Missouri University of Science and Technology
Scholars' Mine

Mathematics and Statistics Faculty Research & Creative Works

Mathematics and Statistics

14 Apr 2014

New POD Error Expressions, Error Bounds, and Asymptotic Results for Reduced Order Model of Parabolic PDEs

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

Follow this and additional works at: https://scholarsmine.mst.edu/math_stat_facwork

Part of the Mathematics Commons, Numerical Analysis and Scientific Computing Commons, and the Statistics and Probability Commons

Recommended Citation

J. R. Singler, "New POD Error Expressions, Error Bounds, and Asymptotic Results for Reduced Order Model of Parabolic PDEs," *SIAM Journal on Numerical Analysis*, vol. 52, no. 2, pp. 852-876, Society for Industrial and Applied Mathematics (SIAM), Apr 2014.
The definitive version is available at https://doi.org/10.1137/120886947

This Article - Journal is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Mathematics and Statistics Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

SIAM J. NUMER. ANAL.
Vol. 52, No. 2, pp. 852–876
© 2014 Society for Industrial and Applied Mathematics

NEW POD ERROR EXPRESSIONS, ERROR BOUNDS, AND ASYMPTOTIC RESULTS FOR REDUCED ORDER MODELS OF PARABOLIC PDEs*

JOHN R. SINGLER†

Abstract. The derivations of existing error bounds for reduced order models of time varying partial differential equations (PDEs) constructed using proper orthogonal decomposition (POD) have relied on bounding the error between the POD data and various POD projections of that data. Furthermore, the asymptotic behavior of the model reduction error bounds depends on the asymptotic behavior of the POD data approximation error bounds. We consider time varying data taking values in two different Hilbert spaces H and V, with $V \subset H$, and prove exact expressions for the POD data approximation errors considering four different POD projections and the two different Hilbert space error norms. Furthermore, the exact error expressions can be computed using only the POD eigenvalues and modes, and we prove the errors converge to zero as the number of POD modes increases. We consider the POD error estimation approaches of Kunisch and Volkwein [*SIAM J. Numer. Anal.*, 40 (2002), pp. 492–515] and Chapelle, Gariah, and Sainte-Marie [*ESAIM Math. Model. Numer. Anal.*, 46 (2012), pp. 731–757] and apply our results to derive new POD model reduction error bounds and convergence results for the two-dimensional Navier–Stokes equations. We prove the new error bounds tend to zero as the number of POD modes increases for POD space $X = H$ in both approaches; the asymptotic behavior of existing error bounds was unknown for this case. Also, for $X = H$, we prove one new error bound tends to zero without requiring time derivative data in the POD data set.

Key words. proper orthogonal decomposition, reduced order models, parabolic equations, Navier–Stokes equations, data approximation

AMS subject classifications. 65M, 41A25

DOI. 10.1137/120886947

1. Introduction.

Proper orthogonal decomposition (POD) is a well-known and widely used model reduction method for partial differential equations (PDEs) and other dynamical systems. POD takes simulation data or experimental data and produces modes that can be used to reduce the model via a Galerkin projection. These reduced order POD models are not well understood—when initial conditions or parameters are changed in the reduced model, the solution of the reduced model can be either surprisingly accurate or completely unrelated to the solution of the full model. Understanding and predicting the accuracy of the reduced model is an extremely important problem.

As a first step toward this goal, Kunisch and Volkwein in [20, 21] proved error bounds for the POD reduced order model of various linear and nonlinear parabolic PDEs under the assumption that the initial conditions and parameters in the reduced model were not varied from the original PDE model. Since these works, other researchers have considered various scenarios (different PDEs, numerical methods, etc.) and modified POD model reduction schemes and have proved related error bounds; see, e.g., [2, 3, 4, 8, 13, 15, 18, 19, 23, 24, 25, 26, 27]. Also, see the many references in these works for more information about POD and applications of POD to many different fields.

*Received by the editors August 3, 2012; accepted for publication (in revised form) January 6, 2014; published electronically April 8, 2014.
http://www.siam.org/journals/sinum/52-2/88694.html

†Department of Mathematics and Statistics, Missouri University of Science and Technology, Rolla, MO 65409-0020 (singlerj@mst.edu).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

A distinguishing feature of parabolic PDEs is that the solution is square integrable in time when taking values in two different Hilbert spaces, H and V. For example, for the diffusion equation with zero Dirichlet boundary conditions on a bounded domain $\Omega \subset \mathbb{R}^n$, the usual choices are $H = L^2(\Omega)$ and $V = H_0^1(\Omega)$. The inner product of two functions f and g on H is the integral of the product fg over Ω, while the inner product on V is the integral of the dot product of the gradients $\nabla f \cdot \nabla g$ over Ω. The POD eigenvalues and modes depend on the Hilbert space X chosen, and therefore different reduced order models arise for $X = H$ and $X = V$.

In [21], Kunisch and Volkwein gave an error bound for the two-dimensional (2D) Navier–Stokes equations for the case $X = V$ that converges to zero as the time step is refined and the number of POD modes is increased. However, they did not obtain a similar asymptotic result for $X = H$ due to a factor in the error bound that tends to infinity as the number of POD modes increases. Also, in many of the recent POD error estimate works listed above, the case $X = H$ is not considered even though numerical results in many works indicate the error converges to zero for this case.

Furthermore, in order to obtain the asymptotic result for the case $X = V$, Kunisch and Volkwein required that solution time derivative data be included in the POD computation. They noted in [20] that including this extra data did not improve the accuracy of the reduced model in numerical experiments unless the time step was coarse. However, the time derivative data was found to increase the accuracy of the reduced model in [15].

In a recent paper [3], Chapelle, Gariah, and Sainte-Marie used a different error estimation technique and produced a new error bound for the case $X = V$ (they did not consider the case $X = H$) that does not depend on the time derivative of the solution. In addition, they proved that the error bound converges to zero under the assumption that the V operator norm of a certain POD projection is bounded as the number of POD modes increases. They noted that it was possible for this quantity to be unbounded; however, they produced analytical and numerical evidence that strongly suggests that this quantity is bounded for the problems they considered.

The main roadblock in the analysis in the works discussed above is the estimation of the error between the solution data used to construct the POD modes and different POD projections of that data. In this work, we consider time varying data taking values in two different Hilbert spaces and we prove *exact expressions* for these POD data approximation errors for four different POD projections and two error norms (H and V). These error expressions can be computed using only the POD eigenvalues and modes. We illustrate these new exact error expression results in section 4 with a numerical example. We also prove that the errors converge to zero as the number of POD modes in the projections increases. However, for one case we must make the same additional assumption that was made in [3] for the case $X = V$.

We show how to use these results with the error estimation approaches of [20, 21] and [3] to prove new POD model reduction error bounds and convergence results (as the number of POD modes increases) for the 2D Navier–Stokes equations in section 6. We prove the new error bounds tend to zero for POD space $X = H$ in both error estimation approaches. Moreover, for $X = H$, we use our new results along with the technique of [3] to prove a new error bound that tends to zero without including time derivative data in the POD data set.

Our results can be applied to derive new error bounds for standard POD Galerkin reduced order models of other parabolic PDEs, and our results should also be

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

applicable to error bounds for nonstandard POD reduced order models such as those considered in [2, 4, 18].

2. POD. We begin with a brief overview of the continuous POD theory for a collection of time varying functions taking values in a Hilbert space. For more details and for the proof of the main POD theory (for both discrete and time varying data), see, e.g., [14, 31, 21, 9, 32]. We also highlight the POD properties that will be important to the proofs of the main results in section 5.

Before we describe the continuous POD, we introduce some notation and background.

2.1. Notation. Let X be a Hilbert space with inner product[1] $(\cdot,\cdot)_X$ and corresponding norm $\|x\|_X = (x,x)_X^{1/2}$. For $X = \mathbb{C}^N$, the dot product can be taken as the standard dot product $(x,y)_X = y^*x$ or a weighted product $(x,y)_X = y^*Mx$, where M is a Hermitian positive definite matrix.

Let $I = (a,b)$ with $-\infty \le a < b \le \infty$. Let the inner product on $L^2(I)$ be given by

$$(f,g)_{L^2(I)} = \int_I f(t)\,\overline{g(t)}\,dt$$

with norm $\|h\|_{L^2(I)} = (h,h)_{L^2(I)}^{1/2}$. Also let I_j be open intervals for $j = 1,\ldots,m$, and let $L^2(I)^m$ denote the Hilbert space $L^2(I_1)\times L^2(I_2)\times\cdots\times L^2(I_m)$ with inner product

$$(f,g)_{L^2(I)^m} = \sum_{j=1}^{m}(f_j,g_j)_{L^2(I_j)}$$

for $f = [f_1,\ldots,f_m]^T$ and $g = [g_1,\ldots,g_m]^T$.

Let $L^2(I;X)$ be the Hilbert space of functions $w(t)$ such that $w(t)\in X$ for $t\in I$ and

$$\|w\|_{L^2(I;X)} = \left(\int_I \|w(t)\|_X^2\,dt\right)^{1/2} < \infty.$$

The inner product on $L^2(I;X)$ is

$$(w,z)_{L^2(I;X)} = \int_I \big(w(t),z(t)\big)_X\,dt.$$

2.2. The singular value decomposition and Hilbert–Schmidt operators. Let $K : X \to Y$ be a compact linear operator, where X and Y are separable Hilbert spaces. Let $\{x_k\}\subset X$ and $\{y_k\}\subset Y$ be orthonormal bases of eigenvectors of $K^*K : X\to X$ and $KK^* : Y\to Y$, where $K^* : Y\to X$ is the Hilbert adjoint operator. The nonzero (positive) eigenvalues $\{\lambda_k\}$ of K^*K and KK^* are the same, and we have

$$Kx_k = \sigma_k y_k, \quad K^*y_k = \sigma_k x_k,$$

where $\sigma_k = \lambda_k^{1/2}$. Furthermore, zero is an eigenvalue of K^*K with eigenvector x_k if and only if $Kx_k = 0$ since

$$K^*Kx_k = 0 \;\Rightarrow\; (K^*Kx_k,x_k)_X = 0 \;\Rightarrow\; \|Kx_k\|_Y^2 = 0 \;\Rightarrow\; Kx_k = 0,$$

and $Kx_k = 0$ implies $K^*Kx_k = 0$. Similarly, zero is an eigenvalue of KK^* with eigenvector y_k if and only if $K^*y_k = 0$.

[1] We assume all inner products are linear in the first argument and conjugate linear in the second argument so that $(\alpha\, x,y)_X = \alpha\,(x,y)_X$ and $(x,\alpha\, y)_X = \overline{\alpha}\,(x,y)_X$ for $\alpha\in\mathbb{C}$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

The singular value decomposition of K consists of the *singular vectors* $\{x_k\}$ and $\{y_k\}$ and the *singular values* $\{\sigma_k\}$, where we include zero in the list of singular values if either K^*K or KK^* has a zero eigenvalue. Furthermore, we order the singular values (and corresponding singular vectors) so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$. In general, it is possible for K to have infinitely many positive singular values and also a zero singular value.

Also, the operator K above is *Hilbert–Schmidt* (see, e.g., [1]) if

$$\|K\|_{HS} := \left(\sum_{k \geq 1} \|K z_k\|_Y^2 \right)^{1/2} < \infty, \tag{2.1}$$

where $\{z_k\} \subset X$ is any orthonormal basis. The above sum is independent of the choice of orthornormal basis, and also

$$\|K\|_{HS} = \left(\sum_{k \geq 1} \sigma_k^2 \right)^{1/2} < \infty.$$

Furthermore, if Z is another separable Hilbert space and $L : Y \to Z$ is bounded, then it follows from (2.1) that $LK : X \to Z$ is also Hilbert–Schmidt.

2.3. Continuous POD of one function. Let data $w \in L^2(I; X)$ be given. The POD problem for the data $w \in L^2(I; X)$ looks for an orthonormal basis $\{\varphi_i\} \subset X$ (the POD modes) minimizing the data approximation error

$$E_r = \|w - P_r w\|_{L^2(I;X)}^2 = \int_I \|w(t) - P_r w(t)\|_X^2 \, dt$$

for the data approximation

$$P_r w(t) = \sum_{i=1}^{r} \big(w(t), \varphi_i\big)_X \, \varphi_i. \tag{2.2}$$

To solve this data approximation problem, introduce the linear operator $K : L^2(I) \to X$ defined by

$$Ku = \int_I u(t) \, w(t) \, dt.$$

It can be shown that K is a compact linear operator, and therefore it has a singular value decomposition: there exist singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ and singular vectors $\{f_i\} \subset L^2(I)$ and $\{\varphi_i\} \subset X$ such that

$$K f_i = \sigma_i \varphi_i, \quad K^* \varphi_i = \sigma_i f_i, \tag{2.3}$$

where the Hilbert adjoint operator $K^* : X \to L^2(I)$ is given by

$$[K^* x](t) = \big(x, w(t)\big)_X.$$

The singular vectors are orthonormal bases for each space. Furthermore, K is known to be Hilbert–Schmidt (see, e.g., [28, Proposition 4.3]), and so the sum of the squares of the singular values is finite:

$$\sum_{i \geq 1} \sigma_i^2 < \infty.$$

The squares of the singular values, $\lambda_i = \sigma_i^2$, are often called the POD eigenvalues of the data.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

It can be shown that the POD modes $\{\varphi_i\} \subset X$ are the above singular vectors of K, and the approximation error is given by the sum of the squares of the neglected singular values, i.e.,

$$E_r = \|w - P_r w\|^2_{L^2(I;X)} = \int_I \|w(t) - P_r w(t)\|^2_X \, dt = \sum_{i>r} \sigma_i^2.$$

Furthermore, from the SVD equations (2.3) for K and the definition of K^*, we have

$$\big(w(t), \varphi_i\big)_X = \overline{[K^*\varphi_i](t)} = \sigma_i \overline{f_i(t)}.$$

Therefore, the data approximation can be rewritten as

$$P_r w(t) = \sum_{i=1}^{r} \sigma_i \, \overline{f_i(t)} \, \varphi_i.$$

The Hilbert–Schmidt property of K and the above form of $P_r w$ play crucial roles in the proofs of the main results in section 5.

2.4. Continuous POD of multiple functions. In model reduction applications, the POD of more than one function is often considered. We consider the case where all functions take values in the same Hilbert space, but they may be defined on different time intervals.[2]

Assume we are given multiple functions $w_j \in L^2(I_j; X)$ for $j = 1, \ldots, m$. The POD problem is again to find an orthonormal basis $\{\varphi_i\} \subset X$ minimizing the approximation error

$$E_r = \sum_{j=1}^{m} \|w_j - P_r w_j\|^2_{L^2(I_j;X)} \tag{2.4}$$

with data approximations $P_r w_j$ defined analogously to (2.2) above. The linear operator $K : L^2(I)^m \to X$ and its adjoint $K^* : X \to L^2(I)^m$ are now given by

$$Ku = \sum_{j=1}^{m} \int_{I_j} u_j(t)\, w_j(t)\, dt, \quad K^* x = \big[\big(x, w_1(t)\big)_X, \ldots, \big(x, w_m(t)\big)_X\big]^T. \tag{2.5}$$

The singular value decomposition $\{\sigma_i, f_i, \varphi_i\}$ of K again gives the solution. We have again that K is Hilbert–Schmidt, $E_r = \sum_{i>r} \sigma_i^2$, and

$$P_r w_j(t) = \sum_{i=1}^{r} \big(w_j(t), \varphi_i\big)_X \, \varphi_i = \sum_{i=1}^{r} \sigma_i \, \overline{f_{i,j}(t)} \, \varphi_i,$$

where $f_{i,j}$ is the jth component of the singular vector $f_i \in L^2(I)^m$.

3. POD projections: Basic properties and new results. The POD Galerkin model reduction error bounds and our main results all center around POD projections. In this section, we define four POD projections, give their basic properties, and present new POD data approximation errors involving these projections. As in the previous section, we consider the POD of multiple functions. To begin, we make a basic assumption that the data takes values in two Hilbert spaces and we fix notation for the POD singular values and singular vectors.

[2]The theory for the case of different time intervals is similar to the case of one time interval.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Main assumption. Let H and V be two Hilbert spaces with $V \subset H$, and let $w_j \in L^2(I_j; H) \cap L^2(I_j; V)$ for $j = 1, \dots, m$ be given data, where $I_j = (a_j, b_j)$ with $-\infty \leq a_j < b_j \leq \infty$.

At times, we assume V is dense in H or V is continuously embedded in H, i.e., there exists a constant $C_V > 0$ (called the embedding constant) such that $\|v\|_H \leq C_V \|v\|_V$ for all $v \in V$. We do not make either of these assumptions unless specified otherwise.

Since the data satisfies $w_j \in L^2(I_j; X)$ for $X = H$ and $X = V$, we can consider two different POD problems with error norm $X = H$ or $X = V$ in (2.4). The solution of the POD problem in each case leads to the operators $K : L^2(I)^m \to X$, where $X = H$ or $X = V$. Since the adjoint operator K^* depends on the X inner product, the singular value decomposition of K is different in each case.

Notation. Let $\{\sigma_k, f_k, \varphi_k\} \subset \mathbb{R} \times L^2(I)^m \times H$ denote the singular values and singular vectors of K, where $K : L^2(I)^m \to H$, and let $\{\mu_k, g_k, \psi_k\} \subset \mathbb{R} \times L^2(I)^m \times V$ denote the singular values and singular vectors of K, where now K is viewed as a mapping from $L^2(I)^m$ to V. We assume throughout that the singular values are ordered so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ and $\mu_1 \geq \mu_2 \geq \cdots \geq 0$.

The following lemma is important to keep in mind in this work.

LEMMA 3.1. *Let the main assumption hold.*

1. *If $\sigma_i > 0$, then $\varphi_i \in V$.*
2. *For $0 < M < \infty$, $\mu_i = 0$ for $i > M$ if and only if $\sigma_i = 0$ for all $i > M$.*
3. *$\sigma_i > 0$ for infinitely many i if and only if $\mu_i > 0$ for infinitely many i.*
4. *If V is densely and continuously embedded in H and $\mu_i > 0$ for all i, then $\sigma_i > 0$ for all i.*

Proof. To prove the first item, recall that φ_i satisfies $K f_i = \sigma_i \varphi_i$, where the operator K is viewed as a mapping from $L^2(I)^m$ to H. Since $w_j \in L^2(I_j; V)$, we have that K maps into V. This gives $\varphi_i \in V$.

For the second item, assume that $\mu_i = 0$ for $i > M$ with $0 < M < \infty$. Since the approximation error (2.4) for $X = V$ is given by

$$E_r = \sum_{i>r} \mu_i^2 = \sum_{i=r+1}^{M} \mu_i^2,$$

we have $E_M = 0$ and therefore $w_j(t) = \sum_{i=1}^{M} \left(w_j(t), \psi_i \right)_V \psi_i$ for each j. Therefore $K : L^2(I)^m \to H$ has rank at most M for both $X = V$ and $X = H$ by the definition of K in (2.5). Since $K : L^2(I)^m \to H$ has rank at most M, we have $\sigma_i = 0$ for all $i > M$. The same argument shows that $\sigma_i = 0$ for $i > M$ implies $\mu_i = 0$ for $i > M$. This also proves $\sigma_i > 0$ for infinitely many i if and only if $\mu_i > 0$ for infinitely many i.

For the final part, since $\mu_i > 0$ for all i, the null space of $K : L^2(I)^m \to V$ and its adjoint $K^{*_V} : V \to L^2(I)^m$ consist of only zero. We show the null space of $K : L^2(I)^m \to H$ and its adjoint $K^{*_H} : H \to L^2(I)^m$ also consist of only zero. First, note that $K : L^2(I)^m \to H$ and $K : L^2(I)^m \to V$ have the same null space. Therefore, we need only show K^{*_H} has zero null space.

For $X = H$ or $X = V$, let $\mathcal{N}(K^{*_X})$ and $\mathcal{R}(K)$ denote the null space of $K^{*_X} : X \to L^2(I)^m$ and the range of $K : L^2(I)^m \to X$. Then we have $X = \mathcal{N}(K^{*_X}) \oplus \overline{\mathcal{R}(K)}$, where the bar denotes the closure in X. Since the null space of K^{*_V} is only zero, the closure in V of $\mathcal{R}(K)$ is all of V.

Let $x \in H$ and let $\varepsilon > 0$. Since V is dense in H, there is a $v \in V$ with $\|x - v\|_H < \varepsilon/2$. Also, there exists $r \in \mathcal{R}(K)$ such that $\|v - r\|_V < \varepsilon/(2C_V)$, where C_V is the embedding constant. Then

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$\|x - r\|_H \leq \|x - v\|_H + \|v - r\|_H < \|x - v\|_H + C_V \|v - r\|_V < \varepsilon,$$

and so the closure in H of $\mathcal{R}(K)$ is all of H. Therefore, the null space of K^{*_H} consists of only zero. $\square$

Our analysis is centered around the following four orthogonal projections.

DEFINITION 3.2. *Define the sets H_r and V_r by*

$$H_r = \operatorname{span}\{\varphi_1, \ldots, \varphi_r\}, \quad V_r = \operatorname{span}\{\psi_1, \ldots, \psi_r\}.$$

Also define the orthogonal projections $P_r^H, Q_r^H : H \to H$ and $P_r^V, Q_r^V : V \to V$ by

1. $P_r^H : H \to H_r$, *and for* $x \in H$, $P_r^H x$ *minimizes* $\inf_{x_r \in H_r} \|x - x_r\|_H$.
2. $P_r^V : V \to H_r$, *and for* $v \in V$, $P_r^V v$ *minimizes* $\inf_{x_r \in H_r} \|v - x_r\|_V$.
3. $Q_r^H : H \to V_r$, *and for* $x \in H$, $Q_r^H x$ *minimizes* $\inf_{v_r \in V_r} \|x - v_r\|_H$.
4. $Q_r^V : V \to V_r$, *and for* $v \in V$, $Q_r^V v$ *minimizes* $\inf_{v_r \in V_r} \|v - v_r\|_V$.

If $\sigma_i > 0$ for $i = 1, \ldots, r$, the lemma above gives that $H_r \subset V$ and so P_r^V is well defined and also P_r^H maps into V.

These projections can be computed using the following expressions:

$$P_r^H x = \sum_{k=1}^r (x, \varphi_k)_H \varphi_k, \qquad P_r^V v = \sum_{k=1}^r c_k \, \varphi_k, \quad S_r c = y,$$

$$Q_r^H x = \sum_{k=1}^r d_k \, \psi_k, \quad M_r d = z, \qquad Q_r^V v = \sum_{k=1}^r (v, \psi_k)_V \psi_k,$$

where $c = [c_1, \ldots, c_r]^T$, $d = [d_1, \ldots, d_r]^T$, and the matrices $S_r, M_r \in \mathbb{C}^{r \times r}$ and vectors $y, z \in \mathbb{C}^r$ have entries

$$S_{r,ij} = (\varphi_j, \varphi_i)_V, \quad M_{r,ij} = (\psi_j, \psi_i)_H, \quad y_i = (v, \varphi_i)_V, \quad z_i = (x, \psi_i)_H. \tag{3.1}$$

The orthogonal projections P_r^H and Q_r^H are bounded operators on H; however, we can also view P_r^H and Q_r^H as operators from H to V or from V to V. Lemma 2 in [20] essentially shows these operators are bounded; in the latter case, we make the additional assumption that V is continuously embedded in H.

LEMMA 3.3. *If the main assumption holds, then $Q_r^H : H \to V$ is bounded. If, in addition, $\sigma_i > 0$ for $i = 1, \ldots, r$, then $P_r^H : H \to V$ is bounded. Furthermore,*

$$\|Q_r^H v\|_V \leq \|M_r^{-1}\|_2^{1/2} \, \|v\|_H, \quad \|P_r^H v\|_V \leq \|S_r\|_2^{1/2} \, \|v\|_H,$$

where $\| \cdot \|_2$ denotes the spectral norm of a matrix, and $M_r, S_r \in \mathbb{C}^{r \times r}$ are defined in (3.1).

If also V is continuously embedded in H with embedding constant C_V, then $Q_r^H : V \to V$ and $P_r^H : V \to V$ are bounded and

$$\|Q_r^H v\|_V \leq C_V \, \|M_r^{-1}\|_2^{1/2} \, \|v\|_V, \quad \|P_r^H v\|_V \leq C_V \, \|S_r\|_2^{1/2} \, \|v\|_V.$$

Proof. Let $v \in V$. Since $\sigma_i > 0$ for $i = 1, \ldots, r$, Lemma 3.1 gives that $H_r \subset V$ so that P_r^H maps into V. The proof now follows the proof of Lemma 2 in [20]. We also use that since P_r^H and Q_r^H are orthogonal projections on H, we have $\|P_r^H v\|_H \leq \|v\|_H$ and $\|Q_r^H v\|_H \leq \|v\|_H$. $\square$

Note that we have *not* proved that the operator norms $\|P_r^H\|_{\mathcal{L}(V)}$ and $\|Q_r^H\|_{\mathcal{L}(V)}$ can be bounded independent of r.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Next, the orthogonal projections P_r^V and Q_r^V are bounded operators on V; however, we can also view P_r^V and Q_r^V as operators from V to H. Again, Lemma 2 in [20] shows that these operators are bounded.

LEMMA 3.4. *If the main assumption holds, then $P_r^V : V \to H$ and $Q_r^V : V \to H$ are bounded and*

$$\|P_r^V v\|_H \le \|S_r^{-1}\|_2^{1/2} \|v\|_V, \quad \|Q_r^V v\|_H \le \|M_r\|_2^{1/2} \|v\|_V.$$

3.1. New results. Our new results give *exact expressions* for various POD data approximation errors $\|w-\pi_r w\|_{L^2(I;W)}$ involving the four different POD projections π_r and different error norms $W = H$ and $W = V$. The error expressions are computable using only the POD singular values and modes. We also give convergence results for the errors as $r \to \infty$.

We give an overview of the results here and give detailed statements of the results, along with the proofs, in section 5. For some of the results, we must assume V is continuously embedded in H.

Recall the well-known POD data approximation error expressions (from section 2 with $X = H$ and $X = V$, respectively):

$$\sum_{j=1}^{m} \|w_j - P_r^H w_j\|_{L^2(I_j;H)}^2 = \sum_{j=1}^{m} \int_{I_j} \|w_j(t) - P_r^H w_j(t)\|_H^2 \, dt = \sum_{i>r} \sigma_i^2 < \infty, \tag{3.2}$$

$$\sum_{j=1}^{m} \|w_j - Q_r^V w_j\|_{L^2(I_j;V)}^2 = \sum_{j=1}^{m} \int_{I_j} \|w_j(t) - Q_r^V w_j(t)\|_V^2 \, dt = \sum_{i>r} \mu_i^2 < \infty. \tag{3.3}$$

Here, the errors tend to zero as $r \to \infty$.

Our results are as follows: If the main assumption on the data holds, then the data approximation errors for the H error norm are

$$\sum_{j=1}^{m} \|w_j - P_r^V w_j\|_{L^2(I_j;H)}^2 = \sum_{i>r} \sigma_i^2 \, \|\varphi_i - P_r^V \varphi_i\|_H^2 < \infty,$$

$$\sum_{j=1}^{m} \|w_j - Q_r^H w_j\|_{L^2(I_j;H)}^2 = \sum_{i>r} \mu_i^2 \, \|\psi_i - Q_r^H \psi_i\|_H^2 < \infty,$$

$$\sum_{j=1}^{m} \|w_j - Q_r^V w_j\|_{L^2(I_j;H)}^2 = \sum_{i>r} \mu_i^2 \, \|\psi_i\|_H^2 < \infty,$$

and for the V error norm the error expressions are

$$\sum_{j=1}^{m} \|w_j - Q_r^H w_j\|_{L^2(I_j;V)}^2 = \sum_{i>r} \mu_i^2 \, \|\psi_i - Q_r^H \psi_i\|_V^2 < \infty,$$

$$\sum_{j=1}^{m} \|w_j - P_r^V w_j\|_{L^2(I_j;V)}^2 = \sum_{i>r} \sigma_i^2 \, \|\varphi_i - P_r^V \varphi_i\|_V^2 < \infty,$$

$$\sum_{j=1}^{m} \|w_j - P_r^H w_j\|_{L^2(I_j;V)}^2 = \sum_{i>r} \sigma_i^2 \, \|\varphi_i\|_V^2 < \infty.$$

In addition, all the above errors tend to zero as $r \to \infty$. However, for the case of Q_r^H with V error norm, we must make an additional assumption to guarantee

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

convergence of the error to zero as $r \to \infty$. (See Remark 1 below.) As we mentioned in the introduction, this convergence result for Q_r^H and the V norm was obtained in [3]. As far as we are aware, all the other results are new.[3]

4. Numerical example. Before we prove the main results, we give a numerical example illustrating the exact expressions for the approximation errors. Consider a 2D scalar Burgers' equation

$$w_t + w\, w_x + w\, w_y = \mu\, (w_{xx} + w_{yy}),$$

on the unit square Ω with zero Dirichlet boundary conditions and piecewise constant initial condition $w(0,x,y) = 1$ if $x \leq 1/2$ and $w(0,x,y) = 0$ otherwise. The solution is square integrable in time with values in the Hilbert spaces $H = L^2(\Omega)$ and $V = H_0^1(\Omega)$, where the latter space is given the inner product $(u,v)_V = \int_\Omega u_x\, v_x + u_y\, v_y$. Also, due to the Poincaré inequality, V is continuously embedded in H.

We take $\mu = 1/100$ and approximate the solution for $0 \leq t \leq 1$ using the group finite element method (see, e.g., [5, 6, 7]) with continuous piecewise bilinear basis functions and `ode23s` of MATLAB for the time stepping scheme.

For the POD computations, we take the approximate solution values at each time step, $w(t_k)$, and form a piecewise constant function in time; we set the value on the kth time interval as the average $(w(t_{k+1}) + w(t_k))/2$ of the approximate solution at the endpoints of the interval. As is well known, the POD eigenvalues and modes can be found exactly in this case [29]. Also, the approximation errors can be computed exactly for this piecewise constant in time function so that we can compare the actual approximation errors with the error formulas.

Due to round-off errors in the POD eigenvalue computation, some of the computed error formulas have extremely small imaginary parts; we report the absolute value of the computed error formulas below. Also, the two POD error formulas (3.2) and (3.3) are of course already known; however, we also present computational results for these formulas below to give a point of comparison for the other error formulas.

Table 1 shows the actual approximation errors and the computed error formulas for all eight scenarios with $r = 5$ and 201 equally spaced finite element nodes in each coordinate direction. To give an example of the notation in the table, the last line in the table gives the values

$$\text{computed actual error} = \int_I \|w(t) - P_r^H w(t)\|_V^2\, dt,$$
$$\text{computed error formula} = \sum_{i>r} \sigma_i^2\, \|\varphi_i\|_V^2.$$

Also, line 1 and line 5 in the table correspond to the two known POD error formulas (3.2) and (3.3). The table shows excellent agreement between all eight actual errors and the error formulas. Rounding error in the POD computations led to small differences in the computed values, but the numerical results illustrate the theoretical findings.

Table 2 again shows excellent agreement between the actual approximation errors and the computed error formulas for all eight scenarios with $r = 15$. We also see all the errors tending to zero as r becomes larger in accordance with the theory. Of

[3]Of course, the convergence of the error for Q_r^V and the H error norm follows directly from V being continuously embedded in H and (3.3).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Table 1
Comparison of actual approximation errors and error formulas with $r = 5$ using 201 equally spaced finite element nodes in each coordinate direction.

Projection	Error norm	Actual error	Error formula
P_r^H	H norm	1.084×10^{-5}	1.148×10^{-5}
P_r^V	H norm	3.867×10^{-5}	3.952×10^{-5}
Q_r^H	H norm	2.040×10^{-5}	2.159×10^{-5}
Q_r^V	H norm	6.754×10^{-5}	6.949×10^{-5}
Q_r^V	V norm	1.848×10^{-2}	1.862×10^{-2}
Q_r^H	V norm	1.017×10^{-1}	1.034×10^{-1}
P_r^V	V norm	3.340×10^{-2}	3.345×10^{-2}
P_r^H	V norm	5.589×10^{-2}	5.614×10^{-2}

Table 2
Comparison of actual approximation errors and error formulas with $r = 15$ using 201 equally spaced finite element nodes in each coordinate direction.

Projection	Error norm	Actual error	Error formula
P_r^H	H norm	9.611×10^{-10}	1.031×10^{-9}
P_r^V	H norm	6.720×10^{-9}	6.826×10^{-9}
Q_r^H	H norm	2.377×10^{-9}	2.603×10^{-9}
Q_r^V	H norm	6.915×10^{-9}	7.354×10^{-9}
Q_r^V	V norm	5.421×10^{-6}	5.530×10^{-6}
Q_r^H	V norm	5.429×10^{-5}	5.688×10^{-5}
P_r^V	V norm	1.682×10^{-5}	1.684×10^{-5}
P_r^H	V norm	6.140×10^{-5}	6.191×10^{-5}

course, as mentioned above, we have not proved the error will tend to zero for Q_r^H in the V norm. Further increasing r gives even smaller errors. Also, increasing the number of finite element nodes gives similar results.

As we discuss in Remark 1 below, we can prove $Q_r^H w$ converges to w in $L^2(I;V)$ as $r \to \infty$ if the quantity $\|Q_r^H \psi_i\|_V^2$ is bounded for all r and $i > r$. Figure 1 shows the values $\|Q_r^H \psi_{r+k}\|_V^2$ for $k = 1, \ldots, 40$ and various values of r. We see that these values remain bounded as r increases. Therefore, we expect convergence as $r \to \infty$, as we have already seen above.

5. Proofs of main results. We now prove the main results: for all four POD projections, we give exact expressions for the data approximation errors in different error norms and prove the errors tend to zero as the number of POD modes tends to infinity. We also consider the convergence of the approximation errors for individual elements of the Hilbert spaces H and V below.

In this section, we assume the general framework of sections 2 and 3 holds. Let V and H be (possibly complex) Hilbert spaces with $V \subset H$. We do not assume the embedding is dense or continuous unless specified otherwise. Assume the given data w_j satisfies $w_j \in L^2(I_j;H) \cap L^2(I_j;V)$, with $V \subset H$, for $j = 1, \ldots, m$.

We begin by approximating w_j by $P_r^V w_j$ in the $L^2(I_j;V)$ norm and by $Q_r^H w_j$ in the $L^2(I_j;H)$ norm. The following lemma plays a crucial role in the proof.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

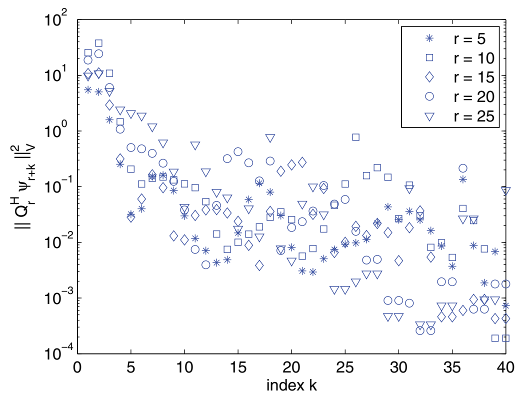

FIG. 1. *The values* $\|Q_r^H \psi_{r+k}\|_V^2$ *for* $k = 1, \ldots, 40$ *and various values of* r *computed using* 201 *equally spaced finite element nodes in each coordinate direction.*

LEMMA 5.1. *Let* $1 \le r, s < \infty$*. If the main assumption holds, then*

$$\sum_{j=1}^{m} \|(I - Q_r^H) P_s^H w_j\|_{L^2(I_j;H)}^2 = \sum_{k=1}^{s} \sigma_k^2 \, \|\varphi_k - Q_r^H \varphi_k\|_H^2, \tag{5.1}$$

$$\sum_{j=1}^{m} \|(I - Q_r^H) Q_s^V w_j\|_{L^2(I_j;H)}^2 = \sum_{k=r+1}^{s} \mu_k^2 \, \|\psi_k - Q_r^H \psi_k\|_H^2. \tag{5.2}$$

In addition, if $\sigma_i > 0$ *for* $i = 1, \ldots, r$*, then*

$$\sum_{j=1}^{m} \|(I - P_r^V) Q_s^V w_j\|_{L^2(I_j;V)}^2 = \sum_{k=1}^{s} \mu_k^2 \, \|\psi_k - P_r^V \psi_k\|_V^2, \tag{5.3}$$

$$\sum_{j=1}^{m} \|(I - P_r^V) P_s^H w_j\|_{L^2(I_j;V)}^2 = \sum_{k=r+1}^{s} \sigma_k^2 \, \|\varphi_k - P_r^V \varphi_k\|_V^2. \tag{5.4}$$

Proof. We prove the results (5.3) and (5.4) involving P_r^V. The remaining proofs are similar.

We start with (5.3). First, since $\sigma_i > 0$ for $i = 1, \ldots, r$, Lemma 3.1 gives that $H_r \subset V$ so that P_r^V is well defined. Recall from section 2 that

$$\big(w_j(t), \psi_i\big)_V = \mu_i \, \overline{g_{i,j}(t)},$$

where $g_{i,j}$ is the jth component of $g_i \in L^2(I)^m$, which is the ith singular vector of the operator $K : L^2(I)^m \to V$. Therefore,

$$Q_s^V w_j(t) = \sum_{i=1}^{s} \big(w_j(t), \psi_i\big)_V \psi_i = \sum_{i=1}^{s} \mu_i \, \overline{g_{i,j}(t)} \, \psi_i, \quad P_r^V Q_s^V w_j(t) = \sum_{i=1}^{s} \mu_i \, \overline{g_{i,j}(t)} \, P_r^V \psi_i,$$

which gives

$$(I - P_r^V) Q_s^V w_j(t) = \sum_{i=1}^{s} \mu_i \, \overline{g_{i,j}(t)} \, (\psi_i - P_r^V \psi_i). \tag{5.5}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Next, use the orthonormality of $\{g_i\}$ in $L^2(I)^m$ and this separated form to directly compute the approximation error:

$$\begin{aligned}
&\sum_{j=1}^{m} \|(I - P_r^V) Q_s^V w_j\|^2_{L^2(I_j;V)} \\
&\quad = \sum_{j=1}^{m} \big((I - P_r^V) Q_s^V w_j, (I - P_r^V) Q_s^V w_j\big)_{L^2(I_j;V)} \\
&\quad = \sum_{j=1}^{m} \sum_{k,\ell=1}^{s} \mu_k \, \mu_\ell \, (g_{\ell,j}, g_{k,j})_{L^2(I_j)} \, (\psi_k - P_r^V \psi_k, \psi_\ell - P_r^V \psi_\ell)_V \\
&\quad = \sum_{k,\ell=1}^{s} \mu_k \, \mu_\ell \, (g_\ell, g_k)_{L^2(I)^m} \, (\psi_k - P_r^V \psi_k, \psi_\ell - P_r^V \psi_\ell)_V \\
&\quad = \sum_{k=1}^{s} \mu_k^2 \, \|\psi_k - P_r^V \psi_k\|_V^2.
\end{aligned}$$

This proves the error expression (5.3).

The proof of the error expression (5.4) is similar; the main difference is that for $k = 1, \ldots, r$, $P_r^V \varphi_k = \varphi_k$ since $\varphi_k \in H_r$. $\square$

THEOREM 5.2. *Let the main assumption hold and let $r \geq 1$. Then*

$$\sum_{j=1}^{m} \|w_j - Q_r^H w_j\|^2_{L^2(I_j;H)} = \sum_{i>r} \mu_i^2 \, \|\psi_i - Q_r^H \psi_i\|_H^2 < \infty. \tag{5.6}$$

If $\sigma_i > 0$ for $i = 1, \ldots, r$, then

$$\sum_{j=1}^{m} \|w_j - P_r^V w_j\|^2_{L^2(I_j;V)} = \sum_{i>r} \sigma_i^2 \, \|\varphi_i - P_r^V \varphi_i\|_V^2 < \infty. \tag{5.7}$$

Furthermore, $Q_r^H w_j \to w_j$ in $L^2(I_j;H)$ and $P_r^V w_j \to w_j$ in $L^2(I_j;V)$ as $r \to \infty$ for each j.

Proof. We prove the results involving P_r^V. The proofs involving Q_r^H are similar. Again, since $\sigma_i > 0$ for $i = 1, \ldots, r$, Lemma 3.1 gives that $H_r \subset V$ so that P_r^V is well defined.

To begin, assume $\sigma_i = 0$ for $i > s$ for some $s \in \mathbb{N}$. In this case, the known POD data approximation error expression (3.2) gives that $w_j = P_s^H w_j$ for each j. Therefore, the error expression (5.4) in Lemma 5.1 proves the main error expression (5.7). Furthermore, this error expression shows that $w_j = P_s^V w_j$ since $\sigma_i = 0$ for $i > s$.

Next, assume $\sigma_i > 0$ for infinitely many i, or equivalently $\mu_i > 0$ for infinitely many i (see Lemma 3.1). We cannot extend the above argument to the case $s = \infty$ since we do not know if $P_s^H w_j$ converges to w_j in $L^2(I_j;V)$ for each j. Instead, we approximate w_j by $Q_s^V w_j$ and derive an alternate expression for the approximation error. Then we use the Hilbert–Schmidt property of K to prove that the alternate error expression actually equals the main error expression (5.7).

To begin, recall $I - P_r^V$ is an orthogonal projection on V, and so $\|I - P_r^V\|_{\mathcal{L}(V)} = 1$. Also recall $Q_s^V w_j \to w_j$ in $L^2(I_j;V)$ as $s \to \infty$ for each j by the known POD data approximation error expression (3.3). This gives

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$\sum_{j=1}^{m} \|(I - P_r^V)(w_j - Q_s^V w_j)\|^2_{L^2(I_j;V)} \le \sum_{j=1}^{m} \|w_j - Q_s^V w_j\|^2_{L^2(I_j;V)} \to 0 \quad \text{as } s \to \infty.$$

Therefore, $(I - P_r^V)Q_s^V w_j \to (I - P_r^V)w_j$ in $L^2(I_j;V)$ as $s \to \infty$ for each j. Now use (5.3) in Lemma 5.1 above to give an alternate error expression:

$$\begin{aligned}\sum_{j=1}^{m} \|w_j - P_r^V w_j\|^2_{L^2(I_j;V)} &= \lim_{s\to\infty} \sum_{j=1}^{m} \|(I - P_r^V)Q_s^V w_j\|^2_{L^2(I_j;V)} \\ &= \lim_{s\to\infty} \sum_{k=1}^{s} \mu_k^2 \, \|\psi_k - P_r^V \psi_k\|_V^2 = \sum_{k\in\mathbb{J}_\mu} \mu_k^2 \, \|\psi_k - P_r^V \psi_k\|_V^2,\end{aligned}$$

where $\mathbb{J}_\mu = \{k : \mu_k > 0\}$. Since $\|I - P_r^V\|_{\mathcal{L}(V)} = 1$ and $\|\psi_i\|_V = 1$, this sum is finite since $\sum_{i\ge1} \mu_i^2 < \infty$.

Next, we prove the main error expression (5.7). Since $\mu_k \psi_k = K g_k$ for $k \in \mathbb{J}_\mu$ and $K g_k = 0$ for $k \notin \mathbb{J}_\mu$, the alternate error expression above can be rewritten as

$$\sum_{j=1}^{m} \|w_j - P_r^V w_j\|^2_{L^2(I_j;V)} = \sum_{k\in\mathbb{J}_\mu} \|(I - P_r^V)K g_k\|_V^2 = \sum_{k=1}^{\infty} \|(I - P_r^V)K g_k\|_V^2.$$

Since $I - P_r^V : V \to V$ is bounded and $K : L^2(I)^m \to V$ is Hilbert–Schmidt, we have $(I - P_r^V)K : L^2(I)^m \to V$ is also Hilbert–Schmidt. Also, since $\{g_k\}$ is an orthonormal basis for $L^2(I)^m$, the right-hand side of the above equation equals the square of the Hilbert–Schmidt norm of $(I - P_r^V)K$. The Hilbert–Schmidt norm is independent of the choice of orthonormal basis, so we can replace $\{g_k\}$ with $\{f_k\}$ to obtain

$$\sum_{j=1}^{m} \|w_j - P_r^V w_j\|^2_{L^2(I_j;V)} = \sum_{k=1}^{\infty} \|(I - P_r^V)K f_k\|_V^2 = \sum_{k\in\mathbb{J}_\sigma} \|(I - P_r^V)K f_k\|_V^2,$$

where $\mathbb{J}_\sigma = \{k : \sigma_k > 0\}$ and we used $K f_k = 0$ for $k \notin \mathbb{J}_\sigma$. This gives

$$\sum_{j=1}^{m} \|w_j - P_r^V w_j\|^2_{L^2(I_j;V)} = \sum_{k\in\mathbb{J}_\sigma} \sigma_k^2 \, \|(I - P_r^V)\varphi_k\|_V^2 = \sum_{k>r} \sigma_k^2 \, \|(I - P_r^V)\varphi_k\|_V^2,$$

where we used $K f_k = \sigma_k \varphi_k$ for $k \in \mathbb{J}_\sigma$ and $P_r^V \varphi_k = \varphi_k$ for $k = 1, \dots, r$ (since $\varphi_k \in H_r$). Note in the right-hand side of the above equation we suppressed that $k \in \mathbb{J}_\sigma$ for ease of notation; we continue to use the simpler notation throughout this work. This proves the main error expression (5.7).

Next, we show $P_r^V w_j \to w_j$ in $L^2(I_j;V)$ as $r \to \infty$ for each j. Since $\|I - P_r^V\|_{\mathcal{L}(V)} = 1$, the right-hand side of the main error expression (5.7) can be bounded as

$$\sum_{i>r} \sigma_i^2 \, \|\varphi_i - P_r^V \varphi_i\|_V^2 \le \sum_{i>r} \sigma_i^2 \, \|\varphi_i\|_V^2 = \sum_{i>r} \|K f_i\|_V^2.$$

Since the Hilbert–Schmidt norm of $K : L^2(I)^m \to V$ equals the square root of $\sum_{i\ge1} \|K f_i\|_V^2$, the right-hand side of the above bound tends to zero as $r \to \infty$. $\square$

Next, we consider approximating w_j by $P_r^H w_j$ and $Q_r^H w_j$ in the $L^2(I_j;V)$ norm. When infinitely many of the singular values of K are positive, we make an additional

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

assumption that V is continuously embedded in H in order to guarantee P_r^H and Q_r^H are bounded when viewed as linear operators on V (see Lemma 3.3). As before, we prove that the errors converge to zero as $r \to \infty$; however, for Q_r^H we must make the additional assumption that $\|Q_r^H\|_{\mathcal{L}(V)}$ is bounded independent of r.

THEOREM 5.3. *Let the main assumption hold and let $r, s \in \mathbb{N}$. If $\mu_i = 0$ for $i > s$, then*

$$\sum_{j=1}^{m} \|w_j - Q_r^H w_j\|_{L^2(I_j;V)}^2 = \sum_{i>r} \mu_i^2 \, \|\psi_i - Q_r^H \psi_i\|_V^2 < \infty. \tag{5.8}$$

If also $\sigma_i > 0$ for $i = 1, \ldots, r$, then

$$\sum_{j=1}^{m} \|w_j - P_r^H w_j\|_{L^2(I_j;V)}^2 = \sum_{i>r} \sigma_i^2 \, \|\varphi_i\|_V^2 < \infty. \tag{5.9}$$

If $\mu_i > 0$ for infinitely many i and V is continuously embedded in H, then (5.8) *and* (5.9) *hold and also $P_r^H w_j \to w_j$ in $L^2(I_j;V)$ as $r \to \infty$ for each j. If, in addition, $\|Q_r^H\|_{\mathcal{L}(V)}$ is bounded independent of r, then $Q_r^H w_j \to w_j$ in $L^2(I_j;V)$ as $r \to \infty$ for each j.*

Remark 1. As mentioned in the introduction, the bound

$$\sum_{j=1}^{m} \|w_j - Q_r^H w_j\|_{L^2(I_j;V)}^2 \le \left(1 + \|Q_r^H\|_{\mathcal{L}(V)}\right)^2 \sum_{i>r} \mu_i^2$$

was obtained in [3] for the case $m = 1$. This bound can be obtained directly from the exact expression for the error in (5.8). The convergence result for Q_r^H as $r \to \infty$ assuming $\|Q_r^H\|_{\mathcal{L}(V)}$ is bounded independent of r was also obtained in [3].

Also, since $(\psi_i, Q_r^H \psi_i)_V = 0$ for $i > r$, we have $\|\psi_i - Q_r^H \psi_i\|_V^2 = 1 + \|Q_r^H \psi_i\|_V^2$ for $i > r$. Therefore, (5.8) shows that if $\|Q_r^H \psi_i\|_V^2$ is bounded for all r and $i > r$, then $Q_r^H w_j \to w_j$ in $L^2(I_j;V)$ for each j.

Proof. The proof is similar to the proof of Theorem 5.2. We only indicate the main differences.

We begin with P_r^H. First consider the case $\sigma_i = 0$ for $i > s$. An argument similar to the proof of Lemma 5.1 gives

$$\sum_{j=1}^{m} \|w_j - P_r^H w_j\|_{L^2(I_j;V)}^2 = \sum_{i=r+1}^{s} \sigma_i^2 \, \|\varphi_i - P_r^H \varphi_i\|_V^2,$$

where we used $P_r^H \varphi_i = \varphi_i$ for $i = 1, \ldots, r$. We also have $P_r^H \varphi_i = 0$ for $i > r$, and this proves the main error expression (5.9) when $\sigma_i = 0$ for $i > s$.

Now assume $\mu_i > 0$ for infinitely many i, and also assume V is continuously embedded in H. Since the operator $P_r^H : V \to V$ is bounded (by Lemma 3.3) and $P_r^H \varphi_i = 0$ for $i > r$, the proof is nearly identical.

Next, consider Q_r^H. The proof is similar except for the convergence as $r \to \infty$. A main difference in this case is that $Q_r^H \psi_i$ does not necessarily equal zero for $i > r$. However, if $\|Q_r^H\|_{\mathcal{L}(V)}$ is bounded independent of r, then $\|I - Q_r^H\|_{\mathcal{L}(V)}$ is also bounded independent of r and the approximation error is bounded as

$$\sum_{i>r} \mu_i^2 \, \|\psi_i - Q_r^H \psi_i\|_V^2 \le \|I - Q_r^H\|_{\mathcal{L}(V)}^2 \sum_{i>r} \mu_i^2$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

since $\|\psi_i\|_V = 1$. The convergence as $r \to \infty$ follows. ∎

Last, we consider approximating w_j by $P_r^V w_j$ and $Q_r^V w_j$ in the $L^2(I_j; H)$ norm. When V is continuously embedded in H, we can simply bound the approximation errors by the errors in the $L^2(I_j; V)$ norm. However, we give the exact error expressions for completeness.

THEOREM 5.4. *Let the main assumption hold and let $r, s \in \mathbb{N}$. If $\sigma_i = 0$ for $i > s$, then*

$$\sum_{j=1}^{m} \|w_j - Q_r^V w_j\|_{L^2(I_j;H)}^2 = \sum_{i>r} \mu_i^2 \, \|\psi_i\|_H^2 < \infty. \tag{5.10}$$

If also $\sigma_i > 0$ for $i = 1, \ldots, r$, then

$$\sum_{j=1}^{m} \|w_j - P_r^V w_j\|_{L^2(I_j;H)}^2 = \sum_{i>r} \sigma_i^2 \, \|\varphi_i - P_r^V \varphi_i\|_H^2 < \infty. \tag{5.11}$$

If $\mu_i > 0$ for infinitely many i and V is continuously embedded in H, then (5.10) *and* (5.11) *hold and also $Q_r^V w_j \to w_j$ and $P_r^V w_j \to w_j$ in $L^2(I_j; H)$ as $r \to \infty$ for each j.*

Proof. Again, the proof is similar to the proofs of Theorems 5.2 and 5.3, and we only indicate the main differences.

First, consider the case of P_r^V. Since V is continuously embedded in H, we have

$$\sum_{j=1}^{m} \|(I - P_r^V)(w_j - P_s^H w_j)\|_{L^2(I_j;H)}^2 \leq C_V^2 \sum_{j=1}^{m} \|w_j - P_s^H w_j\|_{L^2(I_j;V)}^2,$$

where C_V is the embedding constant and we used $\|I - P_r^V\|_{\mathcal{L}(V)} = 1$. By Theorem 5.3, the right-hand side tends to zero as $s \to \infty$ and so $(I - P_r^V) P_s^H w_j \to (I - P_r^V) w_j$ in $L^2(I_j; H)$ as $s \to \infty$ for each j. The main error expression (5.11) follows by proving a result similar to Lemma 5.1.

Next, consider Q_r^V. Similar to above, it can be shown that $(I - Q_r^V) P_s^H w_j \to (I - Q_r^V) w_j$ in $L^2(I_j; H)$ as $s \to \infty$ for each j. An alternate error expression can be shown by proving a similar result to Lemma 5.1. Then, Lemma 3.4 gives that $Q_r^V : V \to H$ is bounded. Since V is continuously embedded in H, the identity operator $I : V \to H$ is bounded, and so $I - Q_r^V : V \to H$ is bounded. Since $K : L^2(I)^m \to V$ is Hilbert–Schmidt and $I - Q_r^V : V \to H$ is bounded, we have $(I - Q_r^V) K : L^2(I)^m \to H$ is Hilbert–Schmidt. This can be used to prove the main error expression (5.10) when $\mu_i > 0$ for infinitely many i.

To show the convergence as $r \to \infty$, use the continuous embedding to bound the approximation errors by the errors in the $L^2(I_j; V)$ norm. These errors tend to zero by (3.3) and Theorem 5.2. ∎

We also consider the approximation capability of the POD projections for individual elements of the Hilbert spaces H and V. We know $\|P_r^H x - x\|_H \to 0$ for all $x \in H$ since $\{\varphi_k\}$ is an orthonormal basis for H, and similarly $\|Q_r^V w - w\|_V \to 0$ for all $w \in V$. Below, we show other convergence results for different POD projections and error norms.

PROPOSITION 5.5. *Let the main assumption hold. If V is densely and continuously embedded in H and $\mu_i > 0$ for all i, then*

(a) *$\sigma_i > 0$ and $\varphi_i \in V$ for all i, and so $H_r \subset V$ for all r;*

(b) *for any $w \in \mathcal{R}(K)$ (the range of K), $\|P_r^H w - w\|_V \to 0$ as $r \to \infty$;*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

(c) *for any* $w \in V$, $\|P_r^V w - w\|_V \to 0$ *as* $r \to \infty$.

If $\sigma_i > 0$ *for all* i, *then*

(d) *for any* $x \in \mathcal{R}(K)$, $\|Q_r^V x - x\|_H \to 0$ *as* $r \to \infty$;

(e) *for any* $x \in H$, $\|Q_r^H x - x\|_H \to 0$ *as* $r \to \infty$.

Proof. First, part (a) follows immediately from Lemma 3.1.

Next, note that $w \in \mathcal{R}(K)$ implies that there is a unique $q \in L^2(I)^m$ such that $w = Kq$. To see this, let $w = Kq_1$ and $w = Kq_2$. Then $K(q_1 - q_2) = 0$. However, since zero is not a singular value of K, we must have that $q_1 - q_2 = 0$ or $q_1 = q_2$.

To prove (b), let w be in the range of K with $w = Kq$. We have

$$\begin{aligned} P_r^H w &= \sum_{k=1}^r (w, \varphi_k)_H \varphi_k = \sum_{k=1}^r (Kq, \varphi_k)_H \varphi_k = \sum_{k=1}^r (q, K^* \varphi_k)_{L^2(I)^m} \varphi_k \\ &= \sum_{k=1}^r (q, f_k)_{L^2(I)^m} \sigma_k \varphi_k = \sum_{k=1}^r (q, f_k)_{L^2(I)^m} K f_k = K \sum_{k=1}^r (q, f_k)_{L^2(I)^m} f_k. \end{aligned}$$

Therefore, we have

$$P_r^H w = K q_r, \quad q_r = \sum_{k=1}^r (q, f_k)_{L^2(I)^m} f_k.$$

This gives

$$\|P_r^H w - w\|_V = \|Kq_r - Kq\|_V \le \|K\|_{\mathcal{L}(L^2(I)^m, V)} \|q_r - q\|_{L^2(I)^m}.$$

Since $w_j \in L^2(I_j; V)$ for $j = 1, \ldots, m$, we have that $\|K\|_{\mathcal{L}(L^2(I)^m, V)} < \infty$. Also, since $\{f_k\}$ is an orthonormal basis for $L^2(I)^m$, this gives that $\|q_r - q\|_{L^2(I)^m} \to 0$ as $r \to \infty$. Therefore, $\|P_r^H w - w\|_V \to 0$ for all w in the range of K.

For (c), first note that $Q_N^V w$ is in the range of K for $w \in V$ since $Kg_k = \mu_k \psi_k$ and $\mu_k \neq 0$ for all k gives

$$Q_N^V w = \sum_{k=1}^N (w, \psi_k)_V \psi_k = K \sum_{k=1}^N \mu_k^{-1} (w, \psi_k)_V g_k.$$

Then for fixed N, $\|P_r^H Q_N^V w - Q_N^V w\|_V \to 0$ as $r \to \infty$ by part (b).

Let $w \in V$ and let $\varepsilon > 0$. We have

$$\begin{aligned} \|P_r^V w - w\|_V &= \inf_{x_r \in H_r} \|x_r - w\|_V \\ &\le \|P_r^H Q_N^V w - w\|_V \le \|P_r^H Q_N^V w - Q_N^V w\|_V + \|Q_N^V w - w\|_V. \end{aligned}$$

Since $\{\psi_k\}$ is an orthonormal basis for V, the second term can be made less than $\varepsilon/2$ for N large enough; then for N fixed, the first term can made less than $\varepsilon/2$ for r large enough. This proves (c) since $P_r^H Q_N^V w$ is in H_r for all $w \in V$.

The proofs of (d)–(e) are similar to (b)–(c). ∎

6. New model reduction error bounds for the 2D Navier–Stokes equations. We now apply the new POD data approximation error expressions to give new error bounds for the 2D Navier–Stokes equations and other fluid dynamics equations. POD model reduction error bounds have been considered for this problem in [21, 8]. We prove error bounds and convergence results using Kunisch and Volkwein's

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

approach [20, 21] in section 6.2 and using the approach of Chapelle, Gariah, and Sainte-Marie [3] in section 6.3.

As mentioned in the introduction, numerical examples have shown that including the time derivative data may or may not increase the accuracy of the reduced order model. This is why we consider both error estimation approaches in this work.

It is possible that one approach could yield better error bounds than the other approach for certain problems. In this work, we do not attempt to closely track all the constants appearing in the error bounds. It may be interesting to give thorough comparisons of the error bounds for various problems. Also, we do not attempt to give the sharpest error bounds possible. We leave these issues to be considered elsewhere.

We assume the POD eigenvalues and modes are computed exactly using exact solution data. We do not consider errors due to time stepping and space discretization. These errors can be treated as in other works; for the 2D Navier–Stokes equations considered here, errors due to discretization in time are considered in [21] and errors in the solution data are treated in [8]. We focus on applying the new error expressions. Some of the error estimation methods below are standard, and we only sketch proofs in those cases.

We obtain error bounds with constants that depend exponentially on the quantity $\|w\|^2_{L^2(0,T;V)}$, where w is the solution of the problem. This is similar to known results for other numerical methods; see, e.g., [11, 22]. For sufficiently small problem data, this quantity can be bounded independent of T. In general however, as Heywood and Rannacher explain in [11], this exponential dependence on the solution magnitude is expected since the true solution may be unstable and therefore nearby approximate solutions can diverge exponentially fast (with T) from the true solution. Error bounds that do not grow with the final time T have been obtained for various numerical methods under a stability assumption on the solution that is to be approximated [10, 12]. It is possible that similar results could be obtained for the reduced order models considered here. We leave this to be investigated elsewhere.

6.1. Problem formulation. The 2D Navier–Stokes equations and other fluid dynamics equations can be placed in the following abstract formulation [30]. Let H and V be two real separable Hilbert spaces with inner products $(\cdot,\cdot)_H$ and $(\cdot,\cdot)_V$ and corresponding norms $\|\cdot\|_H = (\cdot,\cdot)_H^{1/2}$ and $\|\cdot\|_V = (\cdot,\cdot)_V^{1/2}$ such that V is densely and compactly embedded in H. Due to this embedding, there is a positive constant C_V so that $\|v\|_H \leq C_V \|v\|_V$ for all $v \in V$. Also let V' be the dual space of V, and denote the action of a functional $f \in V'$ on $v \in V$ by $\langle f, v\rangle$.

Let $a : V \times V \to \mathbb{C}$ be a symmetric bilinear form that is bounded and coercive, i.e.,

$$a(u, v) \leq C_a \|u\|_V \|v\|_V, \quad a(v, v) \geq c_a \|v\|_V^2,$$

for some constants $c_a, C_a > 0$, and all vectors u and v in V. The bilinear form a can be used to define an unbounded linear operator $A : D(A) \subset H \to H$ defined by

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

for all $v \in V$ and all $u \in D(A)$, where $D(A)$ is the set of all $u \in V$ such that there exists $h \in H$ with $a(u, v) = (h, v)_H$ for all $v \in V$.

Next, consider another linear term given by a bounded linear operator $Q : V \to V'$ such that $Q : D(A) \to H$ and

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$|\langle Qv, v\rangle| \le C_Q \, \|v\|_V^{1+\gamma_1} \, \|v\|_H^{1-\gamma_1} \quad \text{for all } v \in V,$$
$$\|Qv\|_H \le C_Q \, \|v\|_V^{1-\gamma_2} \, \|Av\|_H^{\gamma_2} \quad \text{for all } v \in D(A)$$

for some constants $C_Q > 0$ and $\gamma_1, \gamma_2 \in [0, 1)$. For ease of notation, set

$$a_Q(u, v) := a(u, v) + \langle Qu, v\rangle.$$

Also assume $A + Q$ is coercive:

$$a_Q(v, v) \ge c_{aQ} \, \|v\|_V^2 \tag{6.1}$$

for some constant $c_{aQ} > 0$ and for all $v \in V$.

Finally, for the quadratic nonlinear term, let $B : V \times V \to V'$ be bilinear and continuous such that $B : D(A) \times D(A) \to H$. For $u, v, w \in V$, let $b(u, v, w) = \langle B(u, v), w\rangle$, and assume B satisfies

$$\begin{aligned} b(u, v, v) &= 0, \\ |b(u, v, w)| &\le C_B \, \|u\|_H^{\gamma_3} \, \|u\|_V^{1-\gamma_3} \, \|v\|_V \, \|w\|_V^{\gamma_3} \, \|w\|_H^{1-\gamma_3}, \\ \|B(u, z)\|_H + \|B(z, u)\|_H &\le C_B \, \|u\|_V \, \|z\|_V^{1-\gamma_4} \, \|Az\|_H^{\gamma_4}, \\ \|B(u, z)\|_H &\le C_B \, \|u\|_H^{\gamma_5} \, \|u\|_V^{1-\gamma_5} \, \|z\|_V^{1-\gamma_5} \, \|Az\|_H^{\gamma_5} \end{aligned} \tag{6.2}$$

for all $u, v, w \in V$, $z \in D(A)$, and some constants $C_B > 0$ and $\gamma_3, \gamma_4, \gamma_5 \in [0, 1)$. The first two conditions in (6.2) give

$$b(u, v, w) = -b(u, w, v), \quad |b(u, v, w)| \le C_B \, C_V \, \|u\|_V \, \|v\|_V \, \|w\|_V,$$
$$|b(u, v, u)| \le C_B \, \|u\|_H \, \|u\|_V \, \|v\|_V$$

for all $u, v, w \in V$.

For a given forcing $f \in L^2(0, T; V')$ and initial data $w_0 \in H$, the nonlinear evolution equation is given by

$$\begin{aligned} \frac{d}{dt}(w(t), v)_H + a_Q(w(t), v) + b(w(t), w(t), v) &= \langle f(t), v\rangle \quad \text{for all } v \in V, \\ w(0) &= w_0, \end{aligned} \tag{6.3}$$

where the differential equation holds for almost every $t \in (0, T)$. This problem is known to be well posed.

THEOREM 6.1 (see [30]). *Let $T > 0$ and let the above assumptions hold. If $f \in L^2(0, T; V')$ and $w_0 \in H$, then there exists a unique solution w of* (6.3) *satisfying*

$$w \in C([0, T]; H) \cap L^2(0, T; V), \quad \dot{w} \in L^2(0, T; V').$$

If $f \in L^2(0, T; H)$ and $w_0 \in V$, then the unique solution satisfies

$$w \in C([0, T]; V) \cap L^2(0, T; D(A)), \quad \dot{w} \in L^2(0, T; H).$$

For a given finite-dimensional space $W_r \subset V$, the Galerkin reduced order model is

$$\begin{aligned} \frac{d}{dt}(w_r(t), v_r)_H + a_Q(w_r(t), v_r) + b(w_r(t), w_r(t), v_r) &= \langle f(t), v_r\rangle \quad \text{for all } v_r \in W_r, \\ w_r(0) &= w_{r,0} \in W_r, \end{aligned} \tag{6.4}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

and this equation is known to have a unique solution $w_r \in H^1(0, T; V)$ satisfying $w_r(t) \in W_r$ for all t. Below, we consider W_r to be the span of the POD modes constructed from exact solution data.

Before we bound the error using the two approaches, we give one elementary property of the approximate solution w_r that will be useful later.

LEMMA 6.2. *Let $T > 0$ and let the above assumptions hold. If $f \in L^2(0, T; V')$, $w_0 \in H$, and $\|w_{r,0}\|_H$ is bounded independent of r, then there exists a constant $C > 0$ such that*

$$\|w_r\|_{L^\infty(0,T;H)} \le C$$

for all $r \ge 1$.

This result can be proved using well-known energy estimates: take $v_r = w_r(t)$ in (6.4) and use the inequality (6.1) for $A + Q$ and the property $b(u, u, u) = 0$ for all $u \in V$.

In the error bounds below, we use C to denote a positive constant that does not depend on r. The value of C may change from one line to the next; we do not attempt to give the most precise values of the bounding constants. Also, we repeatedly use Young's inequality: for any $a, b, \varepsilon > 0$,

$$ab \le \frac{\varepsilon}{2}\, a^2 + \frac{1}{2\varepsilon}\, b^2.$$

6.2. Error estimate approach 1: *V*-orthogonal projections. First, we consider Kunisch and Volkwein's error bounding approach from [20, 21]. For $f \in L^2(0, T; H)$ and $w_0 \in V$, let w be the exact solution of the nonlinear evolution equation (6.3) and let $\dot{w}$ denote the time derivative of w. We further assume that f and w_0 are regular enough so that $\dot{w}$ satisfies $\dot{w} \in L^2(0, T; V)$. Also, we assume throughout that the assumptions of section 6.1 above hold.

For $w_1 = w$ and $w_2 = \dot{w}$, let $\{\sigma_k, f_k, \varphi_k\} \subset \mathbb{R} \times L^2(0, T)^2 \times H$ and $\{\mu_k, g_k, \psi_k\} \subset \mathbb{R} \times L^2(0, T)^2 \times V$ be the singular values and singular vectors of the operator $K : L^2(0, T)^2 \to X$ for $X = H$ and $X = V$, respectively. Also as before, define $H_r = \operatorname{span}\{\varphi_1, \dots, \varphi_r\}$ and $V_r = \operatorname{span}\{\psi_1, \dots, \psi_r\}$. We set $W_r = H_r$ if $X = H$, and we set $W_r = V_r$ if $X = V$. When $X = H$, assume $\sigma_i > 0$ for $i = 1, \dots, r$ so that $H_r \subset V$.

To bound the error, we split the error as

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

where $\pi_r^V : V \to V$ is the orthogonal projection onto W_r, i.e., $\pi_r^V = P_r^V$ when $X = H$ and $\pi_r^V = Q_r^V$ when $X = V$. We prove the following result below.

THEOREM 6.3. *Let $T > 0$, $f \in L^2(0, T; H)$, and $w_0 \in V$, and let w and w_r be the exact solutions of the nonlinear evolution equation* (6.3) *and the reduced order model* (6.4)*, respectively, for the set W_r as defined above. If $\dot{w} \in L^2(0, T; V)$ and $\|w_{r,0}\|_H$ is bounded independent of r, then there exists a constant C, independent of r, such that*

$$\|w - w_r\|_{L^2(0,T;V)} \le C_1(T)\, \|w_{r,0} - P_r^V w_0\|_H + C_2(T) \left(\sum_{i>r} \sigma_i^2\, \|\varphi_i - P_r^V \varphi_i\|_V^2 \right)^{1/2}$$

if $X = H$ and

$$\|w - w_r\|_{L^2(0,T;V)} \le C_1(T)\, \|w_{r,0} - Q_r^V w_0\|_H + C_2(T) \left(\sum_{i>r} \mu_i^2 \right)^{1/2}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

if $X = V$, where

$$C_1(T) = c_{aQ}^{-1/2} \exp\left(2\, c_{aQ}^{-1}\, C_B^2\, \|w\|_{L^2(0,T;V)}^2\right), \quad C_2(T) = 1 + C\, C_1(T).$$

Furthermore, if $\|w_{r,0} - \pi_r^V w_0\|_H \to 0$ as $r \to \infty$, then $w_r \to w$ in $L^2(0,T;V)$ as $r \to \infty$ for both $X = H$ and $X = V$.

Before we prove the result, we note the following:

- The result for the case $X = V$ can be obtained without using the new results from this work as it relies on the known POD data approximation error expression (3.3). This result was included for completeness.
- For the initial condition $w_{r,0} = \pi_r^V w_0$, the first term in each error bound equals zero. However, other choices of $w_{r,0} \in W_r$ are possible. For example, for $w_{r,0} = P_r^H w_0 \in H_r$ in the case $X = H$ we have

  $$\|w_{r,0} - P_r^V w_0\|_H \le \|P_r^H w_0 - w_0\|_H + \|w_0 - P_r^V w_0\|_H.$$

  The first term tends to zero as $r \to \infty$ since $\{\varphi_k\}$ is an orthonormal basis for H. Proposition 5.5 above gives that the second term also tends to zero as $r \to \infty$ (since $w_0 \in V$ and V is continuously embedded in H) assuming $\mu_k > 0$ for all k.
- The error can also be split using the Ritz projection as in [20, 21]. Another variation is to construct the POD modes using the Hilbert space $X = V_a$, where $V_a = V$ with alternate inner product $(u, v)_{V_a} = a(u, v)$ for $u, v \in V$. Similar error bounds and convergence results can be obtained for both cases.

To begin, we give a preliminary lemma where we prove an error bound for the nonlinear term. The proof is similar to part of the proof of [8, Theorem 6].

LEMMA 6.4. *Let $f \in L^2(0,T;H)$ and $w_0 \in V$, and let w and w_r be the exact solutions of the nonlinear evolution equation* (6.3) *and the reduced order model* (6.4)*, respectively. If $\dot{w} \in L^2(0,T;V)$ and $\|w_{r,0}\|_H$ is bounded independent of r, then for any $\varepsilon > 0$ there exists a constant C, independent of r, such that*

$$\begin{aligned}|b(w_r(t), w_r(t), \theta_r(t)) - b(w(t), w(t), \theta_r(t))| &\le C\, \|\rho_r(t)\|_V^2 + \varepsilon\, \|\theta_r(t)\|_V^2 \\ &\quad + \frac{C_B^2}{2\,\varepsilon}\, \|w(t)\|_V^2\, \|\theta_r(t)\|_H^2\end{aligned}$$

for all $t \in (0,T)$.

Proof. We sketch the proof. First, recall the exact solution w satisfies $w \in C([0,T];V)$. We have

$$\|\rho_r\|_{L^\infty(0,T;H)} \le C_V\, \|I - \pi_r^V\|_{\mathcal{L}(V)}\, \|w\|_{L^\infty(0,T;V)} = C_V\, \|w\|_{L^\infty(0,T;V)},$$

where we used the continuous embedding of $V \subset H$ and $\|I - \pi_r^V\|_{\mathcal{L}(V)} = 1$. Therefore, ρ_r is uniformly bounded in $L^\infty(0,T;H)$. Also,

$$\begin{aligned}\|\theta_r\|_{L^\infty(0,T;H)} &\le \|\pi_r^V w\|_{L^\infty(0,T;H)} + \|w_r\|_{L^\infty(0,T;H)} \\ &\le C_V\, \|w\|_{L^\infty(0,T;V)} + \|w_r\|_{L^\infty(0,T;H)}.\end{aligned}$$

Lemma 6.2 gives that w_r is uniformly bounded in $L^\infty(0,T;H)$, and so θ_r is uniformly bounded in $L^\infty(0,T;H)$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Next, use the properties of B to show (suppressing the dependence on t for ease)

$$\begin{aligned}|b(w_r, w_r, \theta_r) - b(w, w, \theta_r)| &\leq |b(w, \rho_r, \theta_r)| + |b(\rho_r, w, \theta_r)| + |b(\theta_r, w, \theta_r)| \\ &\quad + |b(\rho_r, \theta_r, \rho_r)| + |b(\theta_r, \rho_r, \theta_r)| \\ &\leq 2\, C_B\, C_V\, \|w\|_{L^\infty(0,T;V)}\, \|\theta_r\|_V\, \|\rho_r\|_V \\ &\quad + C_B\, \|w\|_V\, \|\theta_r\|_H\, \|\theta_r\|_V \\ &\quad + C_B \left(\|\rho_r\|_{L^\infty(0,T;H)} + \|\theta_r\|_{L^\infty(0,T;H)}\right) \|\theta_r\|_V\, \|\rho_r\|_V.\end{aligned}$$

The result now follows from Young's inequality. ∎

In the proof of [8, Theorem 6], it appears the coefficient multiplying $\|\theta_r(t)\|_H^2$ in the above result is taken as an arbitrary constant; we are unable to follow the proof of this.

Now we prove the error bounds and convergence results of Theorem 6.3.

Proof. As above, split the error as

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

For any $v_r \in W_r$, subtract the two equations (6.3) and (6.4) to give

$$(\dot{\theta}_r, v_r)_H + a_Q(\theta_r, v_r) = -(\dot{\rho}_r, v_r)_H - a_Q(\rho_r, v_r) + b(w_r, w_r, v_r) - b(w, w, v_r). \tag{6.5}$$

Take $v_r = \theta_r$ and estimate (using Lemma 6.4 with $\varepsilon = c_{aQ}/4$)

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\|\theta_r(t)\|_H^2 + c_{aQ}\,\|\theta_r(t)\|_V^2 \\ &\qquad \leq \|\dot{\rho}_r(t)\|_H\,\|\theta_r(t)\|_H + \left(C_a + \|Q\|_{\mathcal{L}(V,V')}\right)\|\rho_r(t)\|_V\,\|\theta_r(t)\|_V \\ &\qquad\quad + C\,\|\rho_r(t)\|_V^2 + \frac{2C_B^2}{c_{aQ}}\|w(t)\|_V^2\,\|\theta_r(t)\|_H^2 + \frac{c_{aQ}}{4}\|\theta_r(t)\|_V^2 \\ &\qquad \leq C\left(\|\dot{\rho}_r(t)\|_V^2 + \|\rho_r(t)\|_V^2\right) + \frac{c_{aQ}}{2}\|\theta_r(t)\|_V^2 + \frac{2C_B^2}{c_{aQ}}\|w(t)\|_V^2\,\|\theta_r(t)\|_H^2.\end{aligned}$$

This gives

$$\frac{d}{dt}\|\theta_r(t)\|_H^2 + c_{aQ}\,\|\theta_r(t)\|_V^2 \leq C\left(\|\dot{\rho}_r(t)\|_V^2 + \|\rho_r(t)\|_V^2\right) + \frac{4C_B^2}{c_{aQ}}\|w(t)\|_V^2\,\|\theta_r(t)\|_H^2.$$

Gronwall's inequality yields

$$\begin{aligned}&\|\theta_r(T)\|_H^2 + c_{aQ}\, e^{G(T)}\int_0^T e^{-G(s)}\,\|\theta_r(s)\|_V^2\, ds \\ &\qquad \leq e^{G(T)}\,\|\theta_r(0)\|_H^2 + C\, e^{G(T)}\int_0^T e^{-G(s)}\left(\|\dot{\rho}_r(s)\|_V^2 + \|\rho_r(s)\|_V^2\right) ds,\end{aligned}$$

where

$$G(t) = 4\, c_{aQ}^{-1}\, C_B^2 \int_0^t \|w(s)\|_V^2\, ds = 4\, c_{aQ}^{-1}\, C_B^2\, \|w\|_{L^2(0,t;V)}^2.$$

Neglect the $\|\theta_r(T)\|_H^2$ term and use that the function $G(\cdot)$ is increasing to obtain

$$c_{aQ}\,\|\theta_r\|_{L^2(0,T;V)}^2 \leq e^{G(T)}\,\|\theta_r(0)\|_H^2 + C\, e^{G(T)}\left(\|\dot{\rho}_r\|_{L^2(0,T;V)}^2 + \|\rho_r\|_{L^2(0,T;V)}^2\right).$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

The error bounds follow from the triangle inequality

$$\|w - w_r\|_{L^2(0,T;V)} \leq \|\rho_r\|_{L^2(0,T;V)} + \|\theta_r\|_{L^2(0,T;V)},$$

the inequality $(a+b)^{1/2} \leq a^{1/2}+b^{1/2}$, and the exact error expressions and convergence results from (3.3) for $X = V$ and Theorem 5.2 for $X = H$. ∎

6.3. Error estimate approach 2: *H*-orthogonal projections. Next, we use the approach of Chapelle, Gariah, and Sainte-Marie [3] to bound the error. We do not need to include the time derivative of the solution in the POD data; however, we are unable to conclude convergence of the error as $r \to \infty$ if $X = V$ without assuming $\|Q_r^H\|_{\mathcal{L}(V)}$ is bounded independent of r. We prove convergence of the error as $r \to \infty$ if $X = H$ without making additional assumptions. Also, we eliminate the regularity assumption $\dot{w} \in L^2(0,T;V)$ required in the approach above. However, we do require $f \in L^2(0,T;H)$ and $w_0 \in V$ in order to guarantee the solution satisfies $w \in C([0,T];V)$. We use this fact to bound the nonlinear term.

As before, assume throughout that the assumptions of section 6.1 above hold. For $f \in L^2(0,T;H)$ and $w_0 \in V$, let w be the exact solution of the nonlinear evolution equation (6.3). For $w_1 = w$, let $\{\sigma_k, f_k, \varphi_k\} \subset \mathbb{R}\times L^2(0,T)\times H$ and $\{\mu_k, g_k, \psi_k\} \subset \mathbb{R}\times L^2(0,T)\times V$ be the singular values and singular vectors of the operator $K : L^2(0,T) \to X$ for $X = H$ and $X = V$, respectively. As before, define $H_r = \operatorname{span}\{\varphi_1, \ldots, \varphi_r\}$ and $V_r = \operatorname{span}\{\psi_1, \ldots, \psi_r\}$. Set $W_r = H_r$ if $X = H$, and set $W_r = V_r$ if $X = V$. When $X = H$, assume $\sigma_i > 0$ for $i = 1, \ldots, r$ so that $H_r \subset V$.

Theorem 6.5. *Let $T > 0$, $f \in L^2(0,T;H)$, and $w_0 \in V$, and let w and w_r be the exact solutions of the nonlinear evolution equation* (6.3) *and the reduced order model* (6.4)*, respectively, for the set W_r as defined above. If $\|w_{r,0}\|_H$ is bounded independent of r, then there exists a constant C, independent of r, such that*

$$\|w - w_r\|_{L^2(0,T;V)} \leq C_1(T)\, \|w_{r,0} - P_r^H w_0\|_H + C_2(T) \left(\sum_{i>r} \sigma_i^2\, \|\varphi_i\|_V^2 \right)^{1/2}$$

if $X = H$ and

$$\|w - w_r\|_{L^2(0,T;V)} \leq C_1(T)\, \|w_{r,0} - Q_r^H w_0\|_H + C_2(T) \left(\sum_{i>r} \mu_i^2\, \|\psi_i - Q_r^H \psi_i\|_V^2 \right)^{1/2}$$

if $X = V$, where

$$C_1(T) = c_{aQ}^{-1/2} \exp\left(2\, c_{aQ}^{-1}\, C_B^2\, \|w\|_{L^2(0,T;V)}^2\right), \quad C_2(T) = 1 + C\, C_1(T).$$

Furthermore, if $\|w_{r,0} - P_r^H w_0\|_H \to 0$ as $r \to \infty$, then $w_r \to w$ in $L^2(0,T;V)$ as $r \to \infty$ for $X = H$. Also, if $\|w_{r,0} - Q_r^H w_0\|_H \to 0$ as $r \to \infty$ and $\|Q_r^H\|_{\mathcal{L}(V)}$ is bounded independent of r, then $w_r \to w$ in $L^2(0,T;V)$ as $r \to \infty$ for $X = V$.

The error bounds are very similar to the bounds obtained in approach 1. In fact, even though we have not carefully tracked the values of the constant C appearing in both results, it is not hard to see that the values are comparable in both approaches. Therefore, it appears the primary difference in the error bounds is the values of the POD data approximation errors.

Proof. Let $\pi_r^H : H \to H$ be the orthogonal projection onto W_r, i.e., $\pi_r^H = P_r^H$ when $X = H$ and $\pi_r^H = Q_r^H$ when $X = V$. Split the error as

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

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Similar to the proof of Lemma 6.4, we can show ρ_r and θ_r are uniformly bounded in $L^\infty(0,T;H)$ since $w \in C([0,T];H)$, $\|I - \pi_r^H\|_{\mathcal{L}(H)} = \|\pi_r^H\|_{\mathcal{L}(H)} = 1$, and w_r is uniformly bounded in $L^\infty(0,T;H)$ (by Lemma 6.2). Therefore, the error bound result for the nonlinear term of Lemma 6.4 can be proved for this case using $w \in C([0,T];V)$, but without the requirement $\dot{w} \in L^2(0,T;V)$.

Next, for any $v_r \in W_r$, use $w_r = \pi_r^H w - \theta_r$ in the reduced order model (6.4) to give

$$\frac{d}{dt}(\theta_r, v_r)_H + a_Q(\theta_r, v_r) = -(f, v_r)_H + \frac{d}{dt}(\pi_r^H w, v_r)_H + a_Q(\pi_r^H w, v_r) + b(w_r, w_r, v_r).$$

Since $\pi_r^H : H \to H$ is self-adjoint, we have

$$\frac{d}{dt}(\pi_r^H w, v_r)_H = \frac{d}{dt}(w, \pi_r^H v_r)_H = \frac{d}{dt}(w, v_r)_H$$

since $v_r \in W_r$ implies $\pi_r^H v_r = v_r$. Use the nonlinear evolution equation (6.3) in the above equation to give

$$(\dot{\theta}_r, v_r)_H + a_Q(\theta_r, v_r) = -a_Q(\rho_r, v_r) + b(w_r, w_r, v_r) - b(w, w, v_r)$$

for any $v_r \in W_r$.

This equation is similar to (6.5) obtained in approach 1, but now we do not have the term containing $\dot{\rho}_r$. The remainder of the proof is similar to approach 1, except we use Theorem 5.3 for the POD error expressions and convergence results. $\square$

7. Conclusion. We considered POD data approximation errors of time varying functions taking values in two different Hilbert spaces H and V with $V \subset H$. We considered four different POD projections and both error norms. In all cases we proved exact expressions for the errors that are computable using only the POD eigenvalues and modes. Furthermore, we proved all the data approximation errors tend to zero as the number of POD modes in the projections increases. For the case of the POD projection Q_r^H with the V error norm, we made the additional assumption that $\|Q_r^H\|_{\mathcal{L}(V)}$ is bounded independent of r to prove the convergence. As far as the author is aware, all these results are new except this last case, which was proved in [3].

We applied these results to give new error bounds and convergence results for POD reduced order models of the 2D Navier–Stokes equations. We considered two different error estimation approaches: the first approach from [20, 21] used V-orthogonal projections, and the second approach from [3] used H-orthogonal projections. The first approach requires including the time derivative of the solution in the POD data for the most complete results and it also requires the solution has sufficient regularity. The second approach eliminates the time derivative requirement and lessens the regularity requirement; however, for the POD space $X = V$ we can only conclude the error converges to zero as the number of POD modes increases under the above assumption on Q_r^H. For both approaches with POD space $X = H$, we proved error bounds and convergence without making any additional assumptions. This convergence has been observed numerically for many problems in many papers; however, we believe this is the first work to prove computable error bounds that tend to zero as the number of POD modes increases.

We proved error bounds for different POD reduced order models constructed using different POD techniques. It does not appear that our results directly indicate which of these reduced order models is the most accurate. It is possible that further analysis

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

will indicate which inner product (H or V) is more desirable and also whether it is advantageous to include the time derivative data in the POD computation. Recent work analyzing some of the different techniques can be found in [17].

We did not consider estimating errors due to time and space discretization as has been done in other works. However, our results (possibly in a discrete form) should be applicable when considering these additional sources of error to give improved error bounds and convergence results for these cases.

It is likely that our results are also applicable to POD model reduction error bounds for other types of PDEs. For POD error bounds for hyperbolic PDEs, see [19, 3].

Also, we have only considered approximation errors for the POD projections for one-dimensional time intervals. Similar arguments can give error expressions for POD projections arising from a discrete POD setting with a (possibly weighted) sum replacing the time integrals. Furthermore, it is possible that similar error expressions can be derived for multidimensional parameter domains in place of the one-dimensional time interval.

Our work does not give error estimates or error indicators for the POD reduced order model when model parameters or initial conditions are varied from those used to construct the POD basis. (Some work in this direction can be found in [16].) However, our work gives further insight into the behavior of POD reduced order models for PDEs in the case where the model is not varied. Hopefully, an increased understanding of this case will lead to additional progress in analyzing the effects of problem variations on reduced order models.

Acknowledgment. The author would like to thank Mark Opmeer for very helpful discussions and also for the proof of Lemma 3.1, part 4 (which is simpler than the author's original proof).

REFERENCES

[1] J.-P. Aubin, *Applied Functional Analysis*, 2nd ed., Wiley-Interscience, New York, 2000.

[2] J. Borggaard, T. Iliescu, and Z. Wang, *Artificial viscosity proper orthogonal decomposition*, Math. Comput. Modelling, 53 (2011), pp. 269–279.

[3] 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.

[4] S. Chaturantabut and D. C. Sorensen, *A state space error estimate for POD-DEIM nonlinear model reduction*, SIAM J. Numer. Anal., 50 (2012), pp. 46–63.

[5] I. Christie, D. F. Griffiths, A. R. Mitchell, and J. M. Sanz-Serna, *Product approximation for nonlinear problems in the finite element method*, IMA J. Numer. Anal., 1 (1981), pp. 253–266.

[6] C. A. J. Fletcher, *The group finite element formulation*, Comput. Methods Appl. Mech. Engrg., 37 (1983), pp. 225–244.

[7] C. A. J. Fletcher, *Computational Techniques for Fluid Dynamics* I, Springer-Verlag, Berlin, 1991.

[8] P. Galán del Sastre and R. Bermejo, *Error estimates of proper orthogonal decomposition eigenvectors and Galerkin projection for a general dynamical system arising in fluid models*, Numer. Math., 110 (2008), pp. 49–81.

[9] T. Henri and J.-P. Yvon, *Stability of the POD and Convergence of the POD Galerkin Method for Parabolic Problems*, Prépublication 02-40, Institut de Recherche Mathématique de Rennes, Rennes cedex, France, 2002.

[10] J. G. Heywood, *An error estimate uniform in time for spectral Galerkin approximations of the Navier-Stokes problem*, Pacific J. Math., 98 (1982), pp. 333–345.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[11] J. G. Heywood and R. Rannacher, *Finite element approximation of the nonstationary Navier–Stokes problem.* I. *Regularity of solutions and second-order error estimates for spatial discretization*, SIAM J. Numer. Anal., 19 (1982), pp. 275–311.
[12] J. G. Heywood and R. Rannacher, *Finite element approximation of the nonstationary Navier–Stokes problem, part* II: *Stability of solutions and error estimates uniform in time*, SIAM J. Numer. Anal., 23 (1986), pp. 750–777.
[13] M. Hinze and S. Volkwein, *Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition*, Comput. Optim. Appl., 39 (2008), pp. 319–345.
[14] P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge, UK, 1996.
[15] D. Hömberg and S. Volkwein, *Control of laser surface hardening by a reduced-order approach using proper orthogonal decomposition*, Math. Comput. Modelling, 38 (2003), pp. 1003–1028.
[16] C. Homescu, L. R. Petzold, and R. Serban, *Error estimation for reduced-order models of dynamical systems*, SIAM J. Numer. Anal., 43 (2005), pp. 1693–1714.
[17] T. Iliescu and Z. Wang, *Are the snapshot difference quotients needed in the proper orthogonal decomposition?* SIAM J. Sci. Comput., submitted, available from http://arxiv.org/abs/1303.6012.
[18] T. Iliescu and Z. Wang, *Variational multiscale proper orthogonal decomposition: Convection-dominated convection-diffusion-reaction equations*, Math. Comp., 82 (2013), pp. 1357–1378.
[19] I. Kalashnikova and M. F. Barone, *On the stability and convergence of a Galerkin reduced order model (ROM) of compressible flow with solid wall and far-field boundary treatment*, Internat. J. Numer. Methods Engrg., 83 (2010), pp. 1345–1375.
[20] K. Kunisch and S. Volkwein, *Galerkin proper orthogonal decomposition methods for parabolic problems*, Numer. Math., 90 (2001), pp. 117–148.
[21] 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.
[22] W. Layton, *Introduction to the Numerical Analysis of Incompressible Viscous Flows*, Comput. Sci. Eng. 6, SIAM, Philadelphia, 2008.
[23] Z. Luo, J. Chen, I. M. Navon, and X. Yang, *Mixed finite element formulation and error estimates based on proper orthogonal decomposition for the nonstationary Navier–Stokes equations*, SIAM J. Numer. Anal., 47 (2008), pp. 1–19.
[24] Z. Luo, J. Chen, P. Sun, and X. Yang, *Finite element formulation based on proper orthogonal decomposition for parabolic equations*, Sci. China Ser. A, 52 (2009), pp. 585–596.
[25] Z. Luo, Z. Xie, Y. Shang, and J. Chen, *A reduced finite volume element formulation and numerical simulations based on POD for parabolic problems*, J. Comput. Appl. Math., 235 (2011), pp. 2098–2111.
[26] Z. Luo, X. Yang, and Y. Zhou, *A reduced finite difference scheme based on singular value decomposition and proper orthogonal decomposition for Burgers equation*, J. Comput. Appl. Math., 229 (2009), pp. 97–107.
[27] S. S. Ravindran, *Error analysis for Galerkin POD approximation of the nonstationary Boussinesq equations*, Numer. Methods Partial Differential Equations, 27 (2011), pp. 1639–1665.
[28] J. R. Singler, *Convergent snapshot algorithms for infinite dimensional Lyapunov equations*, IMA J. Numer. Anal., 31 (2011), pp. 1468–1496.
[29] L. Sirovich, *Turbulence and the dynamics of coherent structures.* I. *Coherent structures*, Quart. Appl. Math., 45 (1987), pp. 561–571.
[30] R. Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Springer-Verlag, Berlin, 1997.
[31] S. Volkwein, *Optimal control of a phase-field model using proper orthogonal decomposition*, ZAMM Z. Angew. Math. Mech., 81 (2001), pp. 83–97.
[32] 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 (MTNS), 2004.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.