1}\Pi_r^Y L\}$ is uniformly bounded in operator norm, then the error converges to zero as $r \to \infty$.*

Proof. Let $\{f_k\}$ be an orthonormal basis of S of eigenvectors of K^*K and let $\mathbb{J} = \{k : f_k \notin \ker(K^*K)\}$. Note that $Kf_k = 0$ for all $k \notin \mathbb{J}$, since $\ker(K^*K) = \ker(K)$ by Lemma 2.1. Also, $Kf_k = \sigma_k \varphi_k$ for all $k \in \mathbb{J}$. Then,

$$\begin{aligned}
\left\| LK - L\Pi_r^X K \right\|_{\mathrm{HS}(S,Y)}^2 &= \sum_{k \geq 1} \left\| (LK - L\Pi_r^X K) f_k \right\|_Y^2 \\
&= \sum_{k \in \mathbb{J}} \left\| (LK - L\Pi_r^X K) f_k \right\|_Y^2 \\
&= \sum_{k \in \mathbb{J}} \left\| L\sigma_k \varphi_k - L\Pi_r^X \sigma_k \varphi_k \right\|_Y^2 \\
&= \sum_{k > r,\, k \in \mathbb{J}} \sigma_k^2 \| L\varphi_k \|_Y^2,
\end{aligned}$$

where the last equality holds since $\Pi_r^X \varphi_k = \varphi_k$ for $k \leq r$ and $\Pi_r^X \varphi_k = 0$ for $k > r$. Also,

$$\sum_{k > r,\, k \in \mathbb{J}} \sigma_k^2 \| L\varphi_k \|_Y^2 = \sum_{k > r,\, k \in \mathbb{J}} \| LKf_k \|_Y^2 = \sum_{k > r,\, k \in \mathbb{J}} \| K^Y f_k \|_Y^2,$$

which converges to zero as $r \to \infty$ since K^Y is Hilbert–Schmidt.

Next,

$$\begin{aligned}
\| LK - \Pi_r^Y LK \|_{\mathrm{HS}(S,Y)}^2 &= \sum_{k \geq 1} \left\| (LK - \Pi_r^Y LK) f_k \right\|_Y^2 \\
&= \sum_{k \in \mathbb{J}} \left\| L\sigma_k \varphi_k - \Pi_r^Y L\sigma_k \varphi_k \right\|_Y^2 \\
&= \sum_{k > r,\, k \in \mathbb{J}} \sigma_k^2 \left\| L\varphi_k - \Pi_r^Y L\varphi_k \right\|_Y^2,
\end{aligned}$$

where the last equality holds since $\Pi_r^Y L\varphi_k = L\varphi_k$ for $k \leq r$. For convergence, note

$$
\begin{aligned}
\left\|LK - \Pi_r^Y LK\right\|_{\mathrm{HS}(S,Y)}^2 &= \sum_{k>r,\, k\in\mathbb{J}} \left\|LKf_k - \Pi_r^Y LKf_k\right\|_Y^2 \\
&= \sum_{k>r,\, k\in\mathbb{J}} \left\|K^Y f_k - \Pi_r^Y K^Y f_k\right\|_Y^2 \\
&\leq \sum_{k>r,\, k\in\mathbb{J}} \left\|I - \Pi_r^Y\right\|^2 \left\|K^Y f_k\right\|_Y^2 .
\end{aligned}
$$

Since $\|I - \Pi_r^Y\|$ is uniformly bounded and K^Y is Hilbert–Schmidt, the error converges to zero as $r \to \infty$.

Similarly, for the last equality we have

$$
\begin{aligned}
\left\|K - L^{-1}\Pi_r^Y LK\right\|_{\mathrm{HS}(S,X)}^2 &= \sum_{k\in\mathbb{J}} \left\|\sigma_k\varphi_k - L^{-1}\Pi_r^Y L\sigma_k\varphi_k\right\|_X^2 \\
&= \sum_{k>r,\, k\in\mathbb{J}} \sigma_k^2 \left\|\varphi_k - L^{-1}\Pi_r^Y L\varphi_k\right\|_X^2 ,
\end{aligned}
$$

since $L^{-1}\Pi_r^Y L\varphi_k = L^{-1}L\varphi_k = \varphi_k$ for $k \leq r$.

Assuming L^{-1} is bounded and $\{\Pi_r^Y\}$ is uniformly bounded, the convergence follows from

$$
\begin{aligned}
\left\|K - L^{-1}\Pi_r^Y LK\right\|_{\mathrm{HS}(S,X)}^2 &= \sum_{k>r,\, k\in\mathbb{J}} \left\|L^{-1}(I - \Pi_r^Y)LKf_k\right\|_X^2 \\
&\leq \sum_{k>r,\, k\in\mathbb{J}} \left\|L^{-1}\right\|^2 \left\|I - \Pi_r^Y\right\|^2 \left\|K^Y f_k\right\|_Y^2
\end{aligned}
$$

in a similar manner to the previous case. For the second convergence case, we assume $\{L^{-1}\Pi_r^Y L\}$ is uniformly bounded in operator norm and we have

$$
\begin{aligned}
\left\|K - L^{-1}\Pi_r^Y LK\right\|_{\mathrm{HS}(S,X)}^2 &= \sum_{k>r,\, k\in\mathbb{J}} \left\|(I - L^{-1}\Pi_r^Y L)Kf_k\right\|_X^2 \\
&\leq \sum_{k>r,\, k\in\mathbb{J}} \left\|I - L^{-1}\Pi_r^Y L\right\|^2 \left\|K^Y f_k\right\|_X^2 ,
\end{aligned}
$$

which converges to zero as $r \to \infty$. $\qquad\square$

4.2. Nonorthogonal POD projections. In section 6, we consider pointwise convergence results for the linear operators $L^{-1}\Pi_r^Y L : X \to X$ and $L\Pi_r^X L^{-1} : Y \to Y$. Below, we give conditions that guarantee that these linear operators are bounded, or have bounded extensions, for r fixed. We note that when these operators are bounded we have $L^{-1}\Pi_r^Y L : X \to X$ is a projection onto $X_r = \mathrm{span}\{\varphi\}_{j=1}^r$ and $L\Pi_r^X L^{-1} : Y \to Y$ is a projection onto $Y_r = \mathrm{span}\{L\varphi\}_{j=1}^r$. Even if Π_r^Y is an orthogonal projection, these projections are typically nonorthogonal POD projection operators.

In the simplest case, if L and L^{-1} are bounded, then clearly $L^{-1}\Pi_r^Y L : X \to X$ and $L\Pi_r^X L^{-1} : Y \to Y$ are both bounded for each r. In this case, $\{L\Pi_r^X L^{-1}\}$ is uniformly bounded in operator norm, and $\{L^{-1}\Pi_r^Y L\}$ is also uniformly bounded when $\{\Pi_r^Y\}$ is uniformly bounded.

Below, we consider the case when either L or L^{-1} is unbounded. For each fixed r, we show $L^{-1}\Pi_r^Y L : X \to X$ is bounded when L is bounded, and $L\Pi_r^X L^{-1} : Y \to Y$

is bounded when L^{-1} is bounded. In other cases, we need certain assumptions to be satisfied to construct bounded extensions of the operators for each r. We do *not* show that these nonorthogonal POD projection operators are uniformly bounded in operator norm.

In specific cases, we need certain adjoint operators to exist and therefore we need the operators to be densely defined or bounded. For example, for the operator L^{-*} to exist we must assume that $\mathcal{D}(L^{-1})$ is dense in Y, or L^{-1} is bounded. These type of assumptions must be added to the second and fourth parts of the following theorem, in addition to results later in this paper.

THEOREM 4.3. *Assume L is invertible and $r > 0$ is fixed.*
1. *If L^{-1} is bounded, then $L\Pi_r^X L^{-1} : Y \to Y$ is bounded.*
2. *If $\mathcal{D}(L^{-1})$ is dense in Y and $\{\varphi_k\}_{k=1}^r \subset \mathcal{D}(L^{-*})$, the operator $L\Pi_r^X L^{-1} : Y \to Y$ can be extended to a bounded operator on Y.*
3. *If L is bounded, then $L^{-1}\Pi_r^Y L : X \to X$ is bounded.*
4. *Assume $\Pi_r^Y : Y \to Y$ is the orthogonal projection onto $\operatorname{span}\{L\varphi_k\}_{k=1}^r$. If L^{-1} is bounded, $\mathcal{D}(L)$ is dense, and $\{L\varphi_k\}_{k=1}^r \subset \mathcal{D}(L^*)$, then $L^{-1}\Pi_r^Y L : X \to X$ can be extended to a bounded operator on X.*

Remark 4.4. In the second and fourth items, we assume the POD modes satisfy the regularity properties $\{\varphi_k\}_{k=1}^r \subset \mathcal{D}(L^{-*})$ and $\{L\varphi_k\}_{k=1}^r \subset \mathcal{D}(L^*)$, respectively. See section 3 for conditions on the data in the discrete and continuous cases that guarantee these properties hold.

Proof. For item 1, note that for $y \in Y$ we have

$$(4.4) \qquad L\Pi_r^X L^{-1} y = \sum_{k=1}^r (L^{-1}y, \varphi_k)_X L\varphi_k.$$

Since L^{-1} is a bounded operator and $\varphi_k \in \mathcal{D}(L)$ for all k, the sum in (4.4) is well-defined for all $y \in Y$. Also, it can be checked that

$$(4.5) \qquad \left\| L\Pi_r^X L^{-1} y \right\|_Y \le c\|y\|_Y,$$

where the constant $c := \|L^{-1}\| \left(\sum_{k=1}^r \|L\varphi_k\|_Y^2 \right)^{1/2}$ depends on r. This shows that the operator $L\Pi_r^X L^{-1}$ is bounded when L^{-1} is bounded.

For item 2, the linear operator $L\Pi_r^X L^{-1} : Y \to Y$ is defined by (4.4) for all $y \in \mathcal{D}(L^{-1})$. Using the assumptions, we can rewrite (4.4) for $y \in \mathcal{D}(L^{-1})$ as

$$(4.6) \qquad L\Pi_r^X L^{-1} y = \sum_{k=1}^r (y, L^{-*}\varphi_k)_Y L\varphi_k.$$

It can be checked that (4.5) holds for all $y \in \mathcal{D}(L^{-1})$ with the constant c given by $c := \sum_{k=1}^r \|L^{-*}\varphi_k\|_Y \|L\varphi_k\|_Y$. Note that (4.6) is well-defined for all $y \in Y$ and therefore yields a bounded linear extension of $L\Pi_r^X L^{-1} : Y \to Y$ to all of Y.

For item 3, since Π_r^Y is a projection onto $Y_r = \operatorname{span}\{L\varphi_j\}_{j=1}^r$, we know for $y \in Y$ there exist constants $\{\alpha_j(y)\}$ depending on y such that $\Pi_r^Y y = \sum_{j=1}^r \alpha_j(y)L\varphi_j$. Then

$$(4.7) \qquad \left\| \sum_{j=1}^r \alpha_j(y)L\varphi_j \right\|_Y = \left\| \Pi_r^Y y \right\|_Y \le \left\| \Pi_r^Y \right\| \|y\|_Y.$$

Also,

$$\left\|\sum_{j=1}^{r} \alpha_j(y) L\varphi_j\right\|_Y^2 = \sum_{j,k=1}^{r} \alpha_j(y)(L\varphi_j, L\varphi_k)_Y \overline{\alpha_k(y)} = \alpha(y)^* A_r \alpha(y),$$

where the star denotes complex conjugate, and

$$\alpha(y) = [\alpha_1(y), \ldots, \alpha_r(y)]^T \in \mathbb{K}^r, \quad [A_r]_{i,j} = (L\varphi_i, L\varphi_j)_Y.$$

Since L is invertible and $\{\varphi_j\}_{j=1}^r$ is a linearly independent set, we know $\{L\varphi_j\}_{j=1}^r$ is a linearly independent set; therefore, A_r is symmetric positive definite, which implies there exists $\beta > 0$ such that $\alpha^* A_r \alpha \geq \beta\|\alpha\|_{\mathbb{K}^r}^2$ for all $\alpha \in \mathbb{K}^r$. Note that β may depend on r. Together, the above implies that

$$\beta\|\alpha(y)\|_{\mathbb{K}^r}^2 \leq \left\|\sum_{j=1}^{r} \alpha_j(y) L\varphi_j\right\|_Y^2 \leq \left\|\Pi_r^Y\right\|^2 \|y\|_Y^2.$$

So,

(4.8) $$\|\alpha(y)\|_{\mathbb{K}^r} \leq \beta^{-1/2} \left\|\Pi_r^Y\right\| \|y\|_Y.$$

In this case, $y = Lx$ for $x \in X$ and L is bounded and invertible; thus,

$$L^{-1}\Pi_r^Y(Lx) = L^{-1}\sum_{j=1}^{r} \alpha_j(Lx) L\varphi_j = \sum_{j=1}^{r} \alpha_j(Lx)\varphi_j,$$

where the constants α_j now depend on Lx. Since $\{\varphi_j\} \subset X$ is orthonormal, we have

$$\begin{aligned}
\left\|L^{-1}\Pi_r^Y Lx\right\|_X^2 &= \left\|\sum_{j=1}^{r} \alpha_j(Lx)\varphi_j\right\|_X^2 \\
&= \|\alpha(Lx)\|_{\mathbb{K}^r}^2 \\
&\leq \beta^{-1} \left\|\Pi_r^Y\right\|^2 \|Lx\|_Y^2 \\
&\leq \beta^{-1} \left\|\Pi_r^Y\right\|^2 \|L\|^2\|x\|_X^2.
\end{aligned}$$

Therefore, for all $x \in X$ we have

(4.9) $$\left\|L^{-1}\Pi_r^Y Lx\right\|_X \leq c\|x\|_X,$$

where $c := \beta^{-1/2}\|\Pi_r^Y\|\|L\|$.

For item 4, we obtain a representation of $L^{-1}\Pi_r^Y L$ as follows. First, note that the sets $\{L\varphi_k\}$ and $\{L^{-*}\varphi_k\}$ are biorthogonal, i.e., $(L\varphi_k, L^{-*}\varphi_j)_Y = \delta_{k,j}$, where $\delta_{k,j}$ is the Kronecker delta symbol. Recall from the proof of part 3 that $\Pi_r^Y y = \sum_{k=1}^{r} \alpha_k L\varphi_k$ for some scalars α_k that depend on y. We can calculate the values for α_k by noting

$$\left(\Pi_r^Y y, L^{-*}\varphi_j\right)_Y = \sum_{k=1}^{r} \alpha_k (L\varphi_k, L^{-*}\varphi_j)_Y = \alpha_j.$$

This yields

(4.10) $$\Pi_r^Y y = \sum_{k=1}^{r} (\Pi_r^Y y, L^{-*}\varphi_k)_Y L\varphi_k = \sum_{k=1}^{r} (y, \Pi_r^Y L^{-*}\varphi_k)_Y L\varphi_k,$$

since Π_r^Y is orthogonal and therefore $(\Pi_r^Y)^* = \Pi_r^Y$.

SARAH LOCKE AND JOHN SINGLER

By assumption, $\{L\varphi_j\} \subset \mathcal{D}(L^*)$ and so (4.10) implies $\Pi_r^Y y \in \mathcal{D}(L^*)$ for all $y \in Y$. This gives the following representation for any $x \in \mathcal{D}(L)$:

$$(4.11) \qquad L^{-1}\Pi_r^Y Lx = \sum_{k=1}^r (x, L^*\Pi_r^Y L^{-*}\varphi_k)_X \varphi_k.$$

Also, for all $x \in \mathcal{D}(L)$, the bound (4.9) holds with $c := \left(\sum_{k=1}^r \|L^*\Pi_r^Y L^{-*}\varphi_k\|_X^2\right)^{1/2}$. Equation (4.11) is well-defined for all $x \in X$ and therefore defines a bounded linear extension of $L^{-1}\Pi_r^Y L : X \to X$ to all of X. $\qquad\square$

4.3. POD singular values and POD eigenvalues. The number of nonzero singular values (or eigenvalues) of the POD operators plays an important role throughout the paper. It is also important to note the difference between singular values and eigenvalues. For a POD operator $K : S \to Z$, recall the POD eigenvalues are the eigenvalues of $KK^* : Z \to Z$, the POD singular values are the singular values of K, and $s_Z = \mathrm{rank}(K)$, i.e., s_Z is the number of positive POD singular values of K (or positive POD eigenvalues of KK^*). As discussed in section 2, it is possible to have a zero POD singular value but to have all nonzero POD eigenvalues.

Below, we study various relationships between the POD eigenvalues and POD singular values for the data $\{w_j\}$ and the data $\{Lw_j\}$. Recall that $K : S \to X$ is the POD operator for the data $\{w_j\}$, and $K^Y = LK : S \to Y$ is the POD operator for the data $\{Lw_j\}$. Therefore, $s_X = \mathrm{rank}(K)$ is the number of nonzero POD singular values (or POD eigenvalues) for the data $\{w_j\}$, and $s_Y = \mathrm{rank}(K^Y)$ is the number of nonzero POD singular values (or POD eigenvalues) for the data $\{Lw_j\}$

First, we give a relationship between the POD eigenvalues and the null space of the adjoint POD operator and also give some additional information about s_X and s_Y.

LEMMA 4.5. 1. *All of the POD eigenvalues for the data $\{w_j\}$ are nonzero if and only if $\ker(K^*) = \{0\}$. In this case, $X = \overline{\mathcal{R}(K)}$. In addition, if $s_X < \infty$, then $X = \mathcal{R}(K)$ and $\dim(X) = s_X$.*
  2. *All of the POD eigenvalues for the data $\{Lw_j\}$ are nonzero if and only if $\ker((K^Y)^*) = \{0\}$. In this case, $Y = \overline{\mathcal{R}(K^Y)}$. In addition, if $s_Y < \infty$, then $Y = \mathcal{R}(K^Y)$ and $\dim(Y) = s_Y$.*
  3. *The number of nonzero POD eigenvalues for $\{Lw_j\}$ is less than or equal to the number of nonzero POD eigenvalues for $\{w_j\}$. That is, $s_Y \leq s_X$.*
  4. *If L is invertible, then $s_X = s_Y$.*

Proof. The first two items are proven similarly. Here we show item 1. Lemma 2.1 proves the first statement. To see the rest, note that $X = \ker(K^*) \oplus \overline{\mathcal{R}(K)}$ and $\ker(K^*) = \{0\}$ imply $X = \overline{\mathcal{R}(K)}$. Then if $s_X = \mathrm{rank}(K) = \dim(\mathcal{R}(K))$ is finite, we have $\overline{\mathcal{R}(K)} = \mathcal{R}(K)$ and therefore $X = \mathcal{R}(K)$ and $\dim(X) = s_X$.

For item 3, first if $s_X = \infty$, we are done. Assume $s_X < \infty$. We know

$$Kf = \sum_{j=1}^{s_X} \sigma_j(f, f_j)_S \varphi_j,$$

and therefore

$$K^Y f = LKf = \sum_{j=1}^{s_X} \sigma_j(f, f_j)_S L\varphi_j.$$

Thus, $s_Y = \mathrm{rank}(K^Y) \leq s_X$.

For item 4, because of item 3, we need only show $s_X \leq s_Y$. First, if $s_Y = \infty$, we are done. Assume $s_Y < \infty$. Let the SVD of K^Y be given by

$$K^Y f = LKf = \sum_{j=1}^{s_Y} \sigma_j^Y (f, f_j^Y)_S \varphi_j^Y.$$

Note that $\varphi_j^Y \in \mathcal{D}(L^{-1})$ whenever $\sigma_j^Y > 0$, since $\mathcal{D}(L^{-1}) = \mathcal{R}(L)$ and $\varphi_j^Y = (\sigma_j^Y)^{-1} K^Y f_j^Y = (\sigma_j^Y)^{-1} LK f_j^Y$. Then, since L is invertible,

$$Kf = L^{-1}LKf = L^{-1}K^Y f = \sum_{j=1}^{s_Y} \sigma_j^Y (f, f_j^Y)_S L^{-1} \varphi_j^Y,$$

and therefore $s_X = \operatorname{rank}(K) \leq s_Y$. $\qquad\square$

The following lemma gives further results about the connections between the two main sets of POD eigenvalues under consideration in this paper, i.e., the POD eigenvalues for the data $\{w_j\}$ and the data $\{Lw_j\}$. With extra assumptions, we can use the fact that all the POD eigenvalues are nonzero for one set of data to obtain the same conclusion for the other set of data.

LEMMA 4.6. 1. *If L is bounded, $\mathcal{R}(L)$ is dense in Y, and the POD eigenvalues for $\{w_j\}$ are all nonzero, then the POD eigenvalues for $\{Lw_j\}$ are all nonzero.*
2. *If L^{-1} is bounded, $\mathcal{R}(L^{-1})$ is dense in X, and the POD eigenvalues for $\{Lw_j\}$ are all nonzero, then the POD eigenvalues for $\{w_j\}$ are all nonzero.*

Proof. The proofs of the two items are similar; we prove only the first item.

Since $X = \ker(K^*) \oplus \overline{\mathcal{R}(K)}$ and $\ker(K^*) = \{0\}$ (Lemma 4.5, item 1), we have $X = \overline{\mathcal{R}(K)}$. Let $\varepsilon > 0$ and let $y \in Y$. Since $\mathcal{R}(L)$ is dense in Y, there exists $x \in X$ such that $\|y - Lx\|_Y < \varepsilon/2$. Since $X = \overline{\mathcal{R}(K)}$, for this x there exists $f \in S$ such that $\|x - Kf\|_X < \varepsilon/(2\|L\|)$. This gives

$$\|y - LKf\|_Y < \|y - Lx\|_Y + \|Lx - LKf\|_Y < \frac{\varepsilon}{2} + \|L\| \frac{\varepsilon}{2\|L\|} < \varepsilon,$$

which shows $\overline{\mathcal{R}(K^Y)} = Y$ and $\ker((K^Y)^*) = \{0\}$. Thus, the POD eigenvalues for $\{Lw_j\}$ are all nonzero by Lemma 4.5, Item 2. $\qquad\square$

5. Error formulas. One goal of this paper is to provide exact formulas for POD data approximation errors. The two main results of this section can be found in Theorems 5.1 and 5.4. The section is split between the discrete case, where we can use a more direct proof approach, and the continuous case, which requires more care since the data can have infinitely many nonzero POD eigenvalues.

5.1. Discrete case. First we introduce several representations that will be useful in the proof of Theorem 5.1 below. Recall, $s_X = \operatorname{rank}(K) < \infty$ is the number of nonzero POD singular values (or POD eigenvalues) for the data $\{w_j\}$. By the known POD error formula (2.5), we have

$$w_j = \Pi_{s_X}^X w_j = \sum_{k=1}^{s_X} (w_j, \varphi_k)_X \varphi_k \quad \text{and} \quad Lw_j = \sum_{k=1}^{s_X} (w_j, \varphi_k)_X L\varphi_k.$$

Note that since the sums are finite, $\{\varphi_k\} \subset \mathcal{D}(L)$, and L is linear we can pull L through the sums in this section without any additional assumptions. This is one point where the discrete and continuous cases differ.

Next, from subsection 2.2 we know for all $j \leq s$ and $k \leq s_X$ we have

$$(w_j, \varphi_k)_X = \overline{(\varphi_k, w_j)_X} = \overline{(K^* \varphi_k)_j} = \sigma_k \overline{f_{k,j}},$$

where $f_{k,j}$ denotes the jth component of the singular vector $f_k \in \mathbb{K}^s$. This gives

$$(5.1) \qquad w_j = \sum_{k=1}^{s_X} \sigma_k \overline{f_{k,j}} \, \varphi_k \quad \text{and} \quad L w_j = \sum_{k=1}^{s_X} \sigma_k \overline{f_{k,j}} \, L\varphi_k.$$

Also, recall $\{f_j\}$ are orthonormal in S, which yields

$$\sum_{j=1}^{s} \gamma_j \overline{f_{k,j}} f_{\ell,j} = (f_\ell, f_k)_S = \delta_{\ell,k}.$$

THEOREM 5.1. *The data approximation errors are given by*

$$(5.2) \qquad \sum_{j=1}^{s} \gamma_j \| L w_j - L\Pi_r^X w_j \|_Y^2 = \sum_{k=r+1}^{s_X} \sigma_k^2 \| L\varphi_k \|_Y^2$$

and

$$(5.3) \qquad \sum_{j=1}^{s} \gamma_j \| L w_j - \Pi_r^Y L w_j \|_Y^2 = \sum_{k=r+1}^{s_X} \sigma_k^2 \| L\varphi_k - \Pi_r^Y L\varphi_k \|_Y^2.$$

Also, if L is invertible, then

$$(5.4) \qquad \sum_{j=1}^{s} \gamma_j \| w_j - L^{-1}\Pi_r^Y L w_j \|_X^2 = \sum_{k=r+1}^{s_X} \sigma_k^2 \| \varphi_k - L^{-1}\Pi_r^Y L\varphi_k \|_X^2.$$

Proof. We prove only (5.3). The proofs of the other two results are similar. First, note we can apply Π_r^Y to $L w_j$ given in (5.1) to get

$$L w_j - \Pi_r^Y w_j = \sum_{k=1}^{s_X} \sigma_k \overline{f_{k,j}} \, (L\varphi_k - \Pi_r^Y L\varphi_k).$$

Then

$$\sum_{j=1}^{s} \gamma_j \| L w_j - \Pi_r^Y L w_j \|_Y^2$$

$$= \sum_{j=1}^{s} \gamma_j \left(\sum_{k=1}^{s_X} \sigma_k \overline{f_{k,j}} (L\varphi_k - \Pi_r^Y L\varphi_k), \sum_{\ell=1}^{s_X} \sigma_\ell \overline{f_{\ell,j}} (L\varphi_\ell - \Pi_r^Y L\varphi_\ell) \right)_Y$$

$$= \sum_{j=1}^{s} \gamma_j \sum_{\ell,k=1}^{s_X} \sigma_k \sigma_\ell \overline{f_{k,j}} f_{\ell,j} \left(L\varphi_k - \Pi_r^Y L\varphi_k, L\varphi_\ell - \Pi_r^Y L\varphi_\ell \right)_Y$$

$$= \sum_{\ell,k=1}^{s_X} \sigma_k \sigma_\ell \left(\sum_{j=1}^{s} \gamma_j \overline{f_{k,j}} f_{\ell,j} \right) \left(L\varphi_k - \Pi_r^Y L\varphi_k, L\varphi_\ell - \Pi_r^Y L\varphi_\ell \right)_Y$$

$$= \sum_{k=1}^{s_X} \sigma_k^2 \left(L\varphi_k - \Pi_r^Y L\varphi_k, L\varphi_k - \Pi_r^Y L\varphi_k \right)_Y = \sum_{k=1}^{s_X} \sigma_k^2 \| L\varphi_k - \Pi_r^Y L\varphi_k \|_Y^2.$$

Now note that $\Pi_r^Y L\varphi_k = L\varphi_k$ for $k = 1, \ldots, r$ since Π_r^Y is a projection onto $Y_r = \mathrm{span}\{L\varphi_k\}_{k=1}^r$. Therefore,

$$\sum_{j=1}^s \gamma_j \|Lw_j - \Pi_r^Y Lw_j\|_Y^2 = \sum_{k=r+1}^{s_X} \sigma_k^2 \|L\varphi_k - \Pi_r^Y L\varphi_k\|_Y^2.$$

$\square$

Remark 5.2. In Corollary 6.10, we focus on error bounds for approximating each individual data snapshot w_ℓ with various POD projections. Also, another way to prove Theorem 5.1 is to use the Hilbert–Schmidt norm results in Lemma 4.2. The proof we give above requires less background. However, we do require Lemma 4.2 for the continuous case below.

5.2. Continuous case. For the continuous case we must consider the possibility that the number of nonzero POD eigenvalues is infinite. We approach this case differently from the discrete case above. We show each of the data approximation errors we consider is equal to one of the Hilbert–Schmidt norm errors from Lemma 4.2. We use that result to prove the convergence of the errors to zero in the case of an infinite number of nonzero POD eigenvalues.

For one case, we need to make an additional assumption on L^{-1}.

The L^{-1} assumption. We assume
1. $s_X < \infty$, or
2. $L^{-1}\Pi_r^Y LKf = \sum_{j=1}^m \int_{\mathcal{O}} f_j(t) L^{-1}\Pi_r^Y Lw_j(t)dt$ for all $f \in S$.

Remark 5.3. Note that if $s_X < \infty$, then the proof technique in subsection 5.1 above can be used for the continuous cases, with some minor modifications to deal with the change in the space S. The second condition is similar to the main assumption made in subsection 3.2. Any of the three common conditions in Remark 3.1 that guarantee the main assumption holds also imply that the second condition in the L^{-1} assumption holds.

THEOREM 5.4. *The data approximation errors are given by*

$$(5.5) \qquad \sum_{j=1}^m \left\| Lw_j - L\Pi_r^X w_j \right\|_{L^2(\mathcal{O};Y)}^2 = \sum_{k>r} \sigma_k^2 \|L\varphi_k\|_Y^2$$

and

$$(5.6) \qquad \sum_{j=1}^m \left\| Lw_j - \Pi_r^Y Lw_j \right\|_{L^2(\mathcal{O};Y)}^2 = \sum_{k>r} \sigma_k^2 \left\| L\varphi_k - \Pi_r^Y L\varphi_k \right\|_Y^2.$$

Also, if the L^{-1} assumption holds, then

$$(5.7) \qquad \sum_{j=1}^m \left\| w_j - L^{-1}\Pi_r^Y Lw_j \right\|_{L^2(\mathcal{O};X)}^2 = \sum_{k>r} \sigma_k^2 \left\| \varphi_k - L^{-1}\Pi_r^Y L\varphi_k \right\|_X^2.$$

In the case $s_X = \infty$, the following convergence results hold. For (5.5), the error tends to zero as $r \to \infty$. For (5.6), if $\{\Pi_r^Y\}$ is uniformly bounded in operator norm, then the error goes to zero as $r \to \infty$. For (5.7), if L^{-1} is bounded and $\{\Pi_r^Y\}$ is uniformly bounded in operator norm, then the error tends to zero as $r \to \infty$. For (5.7), if $\{L^{-1}\Pi_r^Y L\}$ is uniformly bounded in operator norm, then the error converges to zero as $r \to \infty$.

Remark 5.5. Note that for the case $s_X = \infty$, the conditions for convergence are exactly the conditions given in Lemma 4.2.

Proof. We prove (5.5) and the associated convergence result. The proofs of the other equalities and convergence results are similar. We first show that the data approximation error has an integral representation, and then we use the two Hilbert–Schmidt results for POD operators from subsection 4.1 to conclude.

By definition, for $f \in S$ we have

$$
\begin{aligned}
L\Pi_r^X Kf &= \sum_{k=1}^r (Kf, \varphi_k)_X L\varphi_k \\
&= \sum_{k=1}^r \left(\sum_{j=1}^m \int_{\mathcal{O}} f_j(t) w_j(t) dt, \varphi_k \right)_X L\varphi_k \\
&= \sum_{j=1}^m \int_{\mathcal{O}} f_j(t) \sum_{k=1}^r (w_j(t), \varphi_k)_X L\varphi_k dt \\
&= \sum_{j=1}^m \int_{\mathcal{O}} f_j(t) L w_j^r(t) dt,
\end{aligned}
$$

where $w_j^r(t) = \Pi_r^X w_j(t) = \sum_{k=1}^r (w_j(t), \varphi_k)_X \varphi_k$. Because of the main assumption, we can pull the operator L inside the integral to give

$$
(LK - L\Pi_r^X K)f = \int_{\mathcal{O}} \sum_{j=1}^m f_j(t)[Lw_j(t) - Lw_j^r(t)] dt.
$$

Since $Lw_j - Lw_j^r \in L^2(\mathcal{O}; Y)$ for each j, by Lemma 4.1 we have

$$
\sum_{j=1}^m \left\| Lw_j - L\Pi_r^X w_j \right\|_{L^2(\mathcal{O};Y)}^2 = \left\| LK - L\Pi_r^X K \right\|_{\mathrm{HS}(S,Y)}^2.
$$

Lemma 4.2 proves both (5.5) and the convergence result in the case $s_X = \infty$.

Note for (5.7), for $f \in S$ the L^{-1} assumption gives

$$
(5.8) \qquad L^{-1}\Pi_r^Y LKf = \int_{\mathcal{O}} \sum_{j=1}^m f_j(t) L^{-1}\Pi_r^Y L w_j(t) dt,
$$

and then we proceed similarly to establish the result. $\qquad \square$

6. Pointwise convergence of POD projections. Recall that $\{\varphi_k\}$ is an orthonormal basis for X, and therefore $\|\Pi_r^X x - x\|_X \to 0$ for all $x \in X$. In this section, we prove various types of pointwise convergence results for the other POD projections, namely, Π_r^Y from subsection 3.3 and $L\Pi_r^X L^{-1}$ and $L^{-1}\Pi_r^Y L$ from subsection 4.2. The majority of this section is not split into the discrete and continuous cases because the proofs are similar for both, and many of the results hold regardless of case. We do focus on the discrete case at the end of this section and address some assumptions made in the literature about approximations of each individual data snapshot using POD projections.

Pointwise convergence results for these POD projections are easiest to obtain when L and L^{-1} are both bounded. We primarily focus on the case when either L or L^{-1} is unbounded.

Range conditions are an important factor in this section. When an element to be approximated by a POD projection is in the range of K or K^Y, we can often get better results. When certain conditions hold, we know these ranges exactly. Recall from Lemma 4.5 that if all the POD eigenvalues for $\{w_j\}$ are nonzero and $s_X < \infty$, then we know $X = \mathcal{R}(K)$ and $\dim(X) = s_X$. Note that in this case, the Hilbert space X must be finite dimensional. If all the POD eigenvalues for $\{w_j\}$ are nonzero and $s_X = \infty$ (i.e., X must be infinite dimensional), then Lemma 4.5 only gives $X = \overline{\mathcal{R}(K)}$. We do not always obtain the better convergence results in this case. Similar statements hold for the spaces Y and $\mathcal{R}(K^Y)$.

Also, as in section 5, we sometimes need to consider different proof techniques in the case $s_X = \infty$.

We begin with a pointwise convergence result for Π_r^Y assuming L is bounded. For another pointwise convergence result for Π_r^Y with different assumptions, see Theorem 6.5 below.

THEOREM 6.1. *Assume L is bounded and $\{\Pi_r^Y\}$ is uniformly bounded in operator norm. If $y \in \mathcal{R}(L)$, then $\Pi_r^Y y \to y$ as r increases. In addition, if $\mathcal{R}(L)$ is dense in Y, then $\Pi_r^Y y \to y$ for all $y \in Y$.*

Proof. Let $y \in \mathcal{R}(L)$, so that $y = Lx$ for some $x \in X$. Note that since $L\Pi_r^X x \in Y_r = \operatorname{span}\{L\varphi_k\}_{k=1}^r$ and Π_r^Y is a projection onto Y_r, we have $\Pi_r^Y L\Pi_r^X x = L\Pi_r^X x$. Then

$$
\begin{aligned}
\left\|\Pi_r^Y y - y\right\|_Y &\le \left\|\Pi_r^Y Lx - \Pi_r^Y L\Pi_r^X x\right\|_Y + \left\|\Pi_r^Y L\Pi_r^X x - Lx\right\|_Y \\
&= \left\|\Pi_r^Y Lx - \Pi_r^Y L\Pi_r^X x\right\|_Y + \left\|L\Pi_r^X x - Lx\right\|_Y \\
&\le \left\|\Pi_r^Y L\right\| \left\|x - \Pi_r^X x\right\|_Y + \|L\| \left\|\Pi_r^X x - x\right\|_Y,
\end{aligned}
$$

which converges to zero as r increases since $\Pi_r^X x \to x$ and $\{\Pi_r^Y\}$ is uniformly bounded in operator norm. The final result follows directly from the Banach–Steinhaus theorem (i.e., the principle of uniform boundedness). □

The next convergence result relies on the boundedness of either L or L^{-1} and certain range conditions involving L.

THEOREM 6.2. 1. *For any $y \in \mathcal{R}(L) = \mathcal{D}(L^{-1})$, if L is bounded, then $\|L\Pi_r^X L^{-1} y - y\|_Y \to 0$ as r increases. In addition, if $\mathcal{R}(L)$ is dense in Y and $\{L\Pi_r^X L^{-1}\}$ is uniformly bounded, then $L\Pi_r^X L^{-1} y \to y$ for all $y \in Y$.*
  2. *For any $x \in \mathcal{D}(L) = \mathcal{R}(L^{-1})$, if L^{-1} is bounded and $\Pi_r^Y y \to y$ for all $y \in Y$ as r increases, then $\|L^{-1}\Pi_r^Y Lx - x\|_X \to 0$ as r increases. In addition, if $\mathcal{D}(L)$ is dense in X and $\{L^{-1}\Pi_r^Y L\}$ is uniformly bounded, then $L^{-1}\Pi_r^Y Lx \to x$ for all $x \in X$.*

Remark 6.3. Note that Theorems 6.1 and 6.5 give two cases where the assumption $\Pi_r^Y y \to y$ for all $y \in Y$ holds. Also, the uniform boundedness of $\{L\Pi_r^X L^{-1}\}$ and $\{L^{-1}\Pi_r^Y L\}$ is not currently known, unless L and L^{-1} are both bounded. Note that when L and L^{-1} are both bounded, Theorem 6.1 gives $\Pi_r^Y y \to y$ for all $y \in Y$ whenever $\{\Pi_r^Y\}$ is uniformly bounded; therefore, in this case Theorem 6.2 gives $L\Pi_r^X L^{-1} y \to y$ for all $y \in Y$ and $L^{-1}\Pi_r^Y Lx \to x$ for all $x \in X$.

Proof. We prove only the first result; the proof of the second is similar. Since $y \in \mathcal{R}(L)$ we have $y = Lx$ for some $x \in X$. Then

$$\left\| L\Pi_r^X L^{-1} y - y \right\|_Y = \left\| L\Pi_r^X x - Lx \right\|_Y$$
$$\leq \|L\| \left\| \Pi_r^X x - x \right\|_X,$$

which converges to zero as r increases. The final convergence result again follows from the principle of uniform boundedness. □

Next, we consider how range conditions involving K and K^Y affect the convergence of POD projections. We are able to obtain convergence rates and at most require either L or L^{-1} to be bounded. We begin with the POD projection Π_r^Y and then consider $L\Pi_r^X L^{-1}$ and $L^{-1}\Pi_r^Y L$. We use the following simple lemma multiple times below.

LEMMA 6.4. *Assume $y \in \mathcal{R}(K^Y)$ so that $y = K^Y g = LKg$ for some $g \in S$. If*

$$(6.1) \qquad y_N = LK_N g = L\Pi_N^X Kg, \quad where \ K_N = \Pi_N^X K,$$

then $y_N \to y$ as N increases.

Proof. As N increases,

$$\|y_N - y\|_Y = \|LKg - LK_N g\|_Y \leq \|LK - L\Pi_N^X K\|_{\mathrm{HS}(S,Y)} \|g\|_S \to 0$$

by Lemma 4.2. □

Recall from Lemma 4.5 that s_Y is always less than or equal to s_X. Thus if we assume $s_X < \infty$, we know that $s_Y < \infty$. For the following proofs, we consider whether s_X is finite or infinite.

THEOREM 6.5. *Assume $\{\Pi_r^Y\}$ is uniformly bounded in operator norm whenever $s_X = \infty$. If $y = K^Y g$ for some $g \in S$, then $\Pi_r^Y y \to y$ as r increases and the following error bound holds:*

$$(6.2) \qquad \|\Pi_r^Y y - y\|_Y \leq \sum_{k>r} \sigma_k |(g,f_k)_S| \, \|\Pi_r^Y L\varphi_k - L\varphi_k\|_Y.$$

Also, if the POD eigenvalues for the data $\{Lw_j\}$ are all nonzero, then $\Pi_r^Y y \to y$ for all $y \in Y$.

Proof. First consider the case $s_X < \infty$, and fix r. Assume $y = K^Y g = LKg$ for some $g \in S$. Thus,
(6.3)
$$\Pi_r^Y y = \Pi_r^Y K^Y g = \Pi_r^Y LKg = \sum_{k=1}^{s_X} \sigma_k(g,f_k)_S \Pi_r^Y L\varphi_k \quad \text{and} \quad y = \sum_{k=1}^{s_X} \sigma_k(g,f_k)_S L\varphi_k.$$

Subtracting gives

$$\Pi_r^Y y - y = \sum_{k=1}^{s_X} \sigma_k(g,f_k)_S \left(\Pi_r^Y L\varphi_k - L\varphi_k \right) = \sum_{k=r+1}^{s_X} \sigma_k(g,f_k)_S \left(\Pi_r^Y L\varphi_k - L\varphi_k \right)$$

since $\Pi_r^Y L\varphi_k = L\varphi_k$ for $k = 1, \ldots, r$. The error bound (6.2) follows directly from this representation and the triangle inequality. Furthermore, since $s_X < \infty$, clearly $\Pi_r^Y y \to y$ as r increases for each $y \in \mathcal{R}(K^Y)$.

Next, assume the POD eigenvalues for the data $\{Lw_j\}$ are all nonzero. By item 2 of Lemma 4.5, since $s_Y \leq s_X < \infty$ we have $Y = \mathcal{R}(K^Y)$. This gives $\Pi_r^Y y \to y$ for all $y \in Y$.

Now consider the case $s_X = \infty$, and fix r. For $y = K^Y g = LKg$ with $g \in S$ as above, recall the definition of $y_N = L\Pi_N^X Kg$ given in (6.1). We have

$$
\begin{aligned}
\left\|\Pi_r^Y y - y\right\|_Y &\leq \left\|\Pi_r^Y y - \Pi_r^Y y_N\right\|_Y + \left\|\Pi_r^Y y_N - y_N\right\|_Y + \left\|y_N - y\right\|_Y \\
&\leq \left(\left\|\Pi_r^Y\right\| + 1\right)\left\|y - y_N\right\|_Y + \left\|\Pi_r^Y y_N - y_N\right\|_Y.
\end{aligned}
$$

Note that for the second term, $\|\Pi_r^Y y_N - y_N\|_Y$, we can obtain representations for $\Pi_r^Y y_N$ and y_N similar to that in (6.3) above. Proceeding in the same way gives

$$
\left\|\Pi_r^Y y_N - y_N\right\|_Y \leq \sum_{k=r+1}^N \sigma_k |(g, f_k)_S| \left\|\Pi_r^Y L\varphi_k - L\varphi_k\right\|_Y.
$$

Since r is fixed and $y_N \to y$ as $N \to \infty$ (Lemma 6.4), the two inequalities above give

$$
\left\|\Pi_r^Y y - y\right\|_Y \leq \sum_{k=r+1}^\infty \sigma_k |(g, f_k)_S| \left\|\Pi_r^Y L\varphi_k - L\varphi_k\right\|_Y.
$$

For convergence, we have

$$
\left\|\Pi_r^Y y - y\right\|_Y \leq \left(\sum_{k>r} |(g, f_k)_S|^2\right)^{1/2} \left(\sum_{k>r} \sigma_k^2 \left\|\Pi_r^Y L\varphi_k - L\varphi_k\right\|_Y^2\right)^{1/2}.
$$

Since $\{f_k\}$ is an orthonormal basis for S, we know $\sum_{k>r} |(g, f_k)_S|^2$ goes to zero as r increases by Parseval's equality. Furthermore, since $\{\Pi_r^Y\}$ is uniformly bounded, Lemma 4.2 gives that $\sum_{k>r} \sigma_k^2 \|\Pi_r^Y L\varphi_k - L\varphi_k\|_Y^2$ goes to zero as r increases. This gives $\Pi_r^Y y \to y$ for each $y \in \mathcal{R}(K^Y)$.

Finally, assume the POD eigenvalues for the data $\{Lw_j\}$ are all nonzero. By item 2 of Lemma 4.5, we have $\mathcal{R}(K^Y)$ is dense in Y. Since $\{\Pi_r^Y\}$ is uniformly bounded, the principle of uniform boundedness gives $\Pi_r^Y y \to y$ for all $y \in Y$. $\qquad\square$

For the next two results we need to assume L or L^{-1} is bounded whenever $s_X = \infty$.

THEOREM 6.6. *Assume $s_X < \infty$, or either L or L^{-1} is bounded. If $y = K^Y g$ for some $g \in S$, then*

$$
(6.4) \qquad \|y - L\Pi_r^X L^{-1} y\|_Y \leq \sum_{k>r} \sigma_k |(g, f_k)_S| \|L\varphi_k\|_Y
$$

and the error converges to zero as r increases. Now assume $\{L\Pi_r^X L^{-1}\}$ is uniformly bounded in operator norm whenever $s_X = \infty$. If the POD eigenvalues for the data $\{Lw_j\}$ are all nonzero, then $L\Pi_r^X L^{-1} y \to y$ for all $y \in Y$.

Proof. Let $y = LKg$ for some $g \in S$, assume $s_X < \infty$, and fix r. As in the proof of Theorem 6.5, it can be shown that

$$
\begin{aligned}
y - L\Pi_r^X L^{-1} y &= LKg - L\Pi_r^X Kg \\
&= \sum_{k=1}^{s_X} \sigma_k (g, f_k)_S L\varphi_k - \sum_{k=1}^r \sigma_k (g, f_k)_S L\varphi_k \\
&= \sum_{k=r+1}^{s_X} \sigma_k (g, f_k)_S L\varphi_k.
\end{aligned}
$$

The triangle inequality gives the error bound (6.4). The convergence results for the case $s_X < \infty$ follow just as in the proof of Theorem 6.5.

Now consider the case $s_X = \infty$, assume $y = LKg$ for some $g \in S$, and fix r. Then for $y_N = L\Pi_N^X Kg$ as in (6.1), with $N \geq r$, we have

$$\|y - L\Pi_r^X L^{-1}y\|_Y \leq \|y - y_N\|_Y + \|y_N - L\Pi_r^X L^{-1}y_N\|_Y + \|L\Pi_r^X L^{-1}y_N - L\Pi_r^X L^{-1}y\|_Y.$$

Lemma 6.4 implies that the first term tends to zero as $N \to \infty$. For the second term, proceed as above and use $\Pi_r^X \Pi_N^X = \Pi_r^X$ (since $N \geq r$) to show

$$\|y_N - L\Pi_r^X L^{-1}y_N\|_Y \leq \sum_{k=r+1}^{N} \sigma_k\,|(g,f_k)_S|\,\|L\varphi_k\|_Y.$$

For the third term, first assume L is bounded. In this case,

$$\|L\Pi_r^X L^{-1}y_N - L\Pi_r^X L^{-1}y\|_Y \leq \|L\|\|\Pi_r^X\|\|L^{-1}y_N - L^{-1}y\|_X$$
$$= \|L\|\|\Pi_r^X\|\|\Pi_N^X Kg - Kg\|_X,$$

which converges to zero as $N \to \infty$, since r is fixed. If instead L^{-1} is bounded, then $L\Pi_r^X L^{-1}$ is bounded by Theorem 4.3 and so

$$\|L\Pi_r^X L^{-1}y_N - L\Pi_r^X L^{-1}y\|_Y \leq \|L\Pi_r^X L^{-1}\|\|y_N - y\|_Y,$$

which converges to zero as $N \to \infty$ by Lemma 6.4, again since r is fixed. Combining the above results gives

$$\|y - L\Pi_r^X L^{-1}y\|_Y \leq \sum_{k=r+1}^{\infty} \sigma_k\,|(g,f_k)_S|\,\|L\varphi_k\|_Y.$$

For convergence, we proceed as in the proof of Theorem 6.5. We have

$$\|y - L\Pi_r^X L^{-1}y\|_Y \leq \left(\sum_{k>r} |(g,f_k)_S|^2\right)^{1/2} \left(\sum_{k>r} \sigma_k^2\|L\varphi_k\|_Y^2\right)^{1/2}.$$

We know $\sum_{k>r} |(g,f_k)_S|^2$ goes to zero as r increases by Parseval's equality. Furthermore, Lemma 4.2 gives that $\sum_{k>r} \sigma_k^2\|L\varphi_k\|_Y^2$ goes to zero as r increases. This implies $L\Pi_r^X L^{-1}y \to y$ for each $y \in \mathcal{R}(K^Y)$. To show convergence for all $y \in Y$, we again use item 2 of Lemma 4.5 and the principle of uniform boundedness. □

We omit the proof of the next result, as it is similar to the proof of the previous result, Theorem 6.6. Note that in Theorem 6.6 the error converges to zero for a fixed $y \in \mathcal{R}(K^Y)$ without any additional assumptions. In this next result, if $s_X = \infty$ we need to require additional conditions to guarantee that the error converges to zero for a fixed $x \in \mathcal{R}(K)$; these conditions come from Lemma 4.2.

THEOREM 6.7. *Assume $s_X < \infty$ or either L or L^{-1} is bounded. If $x = Kg$ for some $g \in S$, then*

$$(6.5) \qquad \|x - L^{-1}\Pi_r^Y Lx\|_X \leq \sum_{k>r} \sigma_k|(g,f_k)_S|\,\|\varphi_k - L^{-1}\Pi_r^Y L\varphi_k\|_X.$$

If $s_X < \infty$, the error converges to zero as r increases. If $s_X = \infty$, then the error goes to zero as r increases when either (i) *L^{-1} is bounded and $\{\Pi_r^Y\}$ is uniformly bounded or* (ii) *$\{L^{-1}\Pi_r^Y L\}$ is uniformly bounded. Now assume $\{L^{-1}\Pi_r^Y L\}$ is uniformly bounded in operator norm whenever $s_X = \infty$. If the POD eigenvalues for the data $\{w_j\}$ are all nonzero, then $L^{-1}\Pi_r^Y Lx \to x$ for all $x \in X$.*

To be complete, we give an exact error formula and an error bound for approximations of elements in the range of K using the POD projection Π_r^X. This result gives an error bound for approximating each individual data snapshot in the discrete case.

THEOREM 6.8. *If $x = Kg$ for some $g \in S$, then*

$$(6.6) \qquad \|x - \Pi_r^X x\|_X = \left(\sum_{k>r} \sigma_k^2 |(g, f_k)_S|^2 \right)^{1/2} \leq \sigma_{r+1} \|g\|_S.$$

Also, in the discrete case, for each $\ell = 1, \ldots, s$ we have

$$(6.7) \qquad \|w_\ell - \Pi_r^X w_\ell\|_X \leq \gamma_\ell^{-1/2} \sigma_{r+1}.$$

Remark 6.9. The bound (6.7) was obtained in [25, Proposition 3.1] for $X = \mathbb{R}^n$ and $\gamma_\ell = 1$ for all ℓ. Recall the constants $\{\gamma_\ell\}$ are the positive weights in the definition of the POD operator K in the discrete case; see subsection 2.2.

Proof. Using the SVD of K gives

$$x - \Pi_r^X x = \sum_{k>r} \sigma_k (g, f_k)_S \varphi_k.$$

Since $\|x - \Pi_r^X x\|_X^2 = (x - \Pi_r^X x, x - \Pi_r^X x)_X$ and $\{\varphi_k\}$ is an orthonormal basis for X, we immediately obtain the exact error formula in (6.6). To obtain the error bound in (6.6), use $\sigma_k \leq \sigma_{r+1}$ for all $k > r$ and also Parseval's equality.

Next, in the discrete case we have $w_\ell = Kg_\ell$ for each $\ell = 1, \ldots, s$, where $g_\ell = \gamma_\ell^{-1} e_\ell$ and e_ℓ is the ℓth standard unit vector for $\mathbb{K}^s$, i.e., the ℓth entry of e_ℓ is one and all other entries are zero. The error bound (6.7) follows from $\|g_\ell\|_S = \gamma_\ell^{-1/2}$ and (6.6). □

In Theorem 6.8, note that the quantity γ_ℓ^{-1} appears in the error bound (6.7) for approximating the snapshot w_ℓ. However, in applications it is typical that each weight γ_ℓ tends to zero as the number s of snapshots increases. Next, we use the above results to prove various approximation error bounds for each individual snapshot w_ℓ in the discrete case that do not depend on γ_ℓ^{-1}. Here, the bounds are valid only if r is sufficiently large. We note that these types of error bounds have been *assumed* to hold in the literature; Iliescu and Wang made this type of assumption in [20, Assumption 3.2] (with $\gamma_\ell = s^{-1}$ for all ℓ) in their analysis of a POD reduced order model of the Navier–Stokes equations, and many others have followed their approach.

COROLLARY 6.10. *In the discrete case, if r is sufficiently large, then for each $\ell = 1, \ldots, s$ we have*

$$(6.8a) \qquad \|w_\ell - \Pi_r^X w_\ell\|_X^2 \leq \sigma_{r+1}^2,$$

$$(6.8b) \qquad \|Lw_\ell - \Pi_r^Y Lw_\ell\|_Y^2 \leq \sum_{k>r} \sigma_k^2 \|L\varphi_k - \Pi_r^Y L\varphi_k\|_Y^2,$$

$$(6.8c) \qquad \|Lw_\ell - L\Pi_r^X w_\ell\|_Y^2 \leq \sum_{k>r} \sigma_k^2 \|L\varphi_k\|_Y^2,$$

$$(6.8d) \qquad \|w_\ell - L^{-1}\Pi_r^Y Lw_\ell\|_X^2 \leq \sum_{k>r} \sigma_k^2 \|\varphi_k - L^{-1}\Pi_r^Y L\varphi_k\|_X^2.$$

Remark 6.11. For a given set of data $\{w_\ell\}_{\ell=1}^s$, we cannot rule out the possibility that r must be taken as large as possible (i.e., $r = s$) in Corollary 6.10.

Proof. We prove only (6.8b); the proofs of the remaining inequalities are similar. As in the proof of Theorem 6.8, we know $w_\ell = Kg_\ell$ for each $\ell = 1, \ldots, s$, where $g_\ell = \gamma_\ell^{-1} e_\ell$. Using the error bound (6.2) in Theorem 6.5, the Cauchy–Schwarz inequality on the sum, and Parseval's inequality gives

$$\|Lw_\ell - \Pi_r^Y Lw_\ell\|_Y^2 \le \|g_\ell - \Pi_r^S g_\ell\|_S^2 \sum_{k>r} \sigma_k^2 \|L\varphi_k - \Pi_r^Y L\varphi_k\|_Y^2,$$

where $\Pi_r^S : S \to S$ is the orthogonal projection onto $S_r := \mathrm{span}\{f_k\}_{k=1}^r$. Since $\{f_k\}_{k\ge 1}$ is an orthonormal basis for S, we know $\Pi_r^S g_\ell \to g_\ell$ for $\ell = 1, \ldots, s$. Since s is fixed, for all sufficiently large r we have $\|g_\ell - \Pi_r^S g_\ell\|_S \le 1$ for all $\ell = 1, \ldots, s$, and this completes the proof. $\quad\square$

7. More examples. We now consider a few additional examples. For all three examples we consider two separable Hilbert spaces, H and V, where V is a proper subset of H, and V is both continuously embedded[2] and dense in H. The linear operator L is a mapping between these two spaces.

For all three examples, we present results for the continuous case only. We assume we have the data $\{w_j\}_{j=1}^m \subset L^2(\mathcal{O}; H) \cap L^2(\mathcal{O}; V) = L^2(\mathcal{O}; V)$. Results for the discrete case can also be obtained using the theory in this work if desired.

The first two examples are from our previous work [40]. Due to the above assumption on the data, the POD operator K can be viewed as a mapping into H or a mapping into V. One can obtain the SVD of $K : S \to H$ or the SVD of $K : S \to V$, i.e., one can choose $X = H$ or $X = V$. The different choices for X give different POD singular values, POD singular vectors, POD modes, and POD projections. In [40], we considered both choices for X and four different POD projections between these spaces and gave exact expressions for the POD data approximation errors in the two different Hilbert space norms. We relate the notation and results for both the error formulas and pointwise convergence from the present work to [40]. We obtain better pointwise convergence results in this work. Also, $\mathcal{O}$ was only an interval in [40], but now we have $\mathcal{O}$ is an open subset of $\mathbb{R}^d$. For these first two examples, $Y_r = \mathrm{span}\{L\varphi_k\}$ and $\Pi_r^Y : Y \to Y$ is the orthogonal projection onto Y_r. Note this implies $\{\Pi_r^Y\}$ is uniformly bounded in operator norm.

For the third example, we consider a case where Π_r^Y is not an orthogonal projection. In particular, we take Π_r^Y to be a Ritz projection, as considered in [18, 37]. All of our results for this case are new.

7.1. Example 1. For the first example, consider the case where $X = H$, $Y = V$, and $L : H \to V$ is defined by $Lv = v$ for all $v \in \mathcal{D}(L) = V$. The operator L is clearly invertible, and $L^{-1} : V \to H$ is given by $L^{-1} v = v$ for all $v \in V$. Note that $L^{-1} : V \to H$ is bounded due to the continuous embedding assumption. Also, the inverse of a bounded operator is closed, so L is closed. Furthermore, the assumption on the data gives $\{w_j\} \subset L^2(\mathcal{O}; X)$ and $\{Lw_j\} \subset L^2(\mathcal{O}; Y)$. Thus, we know that both the main assumption and the L^{-1} assumption hold.

Since $X = H$ and each set of singular vectors of the POD operator $K : S \to H$ are an orthonormal basis, we know the POD modes $\{\varphi_k\}$ are an orthonormal basis for H. Note that $X_r = \mathrm{span}\{\varphi_k\}_{k=1}^r \subset H$, and $Y_r = \mathrm{span}\{L\varphi_k\}_{k=1}^r = \mathrm{span}\{\varphi_k\}_{k=1}^r \subset V$.

---

[2]That is, there exists a constant $C_V > 0$ such that $\|v\|_H \le C_V \|v\|_V$ for all $v \in V$.

Furthermore, the POD modes $\{\varphi_k\}$ may not be orthogonal in V. Also, the operator $K^Y = LK$ is simply the POD operator K viewed as a mapping from S to V. We take $\Pi_r^X : X \to X$ to be the orthogonal projection onto X_r and $\Pi_r^Y : Y \to Y$ to be the orthogonal projection onto Y_r.

In order to discuss the POD projections we pay special attention to the spaces under consideration. Since $V \subset H$, the projections can be considered as mappings from V to V or from H to H. The projections considered in this work are related to the projections P_r^H and P_r^V in [40, Definition 3.2] as follows:

- $\Pi_r^X : X \to X$ is equal to the orthogonal projection $P_r^H : H \to H$.
- $\Pi_r^Y : Y \to Y$ is equal to the orthogonal projection $P_r^V : V \to V$.
- $L\Pi_r^X L^{-1} : Y \to Y$ is equal to the operator $P_r^H : V \to V$.
- $L^{-1}\Pi_r^Y L : X \to X$ is equal to the operator $P_r^V : H \to H$.

Now that we have the relationships between the projections, we compare the results. The error formulas presented here in Theorem 5.4 are essentially the same as the results in [40]. Again, the primary difference here is that $\mathcal{O}$ is an open subset of $\mathbb{R}^d$ instead of an interval. The POD data approximation errors from Theorem 5.4 become the following:

$$(7.1) \qquad \sum_{j=1}^m \int_{\mathcal{O}} \|w_j(t) - P_r^H w_j(t)\|_V^2 dt = \sum_{k>r} \sigma_k^2 \|\varphi_k\|_V^2,$$

$$(7.2) \qquad \sum_{j=1}^m \int_{\mathcal{O}} \|w_j(t) - P_r^V w_j(t)\|_V^2 dt = \sum_{k>r} \sigma_k^2 \|\varphi_k - P_r^V \varphi_k\|_V^2,$$

$$(7.3) \qquad \sum_{j=1}^m \int_{\mathcal{O}} \|w_j(t) - P_r^V w_j(t)\|_H^2 dt = \sum_{k>r} \sigma_k^2 \|\varphi_k - P_r^V \varphi_k\|_H^2.$$

In this example, all three sums converge to zero as r increases.

A larger improvement from [40] can be seen in the results concerning pointwise convergence of POD projections. To illustrate, we give the following result.

PROPOSITION 7.1. *We have the following:*

1. $\|P_r^V y - y\|_V \to 0$ *for all* $y \in \mathcal{R}(K)$, *and for* $y = Kg$ *we have*

$$\|P_r^V y - y\|_V \leq \sum_{k>r} \sigma_k |(g, f_k)_S| \, \|P_r^V \varphi_k - \varphi_k\|_V.$$

2. *If the POD eigenvalues for* $\{w_j\} \subset L^2(\mathcal{O}; V)$ *are all nonzero, then* $P_r^V y \to y$ *in both* H *and* V *for all* $y \in V$.

3. $\|P_r^H y - y\|_V \to 0$ *for all* $y \in \mathcal{R}(K)$, *and for* $y = Kg$ *we have*

$$\|y - P_r^H y\|_V \leq \sum_{k>r} \sigma_k \, |(g, f_k)_S| \, \|\varphi_k\|_V.$$

4. $\|P_r^V x - x\|_H \to 0$ *for all* $x \in \mathcal{R}(K)$, *and for* $x = Kg$ *we have*

$$\|x - P_r^V x\|_H \leq \sum_{k>r} \sigma_k \, |(g, f_k)_S| \, \|\varphi_k - P_r^V \varphi_k\|_H.$$

Note that since Π_r^Y is orthogonal, items 1 and 2 follow from Theorem 6.5 and item 2 of Theorem 6.2. Items 3 and 4 can be obtained from Theorems 6.6 and 6.7 and the fact that L^{-1} is bounded.

The pointwise convergence results above are more complete and sharper than the results in [40, Proposition 5.5]. First, item 2 is shown in [40, Proposition 5.5] under the assumption that all the POD singular values for $\{w_j\} \subset L^2(\mathcal{O}; V)$ are nonzero; as discussed in section 2 this is a more restrictive assumption than the POD eigenvalues all being nonzero, as is required above. Next, the convergence result in item 3 is shown in [40, Proposition 5.5]; however, the error bound in item 3 is new. Also, items 1 and 4 are completely new.

For item 3, we note that an error bound was given in the proof of [40, Proposition 5.5]. However, that error bound does not converge to zero as fast as the error bound given in Theorem 6.6. Specifically, the error bound in [40] is a constant multiple of $(\sum_{k>r} |(g, f_k)_S|^2)^{1/2}$. However, the error bound in item 3 can be bounded above by

$$\|y - P_r^H y\|_V \le \left(\sum_{k>r} |(g, f_k)_S|^2 \right)^{1/2} \left(\sum_{k>r} \sigma_k^2 \|\varphi_k\|_V^2 \right)^{1/2},$$

and both terms in parentheses tend to zero as r increases by Parseval's equality and Lemma 4.2 (see the proof of Theorem 6.6). Therefore, the error bound in item 3 is an improvement over the error bound in [40].

Finally, we consider boundedness of the nonorthogonal POD projections $P_r^H : V \to V$ and $P_r^V : H \to H$. For each fixed r, we showed in [40, Lemma 3.3] that $P_r^H : V \to V$ is bounded. We did not consider the boundedness of $P_r^V : H \to H$ in [40]. Below, we use Theorem 4.3 to show $P_r^H : V \to V$ is bounded and also give a condition guaranteeing $P_r^V : H \to H$ has a bounded extension. However, we still do not know if these nonorthogonal POD projections are uniformly bounded in operator norm.

Define the linear operator $A : \mathcal{D}(A) \subset H \to H$ by

$$(Au, v)_H = (u, v)_V$$

for all $u \in \mathcal{D}(A)$ and $v \in V$ (see, e.g., [42, section II.2]). We know A is closed. Now we apply this to our example. For all $x \in \mathcal{D}(L) = V$ and $y \in \mathcal{D}(L^*)$ we have

$$(x, L^* y)_H = (Lx, y)_V = (x, y)_V$$
$$\Rightarrow (L^* y, x)_H = (y, x)_V.$$

Thus, $L^* = A$ and $\mathcal{D}(L^*) = \mathcal{D}(A)$. For PDE solution data we often have $\{Aw_j\} \subset L^2(\mathcal{O}; H)$ for each j; see [42] for examples. In this case, since $\varphi_k = \sigma_k^{-1} K f_k$ we can use the Bochner integral result in Theorem 2.2 to show $\varphi_k \in \mathcal{D}(A)$ whenever $\sigma_k > 0$.

Therefore, since L^{-1} is bounded, items 1 and 4 of Theorem 4.3 give the following result.

PROPOSITION 7.2. *Let r be fixed. The operator $P_r^H : V \to V$ is bounded, and if $\{Aw_j\}_{j=1}^m \subset L^2(\mathcal{O}; H)$, then the operator $P_r^V : H \to H$ can be extended to a bounded operator.*

7.2. Example 2. Next, consider the case where $X = V, Y = H$, and $L : V \to H$ is defined by by $Lv = v$ for all $v \in V$. Then $L^{-1} : H \to V$ is given by $L^{-1}v = v$ for all $v \in \mathcal{D}(L^{-1}) = V$. Note that in this case L is bounded by the continuous embedding property. Again, the assumption on the data gives $\{w_j\} \subset L^2(\mathcal{O}; X)$ and $\{Lw_j\} \subset L^2(\mathcal{O}; Y)$. Therefore, the main assumption and the L^{-1} assumption hold.

Since $X = V$, in this example the POD modes $\{\varphi_k\}$ are an orthonormal basis for V. We have $X_r = \text{span}\{\varphi_k\}_{k=1}^r \subset V$, and $Y_r = \text{span}\{L\varphi_k\}_{k=1}^r = \text{span}\{\varphi_k\}_{k=1}^r \subset H$.

The POD modes $\{\varphi_k\}$ may not be orthogonal in H. The operator $K^Y = LK$ is the POD operator $K : S \to H$. As in Example 1, $\Pi_r^X : X \to X$ is the orthogonal projection onto X_r, and $\Pi_r^Y : Y \to Y$ is the orthogonal projection onto Y_r.

The projections in this work are related to the projections Q_r^H and Q_r^V from [40, Definition 3.2] as follows:

- $\Pi_r^X : X \to X$ is equal to the orthogonal projection $Q_r^V : V \to V$.
- $\Pi_r^Y : Y \to Y$ is equal to the orthogonal projection $Q_r^H : H \to H$.
- $L\Pi_r^X L^{-1} : Y \to Y$ is equal to the operator $Q_r^V : H \to H$.
- $L^{-1}\Pi_r^Y L : X \to X$ is equal to the operator $Q_r^H : V \to V$.

As before, the main data approximation error results in Theorem 5.4 become

$$\sum_{j=1}^{m} \int_{\mathcal{O}} \left\| w_j(t) - Q_r^V w_j(t) \right\|_H^2 dt = \sum_{k>r} \sigma_k^2 \|\varphi_k\|_H^2,$$

$$\sum_{j=1}^{m} \int_{\mathcal{O}} \left\| w_j(t) - Q_r^H w_j(t) \right\|_H^2 dt = \sum_{k>r} \sigma_k^2 \|\varphi_k - Q_r^H \varphi_k\|_H^2,$$

$$\sum_{j=1}^{m} \int_{\mathcal{O}} \| w_j(t) - Q_r^H w_j(t) \|_V^2 dt = \sum_{k>r} \sigma_k^2 \|\varphi_k - Q_r^H \varphi_k\|_V^2.$$

Here, the first two sums converge to zero as r increases. However, we cannot show convergence of the last sum. This is because we do not know L^{-1} is bounded or $\{Q_r^H\}$ is uniformly bounded as a family of operators mapping V to V. As before, the only improvement here compared to [40] is that $\mathcal{O}$ is not restricted to be an interval.

We also have the following pointwise convergence results.

PROPOSITION 7.3. *As r increases we have*
1. *$\|Q_r^H y - y\|_H \to 0$ for all $y \in H$, and for $y = Kg$ we have*

$$\|Q_r^H y - y\|_H \leq \sum_{k>r} \sigma_k |(g, f_k)_S| \, \|Q_r^H \varphi_k - \varphi_k\|_H.$$

2. *$\|Q_r^V y - y\|_H \to 0$ for all $y \in V$, and for $y = Kg$ we have*

$$\|Q_r^V y - y\|_H \leq \sum_{k>r} \sigma_k |(g, f_k)_S| \, \|\varphi_k\|_H.$$

3. *For $x = Kg$ we have*

$$\|Q_r^H x - x\|_V \leq \sum_{k>r} \sigma_k |(g, f_k)_S| \|Q_r^H \varphi_k - \varphi_k\|_V.$$

If also $s_X < \infty$ or L^{-1} is bounded, then the error goes to zero as r increases.

Since L is bounded, item 1 follows from item 1 of Theorem 6.1 and also Theorem 6.5. Item 2 can be obtained from Theorem 6.6, using L is bounded. Theorem 6.7 gives item 3; note that we cannot guarantee convergence of the error without the extra assumptions since we only know L is bounded.

Again, these results improve on the results in [40, Proposition 5.5]. All of the error bounds are new. The convergence result in item 2 was not stated in [40], but it follows directly from the continuous embedding and $\|Q_r^V y - y\|_V \to 0$ for all $y \in V$. The convergence result in item 1 was given in [40, Proposition 5.5]; however, we made

the assumption that all the POD singular values for $\{w_j\} \subset L^2(\mathcal{O}; V)$ are nonzero. Here, we proved the convergence result in item 1 without that assumption.

Next, we use the technique from subsection 7.1 to determine the boundedness of the nonorthogonal POD projections $Q_r^H : V \to V$ and $Q_r^V : H \to H$. For this example, we have $A = L^{-*} = (L^{-1})^*$. Therefore, if $\{Aw_j\} \subset L^2(\mathcal{O}; H)$, then we have $\{\varphi_k\} \subset \mathcal{D}(L^{-*})$, just as in subsection 7.1. Since L is bounded, items 2 and 3 of Theorem 4.3 give the following result.

PROPOSITION 7.4. *Let r be fixed. The operator $Q_r^H : V \to V$ is bounded, and if $\{Aw_j\}_{j=1}^m \subset L^2(\mathcal{O}; H)$, then the operator $Q_r^V : H \to H$ can be extended to a bounded operator on H.*

7.3. Example 3. In order to demonstrate the usefulness of considering Π_r^Y as a nonorthogonal projection, we consider the case of a Ritz projection as presented in [18, 37].

Consider the situation from Example 1 in subsection 7.1: we have $X = H$, $Y = V$, and $L : X \to Y$ is defined by $Lv = v$ for all $v \in \mathcal{D}(L) = Y$. Assume we have a continuous elliptic sesquilinear form[3] $a : V \times V \to \mathbb{K}$. Define the projection $P_r^V : V \to V$ onto $V_r := Y_r = \text{span}\{L\varphi_k\} = \text{span}\{\varphi_k\} \subset V$ as follows: let $P_r^V u := u_r \in V_r$ be the unique solution of

$$a(u_r, v_r) = a(u, v_r) \quad \text{for all } v_r \in V_r.$$

The existence and uniqueness of such a solution are guaranteed by the Lax–Milgram theorem. We take $\Pi_r^Y = P_r^V$.

Note that the main difference between this example and Example 1 is that the projection P_r^V is not the same. However, for this example it can be checked that the family of projections, $\{\Pi_r^Y\}$, is uniformly bounded. Therefore, the same pointwise convergence results and error formulas from subsection 7.1 hold for this example with $P_r^V : V \to V$ defined as above. We note that these pointwise convergence results and error formulas are all new. Bounds on the POD data approximation errors can be found in Lemma 3.4 in [18] and Lemma 2.9 in [37] in the discrete case; however, we have the exact formulas (7.1)–(7.3) for the POD data approximation errors in the continuous case. Again, analogous error formulas can be derived for the discrete case using our results.

8. Conclusions. We proved new generalized error formulas for POD data approximation errors for both the discrete and continuous cases. We also showed convergence of these errors under certain conditions and obtained new pointwise convergence results for POD projections. We demonstrated the application of our results to several example problems. We leave the application of these results to the numerical analysis of POD model order reduction methods for PDEs to be considered elsewhere.

Some open questions remain. When L^{-1} is unbounded, we had to assume uniform boundedness of the POD projections $\{L^{-1}\Pi_r^Y L\}$ to show that the error formula in (5.7) converges to zero as r increases. We do not know if there is a simpler condition that yields convergence of the approximation error. If L or L^{-1} is unbounded, we also do not know if the POD projections $\{L\Pi_r^X L^{-1}\}$ and $\{L^{-1}\Pi_r^Y L\}$ are uniformly bounded. Both of these issues have been discussed in the context of Example 2 in subsection 7.2 in [5, 40]. The second issue has also been discussed in the context of

---

[3]That is, there exists constants $C_a, c_a > 0$ such that $|a(u,v)| \leq C_a \|u\|_V \|v\|_V$ and $c_a \|u\|_V^2 \leq \text{Re}\, a(u,u)$ for all $u, v \in V$.

Example 1 in subsection 7.1 in [23, 46]; in these works, the H^1 stability of the L^2 POD projection is of interest.

Appendix A. Optimality of discrete and continuous POD. To be complete, we present a brief proof of the optimality of POD for low rank data approximation in both the discrete and continuous cases. Our problem statement and proof strongly rely on ideas from [43] and [7].

POD optimality problem. Let X be a separable Hilbert space, and let $S = \mathbb{K}_\Gamma^s$ in the discrete case or $S = L^2(\mathcal{O}; \mathbb{K}^m)$ in the continuous case, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$; see section 2 for details. Suppose we have given data $\{w_j\}_{j=1}^s \subset X$ in the discrete case or $\{w_j\}_{j=1}^m \subset L^2(\mathcal{O}; X)$ in the continuous case. The POD optimality problem is to find coefficients $\{a_k\} \subset \mathbb{K}$ and basis elements $\{s_k\} \subset S$ and $\{\eta_k\} \subset X$ so that the rth order approximations

$$w_j^r = \sum_{k=1}^r a_k s_{k,j} \eta_k \qquad \text{for } j = 1, \ldots, s \text{ (discrete case),}$$

$$w_j^r(t) = \sum_{k=1}^r a_k s_{k,j}(t) \eta_k \qquad \text{for } j = 1, \ldots, m \text{ (continuous case),}$$

minimize the data approximation error

$$E_r(w^r) = \sum_{j=1}^s \gamma_j \|w_j - w_j^r\|_X^2 \qquad \text{(discrete case),}$$

$$E_r(w^r) = \sum_{j=1}^m \int_{\mathcal{O}} \|w_j(t) - w_j^r(t)\|_X^2 \, dt \qquad \text{(continuous case).}$$

Remark A.1. In many papers on POD, the basis elements $\{\eta_k\} \subset X$ are required to be orthonormal, and w_j^r is also required to equal the orthogonal projection of w_j onto $\text{span}\{\eta_k\}_{k=1}^r$. Therefore, the POD problem above allows more general approximations. The final result is the same.

Notation. For given data $\{y_j\}_{j=1}^s \subset X$ in the discrete case or $\{y_j\}_{j=1}^m \subset L^2(\mathcal{O}; X)$ in the continuous case, we let $K(y) : S \to X$ denote the POD operator for the data and we let $K^*(y) : X \to S$ denote the Hilbert adjoint operator of $K(y)$.

The proof of the next result follows directly from definitions and is omitted.

LEMMA A.2. *If the data is given by*

$$y_j = \sum_{k=1}^p \alpha_k s_{k,j} \eta_k \quad \text{(discrete case)}, \quad y_j(t) = \sum_{k=1}^p \alpha_k s_{k,j}(t) \eta_k \quad \text{(continuous case)}$$

for each j with $\{\alpha_k\} \subset \mathbb{K}$, $\{s_k\} \subset S$, and $\{\eta_k\} \subset X$, then the POD operator $K(y) : S \to X$ is given by

$$K(y)f = \sum_{k=1}^p \alpha_k (f, \overline{s_k})_S \eta_k, \quad f \in S.$$

Next, we present the discrete version of the Hilbert–Schmidt norm result for a continuous POD operator in Lemma 4.1.

LEMMA A.3. *For given data $\{y_j\}_{j=1}^s \subset X$ in the discrete case, the Hilbert–Schmidt norm of the POD operator $K(y) : S \to X$ is given by*

$$\|K(y)\|_{\mathrm{HS}(S,X)}^2 = \sum_{j=1}^s \gamma_j \|y_j\|_X^2.$$

Proof. Let $\{\xi_k\}_{k\geq 1}$ be an orthonormal basis for X. We have

$$\|K(y)\|_{\mathrm{HS}(S,X)}^2 = \|K^*(y)\|_{\mathrm{HS}(X,S)}^2 = \sum_{k\geq 1} \|K^*(y)\xi_k\|_S^2 = \sum_{k\geq 1} \sum_{j=1}^s \gamma_j |(\xi_k, y_j)_X|^2$$

$$= \sum_{j=1}^s \gamma_j \sum_{k\geq 1} |(\xi_k, y_j)_X|^2 = \sum_{j=1}^s \gamma_j \|y_j\|_X^2$$

by Parseval's equality.   □

Now we prove the main optimality result. We rely on the fact that the rank r truncated SVD of $K(y)$ is the optimal rank r approximation to $K(y)$ in the Hilbert–Schmidt norm; see, e.g., [12, section III.7, Theorem 7.1].

THEOREM A.4. *Let $\{w_j\}_{j=1}^s \subset X$ in the discrete case or $\{w_j\}_{j=1}^m \subset L^2(\mathcal{O}; X)$ in the continuous case be given data, and let $\{\sigma_i, f_i, \varphi_i\} \subset \mathbb{R} \times S \times X$ be the ordered singular values of $K(w) : S \to X$ and the corresponding orthonormal bases of singular vectors. A solution of the POD problem is given by $\{w_j^r\}_{j=1}^s \subset X$ in the discrete case or $\{w_j^r\}_{j=1}^m \subset L^2(\mathcal{O}; X)$ in the continuous case, where*

$$w_j^r = \sum_{k=1}^r \sigma_k \overline{f_{k,j}} \varphi_k = \sum_{k=1}^r (w_j, \varphi_k)_X \varphi_k \qquad \text{(discrete case)},$$

$$w_j^r(t) = \sum_{k=1}^r \sigma_k \overline{f_{k,j}(t)} \varphi_k = \sum_{k=1}^r (w_j(t), \varphi_k)_X \varphi_k \qquad \text{(continuous case)}.$$

The minimum approximation error is given by

$$E_r^{\min} := E_r(w^r) = \sum_{k>r} \sigma_k^2 < \infty,$$

and $E_r^{\min} \to 0$ as r increases.

Proof. We first assume $r \leq s_X$ so that $\sigma_k > 0$ for $k = 1, \ldots, r$.

First, the equivalence of the two expressions for w_j^r comes from $K^*(w)\varphi_k = \sigma_k f_k$, $\sigma_k > 0$ for $k = 1, \ldots, r$, and the formulas for $K^*(w)$. Also, for $g \in S$, Lemma A.2 implies

$$K(w^r)g = \sum_{k=1}^r \sigma_k(g, f_k)_S \varphi_k = K_r(w)g.$$

Therefore, $K(w^r) = K_r(w)$, where $K_r(w) : S \to X$ is the rth order truncated SVD of the POD operator $K(w) : S \to X$.

Next, by the Hilbert–Schmidt norm results Lemma 4.1 and Lemma A.3 and since the POD operator is linear in the data we have

$$E_r(w^r) = \|K(w - w^r)\|_{\mathrm{HS}}^2 = \|K(w) - K(w^r)\|_{\mathrm{HS}}^2 = \|K(w) - K_r(w)\|_{\mathrm{HS}}^2 = \sum_{k>r} \sigma_k^2.$$

Also, since $\|K(w)\|_{\mathrm{HS}} = \sum_{k\geq 1} \sigma_k^2 < \infty$, we have $\sum_{k>r} \sigma_k^2 \to 0$ as r increases.

Now we show that this is the smallest value possible for the error. Let coefficients $\{a_k\} \subset \mathbb{K}$ and basis elements $\{s_k\} \subset S$ and $\{\eta_k\} \subset X$ be given, and define the rth order approximation

$$z_j^r = \sum_{k=1}^{r} \alpha_k s_{k,j} \eta_k \quad \text{(discrete case)}, \quad z_j^r(t) = \sum_{k=1}^{p} \alpha_k s_{k,j}(t) \eta_k \quad \text{(continuous case)}.$$

By Lemma A.2, $K(z^r)$ has rank at most r. Therefore, we have

$$E_r(z^r) = \|K(w - z^r)\|_{\mathrm{HS}}^2 = \|K(w) - K(z^r)\|_{\mathrm{HS}}^2 \geq \sum_{k>r} \sigma_k^2.$$

Next, if $s_X < \infty$, then the result is true for $r = s_X$. Therefore, we have $w_j = w_j^{s_X}$ for all j, and this proves the result for $r > s_X$. □

<div align="center">REFERENCES</div>

[1] A. ALLA, M. FALCONE, AND S. VOLKWEIN, *Error analysis for POD approximations of infinite horizon problems via the dynamic programming approach*, SIAM J. Control Optim., 55 (2017), pp. 3091–3115, https://doi.org/10.1137/15M1039596.

[2] J.-P. AUBIN, *Applied Functional Analysis*, 2nd ed., Pure Appl. Math. (N. Y.), Wiley-Interscience, New York, 2000, https://doi.org/10.1002/9781118032725.

[3] A. V. BALAKRISHNAN, *Applied Functional Analysis*, Springer-Verlag, New York, 1976.

[4] M. BERGMANN, L. CORDIER, AND J.-P. BRANCHER, *Optimal rotary control of the cylinder wake using proper orthogonal decomposition reduced-order model*, Phys. Fluids, 17 (2005), 097101, https://doi.org/10.1063/1.2033624.

[5] D. CHAPELLE, A. GARIAH, AND J. SAINTE-MARIE, *Galerkin approximation with proper orthogonal decomposition: New error estimates and illustrative examples*, ESAIM Math. Model. Numer. Anal., 46 (2012), pp. 731–757, https://doi.org/10.1051/m2an/2011053.

[6] P. G. CIARLET, *Linear and Nonlinear Functional Analysis with Applications*, SIAM, Philadelphia, 2013.

[7] S. M. DJOUADI, *On the optimality of the proper orthogonal decomposition and balanced truncation*, in Proceedings of the 47th IEEE Conference on Decision and Control, 2008, pp. 4221–4226.

[8] N. DUNFORD AND J. T. SCHWARTZ, *Linear Operators. I. General Theory*, Pure Appl. Math. 7, Interscience Publishers, New York, 1958.

[9] F. G. EROGLU, S. KAYA, AND L. G. REBHOLZ, *A modular regularized variational multiscale proper orthogonal decomposition for incompressible flows*, Comput. Methods Appl. Mech. Engrg., 325 (2017), pp. 350–368, https://doi.org/10.1016/j.cma.2017.07.017.

[10] S. GIERE, T. ILIESCU, V. JOHN, AND D. WELLS, *SUPG reduced order models for convection-dominated convection-diffusion-reaction equations*, Comput. Methods Appl. Mech. Engrg., 289 (2015), pp. 454–474, https://doi.org/10.1016/j.cma.2015.01.020.

[11] I. GOHBERG, S. GOLDBERG, AND M. A. KAASHOEK, *Classes of Linear Operators. Vol. I*, Oper. Theory Adv. Appl. 49, Birkhäuser Verlag, Basel, 1990, https://doi.org/10.1007/978-3-0348-7509-7.

[12] I. C. GOHBERG AND M. G. KREĬN, *Introduction to the Theory of Linear Nonselfadjoint Operators*, AMS, Providence, RI, 1969.

[13] C. GRÄSSLE AND M. HINZE, *POD reduced-order modeling for evolution equations utilizing arbitrary finite element discretizations*, Adv. Comput. Math., 44 (2018), pp. 1941–1978, https://doi.org/10.1007/s10444-018-9620-x.

[14] M. GUBISCH AND S. VOLKWEIN, *Proper orthogonal decomposition for linear-quadratic optimal control*, in Model Reduction and Approximation, Comput. Sci. Eng. 15, SIAM, Philadelphia, 2017, pp. 3–63, https://doi.org/10.1137/1.9781611974829.ch1.

[15] M. GUNZBURGER, T. ILIESCU, AND M. SCHNEIER, *A Leray regularized ensemble-proper orthogonal decomposition method for parameterized convection-dominated flows*, IMA J. Numer. Anal., 40 (2020), pp. 886–913, https://doi.org/10.1093/imanum/dry094.

[16] M. GUNZBURGER, N. JIANG, AND M. SCHNEIER, *An ensemble-proper orthogonal decomposition method for the nonstationary Navier-Stokes equations*, SIAM J. Numer. Anal., 55 (2017), pp. 286–304, https://doi.org/10.1137/16M1056444.

[17] P. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, 2nd ed., Cambridge Monogr. Mech., Cambridge University Press, Cambridge, UK, 2012, https://doi.org/10.1017/CBO9780511919701.

[18] T. Iliescu and Z. Wang, *Variational multiscale proper orthogonal decomposition: Convection-dominated convection-diffusion-reaction equations*, Math. Comp., 82 (2013), pp. 1357–1378, https://doi.org/10.1090/S0025-5718-2013-02683-X.

[19] T. Iliescu and Z. Wang, *Are the snapshot difference quotients needed in the proper orthogonal decomposition?*, SIAM J. Sci. Comput., 36 (2014), pp. A1221–A1250, https://doi.org/10.1137/130925141.

[20] T. Iliescu and Z. Wang, *Variational multiscale proper orthogonal decomposition: Navier-Stokes equations*, Numer. Methods Partial Differential Equations, 30 (2014), pp. 641–663, https://doi.org/10.1002/num.21835.

[21] B. Jin and Z. Zhou, *An analysis of Galerkin proper orthogonal decomposition for subdiffusion*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 89–113, https://doi.org/10.1051/m2an/2016017.

[22] T. Kato, *Perturbation Theory for Linear Operators*, Classics in Math., Springer-Verlag, Berlin, 1995.

[23] K. Kean and M. Schneier, *Error analysis of supremizer pressure recovery for POD based reduced order models of the time-dependent Navier–Stokes equations*, SIAM J. Numer. Anal., 58 (2019), pp. 2235–2264.

[24] T. Kostova-Vassilevska and G. M. Oxberry, *Model reduction of dynamical systems by proper orthogonal decomposition: Error bounds and comparison of methods using snapshots from the solution and the time derivatives*, J. Comput. Appl. Math., 330 (2018), pp. 553–573, https://doi.org/10.1016/j.cam.2017.09.001.

[25] T. Kostova-Vassilevska and G. M. Oxberry, *Model reduction of dynamical systems by proper orthogonal decomposition: Error bounds and comparison of methods using snapshots from the solution and the time derivatives*, J. Comput. Appl. Math., 330 (2018), pp. 553–573, https://doi.org/10.1016/j.cam.2017.09.001.

[26] K. Kunisch and S. Volkwein, *Galerkin proper orthogonal decomposition methods for parabolic problems*, Numer. Math., 90 (2001), pp. 117–148, https://doi.org/10.1007/s002110100282.

[27] K. Kunisch and S. Volkwein, *Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics*, SIAM J. Numer. Anal., 40 (2002), pp. 492–515, https://doi.org/10.1137/S0036142900382612.

[28] P. D. Lax, *Functional Analysis*, Pure Appl. Math. (N. Y.), Wiley-Interscience, New York, 2002.

[29] R. Li, Q. Wu, and S. Zhu, *Proper orthogonal decomposition with SUPG-stabilized isogeometric analysis for reduced order modelling of unsteady convection-dominated convection-diffusion-reaction problems*, J. Comput. Phys., 387 (2019), pp. 280–302, https://doi.org/10.1016/j.jcp.2019.02.051.

[30] Y. C. Liang, H. P. Lee, S. P. Lim, W. Z. Lin, K. H. Lee, and C. G. Wu, *Proper orthogonal decomposition and its applications*. I. *Theory*, J. Sound Vibration, 252 (2002), pp. 527–544, https://doi.org/10.1006/jsvi.2001.4041.

[31] M. Miklavčič, *Applied Functional Analysis and Partial Differential Equations*, World Scientific, River Edge, NJ, 1998, https://doi.org/10.1142/9789812796233.

[32] M. Mohebujjaman, L. G. Rebholz, X. Xie, and T. Iliescu, *Energy balance and mass conservation in reduced order models of fluid flows*, J. Comput. Phys., 346 (2017), pp. 262–277, https://doi.org/10.1016/j.jcp.2017.06.019.

[33] V. B. Nguyen, H.-S. Dou, K. Willcox, and B.-C. Khoo, *Model order reduction for reacting flows: Laminar Gaussian flame applications*, in 30th International Symposium on Shock Waves 1, Springer, New York, 2017, pp. 337–343, https://doi.org/10.1007/978-3-319-46213-4_57.

[34] A. Quarteroni, A. Manzoni, and F. Negri, *Reduced Basis Methods for Partial Differential Equations*, Unitext 92, Springer, Cham, 2016, https://doi.org/10.1007/978-3-319-15431-2.

[35] M. Reed and B. Simon, *Methods of Modern Mathematical Physics* I: *Functional Analysis*, 2nd ed., Academic Press, New York, 1980.

[36] C. W. Rowley, *Model reduction for fluids, using balanced proper orthogonal decomposition*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 15 (2005), pp. 997–1013, https://doi.org/10.1142/S0218127405012429.

[37] S. Rubino, *A streamline derivative POD-ROM for advection-diffusion-reaction equations*, ESAIM Proc. Surv., 64 (2018), pp. 121–136, https://doi.org/10.1051/proc/201864121.

[38] J. Shen, J. R. Singler, and Y. Zhang, *HDG-POD reduced order model of the heat equation*, J. Comput. Appl. Math., 362 (2019), pp. 663–679, https://doi.org/10.1016/j.cam.2018.09.031.

[39] J. R. Singler, *Convergent snapshot algorithms for infinite-dimensional Lyapunov equations*, IMA J. Numer. Anal., 31 (2011), pp. 1468–1496, https://doi.org/10.1093/imanum/drq028.

[40] J. R. Singler, *New POD error expressions, error bounds, and asymptotic results for reduced order models of parabolic PDEs*, SIAM J. Numer. Anal., 52 (2014), pp. 852–876, https://doi.org/10.1137/120886947.

[41] J. R. Singler, *Optimality of balanced proper orthogonal decomposition for data reconstruction II: Further approximation results*, J. Math. Anal. Appl., 421 (2015), pp. 1006–1020, https://doi.org/10.1016/j.jmaa.2014.07.059.

[42] R. Temam, *Infinite-dimensional Dynamical Systems in Mechanics and Physics*, 2nd ed., Appl. Math. Sci. 68, Springer-Verlag, New York, 1997, https://doi.org/10.1007/978-1-4612-0645-3.

[43] S. Volkwein, *Interpretation of proper orthogonal decomposition as singular value decomposition and HJB-based feedback design*, in Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems, 2004.

[44] Z. Wang, *Nonlinear model reduction based on the finite element method with interpolated coefficients: Semilinear parabolic equations*, Numer. Methods Partial Differential Equations, 31 (2015), pp. 1713–1741, https://doi.org/10.1002/num.21961.

[45] K. Willcox and J. Peraire, *Balanced model reduction via the proper orthogonal decomposition*, AIAA J., 40 (2002), pp. 2323–2330, https://doi.org/10.2514/2.1570.

[46] X. Xie, D. Wells, Z. Wang, and T. Iliescu, *Numerical analysis of the Leray reduced order model*, J. Comput. Appl. Math., 328 (2018), pp. 12–29, https://doi.org/10.1016/j.cam.2017.06.026.

[47] C. Zerfas, L. G. Rebholz, M. Schneier, and T. Iliescu, *Continuous data assimilation reduced order models of fluid flow*, Comput. Methods Appl. Mech. Engrg., 357 (2019), p. 112596, https://doi.org/10.1016/j.cma.2019.112596.

[48] S. Zhu, L. Dedè, and A. Quarteroni, *Isogeometric analysis and proper orthogonal decomposition for parabolic problems*, Numer. Math., 135 (2017), pp. 333–370, https://doi.org/10.1007/s00211-016-0802-5.