SIAM J. SCI. COMPUT.
Vol. 36, No. 3, pp. A1221–A1250




# ARE THE SNAPSHOT DIFFERENCE QUOTIENTS NEEDED IN THE PROPER ORTHOGONAL DECOMPOSITION?*

TRAIAN ILIESCU† AND ZHU WANG‡

**Abstract.** This paper presents a theoretical and numerical investigation of the following practical question: Should the time difference quotients (DQs) of the snapshots be used to generate the proper orthogonal decomposition (POD) basis functions? The answer to this question is important, since some published numerical studies use the time DQs, whereas other numerical studies do not. The criterion used in this paper to answer this question is the optimality of the convergence rate of the error of the reduced order model with respect to the number of POD basis functions. Since to the best of our knowledge a definition of the optimality of the convergence rate is not available, we propose one in Definition 3.1 in this paper. Two cases are considered: the no_DQ case, in which the snapshot DQs are not used, and the DQ case, in which the snapshot DQs are used. For each case, two types of POD bases are used: the $L^2$-POD basis, in which the basis is generated in the $L^2$-norm, and the $H^1$-POD basis, in which the basis is generated in the $H^1$-norm. The error estimates suggest that the convergence rates in the $C^0(L^2)$-norm and in the $C^0(H^1)$-norm are optimal for the DQ case, but suboptimal for the no_DQ case. The theoretical investigation, which uses two numerically validated assumptions on the POD projection error and the POD Ritz projection error, suggests the following convergence rates: In the DQ case, the error estimates are optimal in all norms (i.e., the $C^0(L^2)$-norm, the $C^0(H^1)$-norm, and the $L^2(H^1)$-norm) for both the $L^2$-POD basis and the $H^1$-POD basis. In the no_DQ case, however, the error estimates are suboptimal in the $C^0(L^2)$-norm for the $L^2$-POD basis and in the $C^0(H^1)$-norm for both the $L^2$-POD basis and the $H^1$-POD basis. Numerical tests are conducted on the heat equation and on the Burgers equation. The numerical results support the conclusions drawn from the theoretical error estimates. For both the heat equation and the Burgers equation, and for all norms and bases considered, the convergence rates for the DQ case are much higher than (and usually twice as high as) the corresponding convergence rates for the no_DQ case. Overall, the theoretical and numerical results strongly suggest that, in order to achieve optimal pointwise-in-time rates of convergence with respect to the number of POD basis functions, one should use the snapshot DQs.

**Key words.** proper orthogonal decomposition, reduced order modeling, error analysis

**AMS subject classifications.** 65M15, 65M60

**DOI.** 10.1137/130925141

## 1. Introduction.

This paper addresses the following question: *"Should the time difference quotients (DQs) of the snapshots be used in the generation of the proper orthogonal decomposition (POD) basis functions?"*

We emphasize that this is an important question. There are two schools of thought: one uses the DQs (see, e.g., [23, 24, 11, 12]), the other does not (see, e.g., [28, 27, 10, 33]).

To our knowledge, the first instance in which the snapshot DQs were incorporated in the generation of the POD basis functions was the pioneering paper of Kunisch and Volkwein [23]. In that report, the authors considered two types of errors for a general

*Submitted to the journal's Methods and Algorithms for Scientific Computing section June 17, 2013; accepted for publication (in revised form) April 1, 2014; published electronically June 17, 2014.
http://www.siam.org/journals/sisc/36-3/92514.html

†Department of Mathematics, Virginia Tech, 456 McBryde Hall, Blacksburg, VA 24061 (iliescu@vt.edu). This author's work was supported in part by the National Science Foundation through grant DMS-1025314.

‡Institute for Mathematics and its Applications, University of Minnesota, 354 Lind Hall, Minneapolis, MN 55455 (wangzhu@ima.umn.edu). This author's work was supported in part by the Institute for Mathematics and its Applications with funds provided by the National Science Foundation.

parabolic equation: the time discretization errors and the POD truncation errors. They argued that one needs to include the temporal DQs in the set of snapshots; otherwise, the error bound of the approximation of $u_t$ will contain an extra $\frac{1}{\Delta t^2}$ factor (see Remark 1 in [23]). Thus, the motivation for using the temporal DQs was purely theoretical. In numerical investigations, however, the authors reported contradictory findings: in [23], the use of the DQs did not improve the quality of the reduced order model (unless the snapshots were taken from a coarse temporal grid); in [17], however, it did. Kunisch and Volkwein used again the snapshot DQs when they considered the Navier–Stokes equations [24].

The snapshot DQs were also used in the *discrete empirical interpolation method* of Chaturantabut and Sorensen [11, 12] (which is a discrete variant of the *empirical interpolation method* [5]). The motivation in [11, 12], however, was different from that in [23]. Indeed, the authors considered in [11, 12] a general, nonlinear system of equations of the form $\mathbf{y}' = \mathbf{f}(\mathbf{y}, t)$ and used the *nonlinear* snapshots $\mathbf{f}(\mathbf{y}, t)$. They further noted (see page 48 in [12]) that, since $\mathbf{f}(\mathbf{y}, t) = \mathbf{y}'$ and $(\mathbf{y}^{n+1} - \mathbf{y}^n)/\Delta t \sim \mathbf{y}'$, this is similar to including the temporal DQs, as done in [23, 24].

To our knowledge, the first reports on POD analysis in which the DQs were not used were [28] for the heat equation and [27] for the Navier–Stokes equations. Chapelle, Gariah, and Sainte-Marie [10] used a different approach that did not utilize the DQs either. This approach employed the $L^2$ projection instead of the standard $H^1$ projection used in, e.g., [23, 24]. Further improvements to the approach used in [10] (as well as that used in [23, 24]) were made by Singler in [33].

From the above discussion, it is clear that the question whether the snapshot DQs should be included or not in the set of snapshots is important. To our knowledge, this question is still open. This report represents a first step in answering this question.

From a theoretical point of view, the only motivation for using the snapshot DQs was given in Remark 1 in [23]. The main point of this remark is the following: In the error analysis of the evolution equation, to approximate $u_t(\cdot, t^n)$, the time derivative of the exact solution $u$ evaluated at time $t^n$, one usually uses the DQ $\overline{\partial} u(\cdot, t^n) := \frac{u(\cdot,t^n) - u(\cdot,t^{n-1})}{\Delta t}$. To approximate the DQ $\overline{\partial} u(\cdot, t^n)$ in the POD space, one naturally uses the POD DQ $\overline{\partial} u_r(\cdot, t^n) := \frac{u_r(\cdot,t^n) - u_r(\cdot,t^{n-1})}{\Delta t}$, where $u_r$ is the POD reduced order model approximation. We assume that $u_r(\cdot, t^n)$ is an optimal approximation for $u(\cdot, t^n)$ and that $u_r(\cdot, t^{n-1})$ is an optimal approximation for $u(\cdot, t^{n-1})$, where the optimality is with respect to $r$ (the number of POD basis functions) and $\Delta t$ (the time step). Then, it would appear that, with respect to $\Delta t$, $\overline{\partial} u_r(\cdot, t^n)$ is a *suboptimal* approximation for $\overline{\partial} u(\cdot, t^n)$, because of the $\Delta t$ in the denominator of the two difference quotients.

Although the argument above, used in Remark 1 in [23] to motivate the inclusion of the snapshot DQs in the derivation of the POD basis, seems natural, we point out that this issue should be treated more carefully. Indeed, in the finite element approximation of parabolic equations, it is well known that the DQs $\overline{\partial} u_h(\cdot, t^n) := \frac{u_h(\cdot,t^n) - u_h(\cdot,t^{n-1})}{\Delta t}$ are actually *optimal* (with respect to $\Delta t$) approximations of the DQs $\overline{\partial} u(\cdot, t^n)$ (see, e.g., [25, 32]). Thus, since the POD and finite element approximations are similar (both use a Galerkin projection in the spatial discretization), one could question the validity of the argument used in Remark 1 in [23]. We emphasize that we are not claiming that the above argument is not valid in a POD setting; we are merely pointing out that a rigorous numerical analysis is needed before drawing any conclusions.

Our preliminary numerical studies indicate that not using the DQs does *not* yield suboptimal (with respect to $\Delta t$) convergence rates. For the heat equation (see sec-

TABLE 1
*Errors of the no_DQ and DQ approaches when $\Delta t$ varies.*

| $\Delta t$ | no_DQ | | | DQ | | |
|---|---|---|---|---|---|---|
| | r | $\mathcal{E}_{L^2(L^2)}$ | $\mathcal{E}_{L^2(H_1)}$ | r | $\mathcal{E}_{L^2(L^2)}$ | $\mathcal{E}_{L^2(H_1)}$ |
| 2.00e-01 | 6 | 3.71e-02 | 9.26e-01 | 6 | 3.71e-02 | 9.26e-01 |
| 1.00e-01 | 11 | 1.27e-02 | 5.81e-01 | 11 | 1.27e-02 | 5.81e-01 |
| 5.00e-02 | 21 | 2.99e-03 | 1.97e-01 | 21 | 2.99e-03 | 1.97e-01 |
| 2.50e-02 | 41 | 6.53e-04 | 3.81e-02 | 41 | 6.53e-04 | 3.81e-02 |
| 1.00e-02 | 59 | 1.03e-04 | 1.15e-02 | 88 | 1.03e-04 | 1.15e-02 |

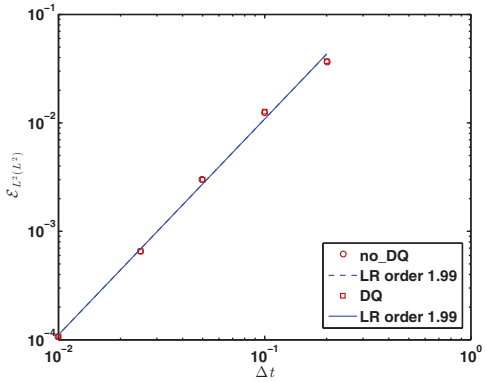


FIG. 1. *Heat equation, $L^2$-POD basis. Plots of the errors in the $L^2(L^2)$-norm with respect to the time step $\Delta t$ when the DQs are used (denoted by DQ) and when the DQs are not used (denoted by no_DQ).*

tion 4 for details regarding the numerical simulation), we monitor the rates of convergence with respect to $\Delta t$ for the POD reduced order model. To this end, we choose the discretization parameters to ensure that the time discretization error dominates the other error components. We consider two cases: the no_DQ case, in which the DQs are not used, and the DQ case, in which the DQs are used. The errors (defined in section 4) are listed in Table 1 and plotted in Figure 1 with associated linear regressions. Both no_DQ and DQ approaches yield an optimal approximation order $\mathcal{O}(\Delta t^2)$ in the $L^2$-norm.

The main goal of the paper is to answer the question in its title, i.e., whether the DQs should be used to generate the POD basis. The criterion used to answer this question is the optimality of the convergence rate of the error of the POD reduced order model with respect to $r$. Since to the best of our knowledge a definition of the optimality of the convergence rate is not available, we propose one in Definition 3.1. We consider the no_DQ case and the DQ case. For each case, we use two types of POD basis: the $L^2$-POD basis, in which the basis is generated in the $L^2$-norm, and the $H^1$-POD basis, in which the basis is generated in the $H^1$-norm.

The rest of the paper is organized as follows: In section 2, we sketch the derivation of the POD reduced order model for the heat equation. In section 3, we derive error estimates for the POD reduced order model in the no_DQ and the DQ cases for the heat equation. In the DQ case, we use the POD Ritz projection to prove optimal error estimates for both the $L^2$-POD basis and the $H^1$-POD basis in all norms. In the no_DQ case, we employ the $L^2$ projection proposed by Chapelle, Gariah, and Sainte-Marie [10] (see also [33]) to avoid the challenge posed by the POD approximation of

$u_t$. Using classical error analysis for Galerkin methods, however, yields suboptimal error estimates in the $C^0(L^2)$-norm for the $L^2$-POD basis and in the $C^0(H^1)$-norm for both the $L^2$-POD basis and the $H^1$-POD basis. Although all the error analysis in section 3 is done exclusively for the (linear) heat equation, in section 4, we present numerical results for both the heat equation and the (nonlinear) Burgers equation. For these two equations, the errors and the convergence rates in the no_DQ and the DQ cases are displayed for both the $L^2$-POD basis and the $H^1$-POD basis. The optimality of the convergence rates suggested by the error estimates in section 3 is confirmed by the numerical results. Furthermore, for all norms and bases considered, the convergence rates for the DQ case are much higher than (and usually twice as high as) the corresponding convergence rates for the no_DQ case. Finally, in section 5, we draw several conclusions regarding the theoretical and numerical results for the no_DQ and the DQ cases and we outline several future research directions.

**2. Proper orthogonal decomposition reduced order modeling.** In this section, we sketch the derivation of the standard POD Galerkin reduced order model for the heat equation. For a detailed presentation of reduced order modeling in general settings, the reader is referred to, e.g., [16, 22, 7, 2, 6, 38, 4].

Let $X := H_0^1(\Omega)$, where $\Omega$ is the computational domain. Let $u(\cdot, t) \in X, t \in [0, T]$ be the weak solution of the weak formulation of the heat equation with homogeneous Dirichlet boundary conditions:

$$(u_t, v)_{L^2} + \nu\,(\nabla u, \nabla v)_{L^2} = (f, v)_{L^2} \quad \forall v \in X. \tag{2.1}$$

Given the time instances $t_0, \ldots, t_N \in [0, T]$, we consider the following two ensembles of snapshots:

$$V^{no\text{-}DQ} := \operatorname{span}\left\{u(\cdot, t_0), \ldots, u(\cdot, t_N)\right\}, \tag{2.2}$$

$$V^{DQ} := \operatorname{span}\left\{u(\cdot, t_0), \ldots, u(\cdot, t_N), \overline{\partial} u(\cdot, t_1), \ldots, \overline{\partial} u(\cdot, t_N)\right\}, \tag{2.3}$$

where $\overline{\partial} u(\cdot, t_n) := \frac{u(\cdot, t_n) - u(\cdot, t_{n-1})}{\Delta t}$, $n = 1, \ldots, N$, are the time DQs. The two ensembles of snapshots correspond to the two cases investigated in this paper: (i) with the DQs not included in the snapshots (i.e., $V^{no\text{-}DQ}$), and (ii) with the DQs included in the snapshots (i.e., $V^{DQ}$). As pointed out in Remark 1 in [23], although the DQs in $V^{DQ}$ can be expressed as linear combinations of $u(\cdot, t_n)$, $n = 0, \ldots, N$, the ensemble of snapshots $V^{no\text{-}DQ}$ and $V^{DQ}$ yield *different* POD bases. This is clearly illustrated by Figures 2–3, which display POD basis functions for the heat equation and the Burgers equation, respectively. Details regarding the corresponding numerical simulations are given in section 4. Since the time step $\Delta t$ is used in the definition of the DQs, here we only briefly mention that the POD basis functions do not change when $\Delta t$ is varied around its original fine resolution value.

*Remark* 2.1. From a theoretical point of view, it would seem more natural to consider the continuous version of the POD method [24]. In this case, the snapshots $u_t(\cdot, t_n)$, $n = 1, \ldots, N$, could not be expressed anymore as linear combinations of $u(\cdot, t_n)$, $n = 0, \ldots, N$. Furthermore, the error analysis in section 3 would probably become clearer. We emphasize, however, that although driven by a theoretical analysis, the question whether the DQs should be used is a practical question that should be investigated numerically. Since the snapshots $u_t(\cdot, t_n)$, $n = 1, \ldots, N$, would have to be discretized in the numerical tests, for consistency reasons, we decided to exclusively use the DQs throughout the paper.

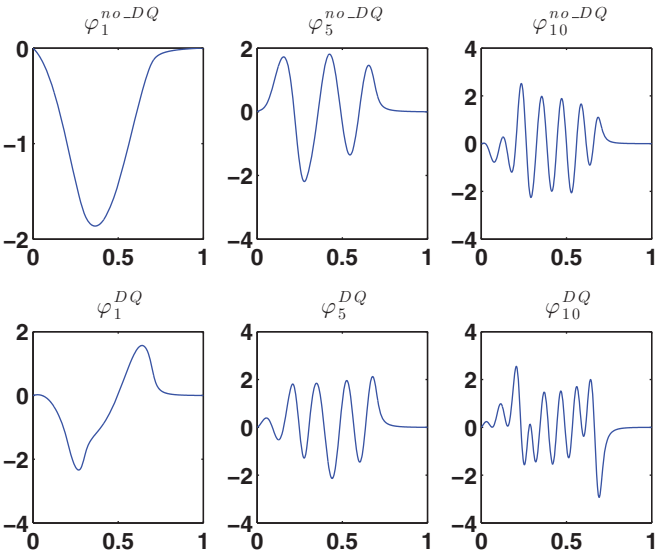


FIG. 2. *Heat equation, $L^2$-POD basis. Plots of the POD basis functions when the DQs are used (denoted by DQ) and when the DQs are not used (denoted by no_DQ).*

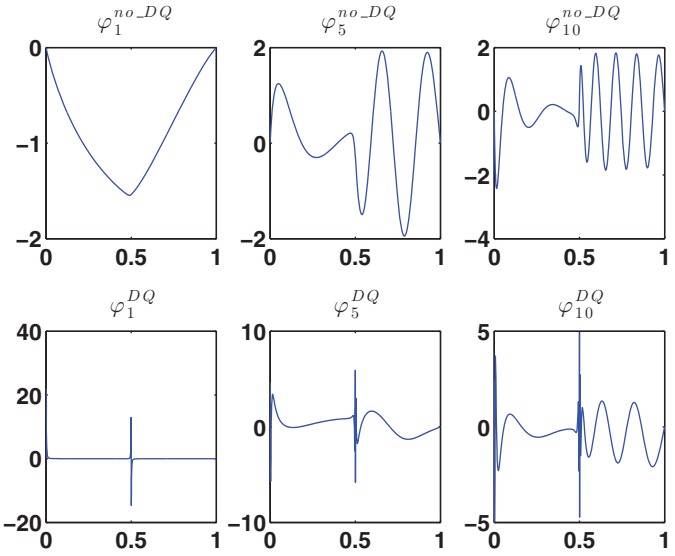


FIG. 3. *Burgers equation, $L^2$-POD basis. Plots of the POD basis functions when the DQs are used (denoted by DQ) and when the DQs are not used (denoted by no_DQ).*

We note that, in this study, the heat equation (2.1) is assumed to be in nondimensionalized form. Thus, in the definition of $V^{DQ}$ in (2.3), the snapshots $u(\cdot, t_i)$, $i = 0, \ldots, N$, and the DQs $\overline{\partial} u(\cdot, t_i)$, $i = 1, \ldots, N$, are both nondimensional quantities.

We also note that the solution of (2.1) might scale differently in space and time. In this case, to ensure a balanced scaling of the two terms in (2.1) (i.e., $u_t$ and $\Delta u$), one might consider, e.g., a nondimensionalization that uses the following scales: $U$ (the characteristic scale of the variable $u$) and $T$ (the characteristic time scale of

the system, e.g., the period in a time periodic system). In this case, the snapshots $u(\cdot,t_i),\ i=0,\dots,N$, and the DQs $\overline{\partial}u(\cdot,t_i),\ i=1,\dots,N$, used in $V^{DQ}$ would scale similarly.

To simplify the presentation, we denote both sets of snapshots (i.e., $V^{no\_DQ}$ and $V^{DQ}$) by

$$V = \operatorname{span}\{s_1, s_2, \dots, s_M\},$$

where $M = N+1$ when $V^{no\_DQ}$ is considered and $M = 2N+1$ when $V^{DQ}$ is considered. We use the specific notation (i.e., $V^{no\_DQ}$ or $V^{DQ}$) only when this is necessary. Let $\dim V = d$. Given a Hilbert space $\mathcal{H}$, the POD method seeks a low-dimensional basis $\{\varphi_1,\dots,\varphi_r\}$, with $r \le d$, which optimally approximates the input collection:

$$\min \frac{1}{M}\sum_{i=1}^{M}\left\| s_i - \sum_{j=1}^{r}\big(s_i\,,\,\varphi_j(\cdot)\big)_{\mathcal{H}}\,\varphi_j(\cdot)\right\|_{\mathcal{H}}^2, \tag{2.4}$$

subject to the conditions that $(\varphi_i,\varphi_j)_{\mathcal{H}} = \delta_{ij},\ 1 \le i,j \le r$. In order to solve (2.4), we consider the eigenvalue problem

$$K\,v = \lambda\,v, \tag{2.5}$$

where $K \in \mathbb{R}^{M\times M}$, with $K_{ij} = \frac{1}{M}(s_j,s_i)_{\mathcal{H}}$, is the snapshot correlation matrix, $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_d > 0$ are the positive eigenvalues, and $v_k$, $k=1,\dots,d$, are the associated eigenvectors. It can then be shown (see, e.g., [16, 22]), that the solution of (2.4) is given by $\varphi_k(\cdot) = \frac{1}{\sqrt{\lambda_k}}\sum_{j=1}^{M}(v_k)_j\,s_j,\ 1 \le k \le r$, where $(v_k)_j$ is the $j$th component of the eigenvector $v_k$.

DEFINITION 2.1. *The POD projection error is defined as*

$$\eta^{proj}(x,t) := u(x,t) - \sum_{j=1}^{r}\big(u(\cdot,t)\,,\,\varphi_j(\cdot)\big)_{\mathcal{H}}\,\varphi_j(x). \tag{2.6}$$

It can also be shown [23] that the *POD projection error* satisfies the following equalities:

$$\begin{aligned}\frac{1}{N+1}\sum_{i=0}^{N}\left\| u(\cdot,t_i) - \sum_{j=1}^{r}\big(u(\cdot,t_i)\,,\,\varphi_j(\cdot)\big)_{\mathcal{H}}\,\varphi_j(\cdot)\right\|_{\mathcal{H}}^2 \\ = \sum_{j=r+1}^{d}\lambda_j^{no\_DQ} \text{ if } V = V^{no\_DQ},\end{aligned} \tag{2.7}$$

$$\begin{aligned}&\frac{1}{2N+1}\sum_{i=0}^{N}\left\| u(\cdot,t_i) - \sum_{j=1}^{r}\big(u(\cdot,t_i)\,,\,\varphi_j(\cdot)\big)_{\mathcal{H}}\,\varphi_j(\cdot)\right\|_{\mathcal{H}}^2 \\ &\quad + \frac{1}{2N+1}\sum_{i=1}^{N}\left\| \overline{\partial}u(\cdot,t_i) - \sum_{j=1}^{r}\big(\overline{\partial}u(\cdot,t_i)\,,\,\varphi_j(\cdot)\big)_{\mathcal{H}}\,\varphi_j(\cdot)\right\|_{\mathcal{H}}^2 \\ &\quad = \sum_{j=r+1}^{d}\lambda_j^{DQ} \text{ if } V = V^{DQ}.\end{aligned} \tag{2.8}$$

*Remark* 2.2. In section 3, we only consider the semidiscretization of the POD reduced order model, i.e., the discretization in space (only with respect to the POD truncation), but not in time. Thus, the parameters $M$ (the number of snapshots) and $\Delta t$ (the time step, which is related to $M$) do not affect the theoretical analysis in section 3. The full discretization used in the numerical tests in section 4 does obviously depend on the parameters $M$ and $\Delta t$. To keep the length of the paper within reasonable limits, we decided to fix the parameters $M$, $\Delta t$, and $d$ (the dimension of the set of snapshots), and vary only the parameter $r$ (the number of modes used in the POD reduced order model). In sections 3 and 4 we show that this setting allows us to answer the question in the title of this paper.

For clarity, in what follows, we will denote by $C$ a generic constant that can depend on all the parameters in the system, except on $r$. In order to be able to prove pointwise-in-time error estimates in section 3, we make the following assumption.

*Assumption* 2.1. We assume that, for $i = 1, \ldots, N$, the POD projection error satisfies the following estimates:

$$\left\| u(\cdot,t_i) - \sum_{j=1}^{r} \big(u(\cdot,t_i)\,,\,\varphi_j(\cdot)\big)_{L^2}\,\varphi_j(\cdot) \right\|_{\mathcal{H}}^2 \leq C \sum_{j=r+1}^{d} \lambda_j^{no\_DQ} \qquad \text{if } \; V = V^{no\_DQ}\,, \tag{2.9}$$

$$\begin{aligned} &\left\| u(\cdot,t_i) - \sum_{j=1}^{r} \big(u(\cdot,t_i)\,,\,\varphi_j(\cdot)\big)_{L^2}\,\varphi_j(\cdot) \right\|_{\mathcal{H}}^2 \\ &\quad + \left\| \overline{\partial} u(\cdot,t_i) - \sum_{j=1}^{r} \big(\overline{\partial} u(\cdot,t_i)\,,\,\varphi_j(\cdot)\big)_{L^2}\,\varphi_j(\cdot) \right\|_{\mathcal{H}}^2 \\ &\leq C \sum_{j=r+1}^{d} \lambda_j^{DQ} \quad \text{if } \; V = V^{DQ}\,. \end{aligned} \tag{2.10}$$

Definition 2.2 ($L^2$-POD basis and $H^1$-POD basis). *If $\mathcal{H} = L^2$ in* (2.4), (2.6), (2.7), (2.8), (2.9), *and* (2.10), *then the resulting POD basis is called an $L^2$-POD basis. If $\mathcal{H} = H^1$ in* (2.4), (2.6), (2.7), (2.8), (2.9), *and* (2.10), *then the resulting POD basis is called an $H^1$-POD basis.*

*Remark* 2.3. Assumption 2.1 says that in the sums in (2.7) and (2.8) no individual term is much larger than the other terms. We also note that Assumption 2.1 would follow directly from the POD approximation property (2.7)–(2.8) if we dropped the $\frac{1}{M}$ factor in the snapshot correlation matrix $K$. In fact, this approach is used in, e.g., [24, 37]. We mention, however, that this would increase the magnitudes of the eigenvalues on the right-hand side (RHS) of the POD approximation property (2.7)–(2.8).

Next, we present a numerical investigation of Assumption 2.1. We consider the heat equation and the Burgers equation (see section 4 for details regarding the numerical simulations), and both $V = V^{no\_DQ}$ and $V = V^{DQ}$ cases. Since the numerical results for the $H^1$-POD basis are similar to those for the $L^2$-POD basis, we only present the latter. We calculate the ratios $(\mathcal{R}_1^{no\_DQ})_{r,i} = \frac{\|u(\cdot,t_i) - \sum_{j=1}^{r} (u(\cdot,t_i)\,,\,\varphi_j(\cdot))_{L^2}\,\varphi_j(\cdot)\|_{L^2}^2}{\sum_{j=r+1}^{d} \lambda_j^{no\_DQ}}$,

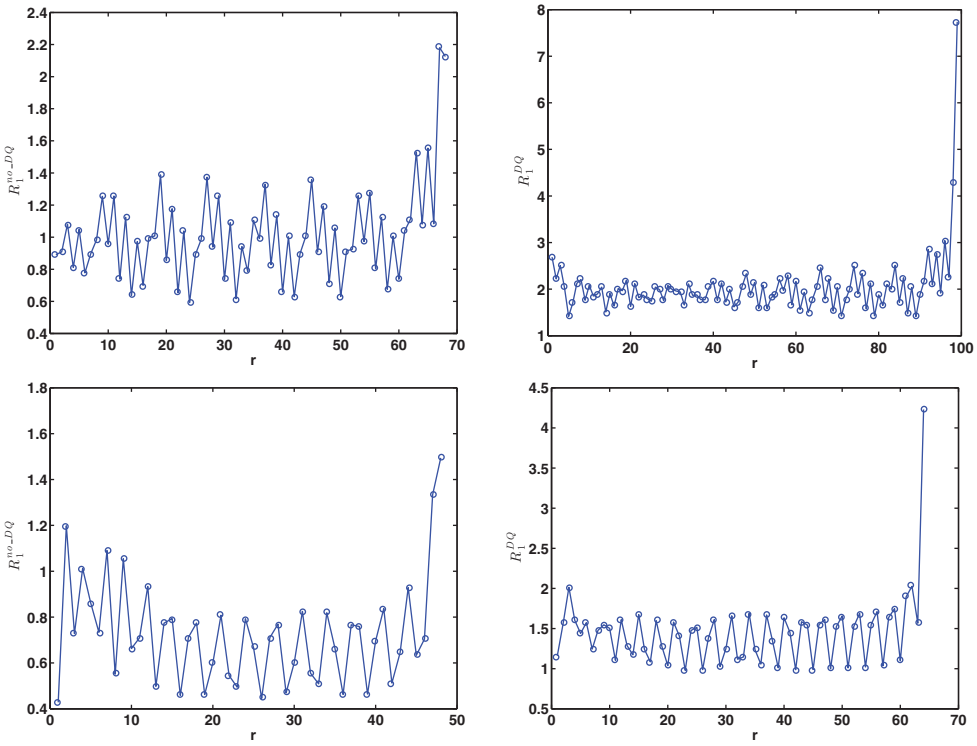


FIG. 4. *Heat equation, $L^2$-POD basis (top); Burgers equation, $L^2$-POD basis (bottom). Plots of the ratios $\mathcal{R}_1^{no\_DQ}$ and $\mathcal{R}_1^{DQ}$ at $t = 0.5$.*

$$(\mathcal{R}_1^{DQ})_{r,i} = \frac{\|u(\cdot,t_i) - \sum_{j=1}^{r} (u(\cdot,t_i)\,,\varphi_j(\cdot))_{L^2}\,\varphi_j(\cdot)\|_{L^2}^2}{\sum_{j=r+1}^{d} \lambda_j^{DQ}} + \frac{\|\overline{\partial} u(\cdot,t_i) - \sum_{j=1}^{r} (\overline{\partial} u(\cdot,t_i)\,,\varphi_j(\cdot))_{L^2}\,\varphi_j(\cdot)\|_{L^2}^2}{\sum_{j=r+1}^{d} \lambda_j^{DQ}}$$

to investigate assumptions (2.9) and (2.10), respectively. The ratios $(\mathcal{R}_1^{no\_DQ})_{r,i}$ and $(\mathcal{R}_1^{DQ})_{r,i}$ at $t = 0.5$ are plotted in Figure 4. Both ratios are uniformly bounded with respect to $r$. For the highest $r$ values, the denominators in the two ratios significantly decrease, which results in a natural increase of the two ratios. We also note that the highest values of the two ratios are attained at $t = 0$ and $t = 1$ for the heat equation, and at $t = 0$ for the Burgers equation. We believe that this behavior is due to the fact that the solution at $t = 0$ and $t = 1$ is not accurately captured by the first few POD basis functions.

Overall, the plots in Figure 4 show that Assumption 2.1 is valid in our setting.

In what follows, we will use the notation $X^r = \text{span}\{\varphi_1, \varphi_2, \ldots, \varphi_r\}$. To derive the POD reduced order model for the heat equation (2.1), we employ the Galerkin truncation, which yields the following approximation $u_r \in X^r$ of $u$:

$$u_r(x,t) := \sum_{j=1}^{r} a_j(t)\,\varphi_j(x). \tag{2.11}$$

Plugging (2.11) into (2.1) and multiplying by test functions in $X^r \subset X$ yields the *POD Galerkin reduced order model (POD-G-ROM)*:

$$(u_{r,t}, v_r)_{L^2} + \nu\,(\nabla u_r, \nabla v_r)_{L^2} = (f, v_r)_{L^2} \quad \forall\, v_r \in X^r. \tag{2.12}$$

The main advantage of the POD-G-ROM (2.12) over a straightforward finite element discretization of (2.1) is clear—the computational cost of the former is dramatically lower than that of the latter.

**3. Error estimates.** In this section, we prove estimates for the error $u - u_r$, where $u$ is the solution of the weak formulation of the heat equation (2.1) and $u_r$ is the solution of the POD-G-ROM (2.12). Error estimates for the POD reduced order modeling of general systems were derived in, e.g., [23, 24, 36, 18, 28, 30, 21, 13, 35, 1, 31, 15]. We perform the error analysis only for the *semidiscretization* of the POD-G-ROM (2.12). In fact, in this semidiscretization, we only consider the error component corresponding to the POD truncation. Of course, in practical numerical simulations, the semidiscretization also has a spatial component (e.g., due to the finite element discretization). Furthermore, when considering the full discretization, the error also has a time discretization component (due to the time stepping algorithm). All these error components should be included in a rigorous error analysis of the discretization of the POD-G-ROM (2.12) (see, e.g., [28, 19, 20]). For clarity of presentation, however, we only consider the error component corresponding to the POD truncation. In what follows, we will show that this is sufficient for answering the question asked in the title of this paper.

In our theoretical analysis, we consider two cases, depending on the type of snapshots used in the derivation of the POD basis: Case I: $V = V^{DQ}$ (i.e., with the DQs); and Case II: $V = V^{no\text{-}DQ}$ (i.e., without the DQs). For each case, we consider both the $L^2$-POD basis and the $H^1$-POD basis. The main goal of this section is to investigate whether Case I, Case II, or both Case I and Case II, yield *error estimates that are optimal with respect to* $r$.

Since, to the best of our knowledge, a definition for the optimality of the POD error estimates is not available, we propose such a definition below.

DEFINITION 3.1 (optimal POD error estimate). *The POD-G-ROM error estimate is optimal with respect to* $r$ *if the following inequalities hold:*

$$\|u - u_r\|_{L^2} \leq C \|\eta^{proj}\|_{L^2}, \tag{3.1}$$

$$\|\nabla(u - u_r)\|_{L^2} \leq C \|\nabla\eta^{proj}\|_{L^2}, \tag{3.2}$$

*where* $\eta^{proj}$ *is the POD projection error defined in* (2.6).

*Remark* 3.1. We note that, since in this section we only consider the error component corresponding to the POD truncation, the optimality in Definition 3.1 is with respect to $r$. Of course, when the other error components are considered, the optimality with respect to other discretization parameters (e.g., $h$ and $\Delta t$) should also be considered.

We also note that other definitions for the optimality of the POD error estimates are possible. For example, one could replace inequalities (3.1) and (3.2) in Definition 3.1 with

$$\|u - u_r\|_{L^2} \leq C \inf_{v_r \in X^r} \|v - v_r\|_{L^2}, \tag{3.3}$$

$$\|\nabla(u - u_r)\|_{L^2} \leq C \inf_{v_r \in X^r} \|\nabla(v - v_r)\|_{L^2}, \tag{3.4}$$

respectively. One advantage of the alternative definition in (3.3) and (3.4) is that it resembles the standard finite element error estimates in Céa's lemma [34] more than Definition 3.1. One potential drawback of the alternative definition, however, is that the evaluation of the RHS of (3.3) or (3.4) might not be straightforward. For example,

when the $L^2$-POD basis is used, $\inf_{v_r \in X^r} \|v - v_r\|_{L^2} = \|\eta^{proj}\|_{L^2}$, so the RHS of (3.3) can be easily evaluated by using (2.6). To evaluate the factor $\inf_{v_r \in X^r} \|\nabla(v - v_r)\|_{L^2}$ on the RHS of (3.4), however, one would have to solve a minimization problem at each time step. Since the optimality of the POD error estimates is monitored in the numerical experiments in section 4, in the remainder of the paper we exclusively use Definition 3.1.

*Remark* 3.2 (optimal error estimate: $L^2$-POD basis vs. $H^1$-POD basis). When the $L^2$-POD basis is used, $\|\eta^{proj}\|_{L^2}$ and $\|\nabla\eta^{proj}\|_{L^2}$ scale differently with respect to $r$: The scaling of $\|\eta^{proj}\|_{L^2}$ is given by the POD projection error equalities (2.9)–(2.10) in Assumption 2.1. The scaling of $\|\nabla\eta^{proj}\|_{L^2}$, however, is *not* given by (2.9)–(2.10). There are several approaches that one can use to derive an estimate for $\|\nabla\eta^{proj}\|_{L^2}$. For example, one can employ a brute force approach and use the fact that the projection error lives in a finite dimensional space, i.e., the space spanned by the snapshots. Using an inverse estimate similar to that presented in Lemma 3.2 but for the entire space of snapshots (of dimension $d$), we get the following estimate:

$$\|\nabla\eta^{proj}\|_{L^2} \leq C_{inv}(d)\, \|\eta^{proj}\|_{L^2}\,, \tag{3.5}$$

where $C_{inv}(d)$ is the constant in the inverse estimate in Lemma 3.2. Following the discussion in Remark 3.3, we conclude that the scaling of $\|\nabla\eta^{proj}\|_{L^2}$ is of lower order with respect to $r$ than the scaling of $\|\eta^{proj}\|_{L^2}$. Thus, if the error analysis yields estimates of the form

$$\|u - u_r\|_{L^2} \leq C\, \|\nabla\eta^{proj}\|_{L^2}\,, \tag{3.6}$$

then these estimates will be called *suboptimal* with respect to $r$.

To derive an estimate for $\|\nabla\eta^{proj}\|_{L^2}$, one can alternatively employ the approach used in Theorem 2 in [33] (see also Lemma 3.2 in [20]). This approach yields an estimate that is sharper than the estimate in (3.5). We note, however, that even in this case, the scaling of $\|\nabla\eta^{proj}\|_{L^2}$ is of lower order with respect to $r$ than the scaling of $\|\eta^{proj}\|_{L^2}$. Thus, the conclusions regarding the optimality of the error estimates remain unchanged: If the error analysis yields estimates such as (3.6), then these estimates will be called suboptimal with respect to $r$.

When the $H^1$-POD basis is used, the POD approximation property in Assumption 2.1 shows that $\|\eta^{proj}\|_{L^2}$ and $\|\nabla\eta^{proj}\|_{L^2}$ have the same scaling with respect to $r$, since they are both bounded by $\sqrt{\sum_{j=r+1}^{d} \lambda_j^{no\_DQ}}$ (if $V = V^{no\_DQ}$; see (2.9)) or by $\sqrt{\sum_{j=r+1}^{d} \lambda_j^{DQ}}$ (if $V = V^{DQ}$; see (2.10)).

Thus, if the error analysis yields estimates of the form

$$\|u - u_r\|_{L^2} \leq C\, \|\nabla\eta^{proj}\|_{L^2}\,, \tag{3.7}$$

then these estimates will be called *optimal* with respect to $r$. We emphasize, however, that if the constant $C$ in (3.7) does depend on $r$ in a suboptimal way (e.g., it increases when $r$ increases), then the estimate (3.7) will be called *suboptimal* with respect to $r$.

We introduce some notation and we list several results that will be used throughout this section. The first result is a POD inverse estimate, which was proved in Lemma 2 and Remark 2 in [23] and in Lemma 3.1 and Remark 3.2 in [19]. Let $M_r \in \mathbb{R}^{r\times r}$ with $M_{ij} = (\varphi_j, \varphi_i)_{L^2}$ be the POD mass matrix and $S_r \in \mathbb{R}^{r\times r}$ with

$S_{ij} = (\nabla\varphi_j, \nabla\varphi_i)_{L^2}$ be the POD stiffness matrix. Let $\|\cdot\|_2$ denote the matrix 2-norm.

LEMMA 3.2 (POD inverse estimate). *For all $v_r \in X^r$, the following POD inverse estimates hold:*

$$\|\nabla v_r\|_{L^2} \le C_{inv}^{L^2}(r)\,\|v_r\|_{L^2} \qquad \textit{for the } L^2\textit{-POD basis, and} \tag{3.8}$$

$$\|\nabla v_r\|_{L^2} \le C_{inv}^{H^1}(r)\,\|v_r\|_{L^2} \qquad \textit{for the } H^1\textit{-POD basis,} \tag{3.9}$$

*where $C_{inv}^{L^2}(r) := \sqrt{\|S_r\|_2}$ and $C_{inv}^{H^1}(r) := \sqrt{\|M_r^{-1}\|_2}$.*

*Remark* 3.3 (POD inverse estimate scalings). Since the $r$ dependency of the error estimates presented in this section will be carefully monitored, we try to get some insight into the scalings of the constants $C_{inv}^{L^2}(r)$ and $C_{inv}^{H^1}(r)$ in (3.8) and (3.9), respectively, i.e., the scalings of $\|S_r\|_2$ and $\|M_r^{-1}\|_2$ with respect to $r$.

We note that, since the POD basis significantly varies from test case to test case, it would be difficult to derive general scalings of $\|S_r\|_2$ and $\|M_r^{-1}\|_2$. When the underlying system is *homogeneous* (i.e., invariant to spatial translations), however, the $L^2$-POD basis is identical to the Fourier basis (see, e.g., section 3.3.1 in [16]). In this case, one can derive the scalings of $\|S_r\|_2$. Thus, to get some insight into the scaling of $C_{inv}^{L^2}(r)$ in a general setting, we first consider the homogeneous case. Without loss of generality, we assume that the domain is $[0,1] \subset \mathbb{R}^1$ and that the boundary conditions are homogeneous Dirichlet. In this case, the Fourier basis functions are given by $\varphi_j(x) = \sin(j\,\pi\,x)$. The matrix $S_r$ is diagonal and its diagonal entries are given by $S_{jj} = \int_0^1 (j\,\pi)^2\,\cos^2(j\,\pi\,x)\,dx = \frac{1}{2}\,(j\,\pi)^2$. Since the POD stiffness matrix $S_r$ is symmetric, its matrix 2-norm is given by $\|S_r\|_2 = \lambda_{\max}$, where $\lambda_{\max}$ is the largest eigenvalue of $S_r$. Thus, we have

$$\|S_r\|_2 = \frac{1}{2}\,(r\,\pi)^2 = \mathcal{O}(r^2)\,. \tag{3.10}$$

Thus, we conclude that, when the underlying system is homogeneous, the 2-norm of the POD stiffness matrix $S_r$ scales as $\mathcal{O}(r^2)$.

Next, we use the numerical tests in section 4 to get some insight into the scalings for $C_{inv}^{L^2}(r)$ and $C_{inv}^{H^1}(r)$ for general, nonhomogeneous systems. For the heat equation and the Burgers equation (see section 4 for details regarding the numerical simulations), we monitor the scalings of $\|S_r\|_2$ and $\|M_r^{-1}\|_2$. We consider Case I ($V = V^{DQ}$) and Case II ($V = V^{no\text{-}DQ}$). For each case, we also consider both the $L^2$-POD basis and the $H^1$-POD basis. The scalings are plotted in Figure 5. Although the theoretical scaling (3.10) predicted for the homogeneous flow fields is not recovered, all four plots show a clear increase of $\|S_r\|_2$ and $\|M_r^{-1}\|_2$ (and $C_{inv}^{L^2}(r)$ and $C_{inv}^{H^1}(r)$, respectively) with respect to $r$. We also note that, for the Burgers equation, the scalings reach a plateau for large $r$ values. We believe that this is due to the fact that the spatial and temporal discretization error components become dominant for large $r$ values.

The theoretical scaling in (3.10) valid for the homogeneous case and the numerical results in Figure 5 for the (general, nonhomogeneous) heat equation and Burgers equation strongly suggest that $C_{inv}^{L^2}(r)$ and $C_{inv}^{H^1}(r)$ increase when $r$ increases.

After these preliminaries, we are ready to derive the error estimates. The error analysis will proceed along the same lines as the error analysis of the finite element semidiscretization [14, 26, 34]. The main difference between the two settings is that the finite element approximation property is *global* [14, 26], whereas the POD projection

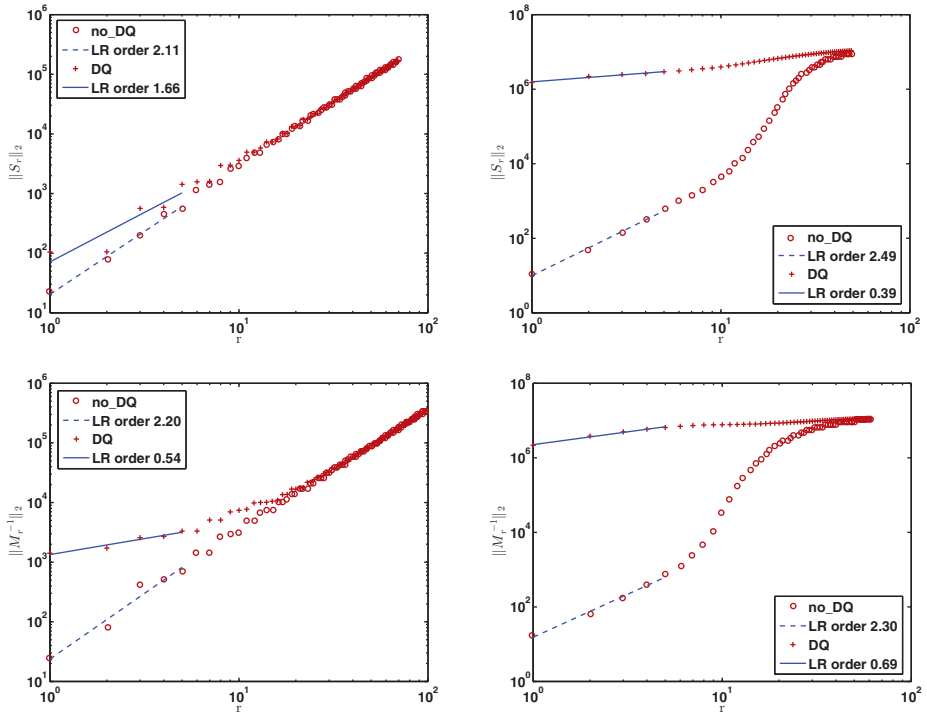


FIG. 5. *Heat equation (left), Burgers equation (right); $L^2$-POD basis top, $H^1$-POD basis bottom. Plots of the scalings of $\|S_r\|_2$ and $\|M_r^{-1}\|_2$ with respect to $r$ when the DQs are used (denoted by DQ) and when the DQs are not used (denoted by no_DQ).*

error equality is *local*, i.e., it is only valid at the time instances at which the snapshots were taken (see (2.7)–(2.10)). Thus, in order to be able to use the POD projection error estimates (2.7)–(2.10), in what follows we assume that the final error estimates for the semidiscretization are considered only at the time instances $t_i$, $i = 1, \ldots, N$. We also consider both the $L^2$-POD basis and the $H^1$-POD basis. For clarity, we first present the analysis for the $L^2$-POD basis, and then highlight the differences in the analysis for the $H^1$-POD basis.

We start by considering the error equation:

$$(e_t, v_r)_{L^2} + \nu\,(\nabla e, \nabla v_r)_{L^2} = 0 \qquad \forall\, v_r \in X^r, \tag{3.11}$$

where $e := u - u_r$ is the error. The error is split into two parts:

$$e = u - u_r = (u - w_r) - (u_r - w_r) = \eta - \phi_r, \tag{3.12}$$

where $w_r$ is an *arbitrary* function in $X^r$, $\eta := u - w_r$, and $\phi_r := u_r - w_r$. Using this decomposition in the error equation (3.11), we get

$$(\phi_{r,t}, v_r)_{L^2} + \nu\,(\nabla \phi_r, \nabla v_r)_{L^2} = (\eta_t, v_r)_{L^2} + \nu\,(\nabla \eta, \nabla v_r)_{L^2}. \tag{3.13}$$

The analysis proceeds by using (3.13) to show that

$$\|\phi_r\|_{L^2} \leq C\,\|\eta\|_{L^2}. \tag{3.14}$$

Using the triangle inequality, one then gets

$$\|e\|_{L^2} \le \|\eta\|_{L^2} + \|\phi_r\|_{L^2} \le (1+C)\,\|\eta\|_{L^2}. \tag{3.15}$$

Since $w_r$ was chosen arbitrarily, we choose $w_r = \sum_{j=1}^{r}(u\,,\,\varphi_j)_{L^2}\,\varphi_j$ in (3.15), use (2.6), and get the following error estimate:

$$\|e\|_{L^2} \le (1+C)\,\|\eta^{proj}\|_{L^2}\,. \tag{3.16}$$

Using Definition 3.1 and Remark 3.2, we conclude that the error estimate (3.16) is optimal.

In the remainder of this section, we investigate whether optimal error estimates can be obtained with or without including the DQs in the set of snapshots. To this end, in section 3.1 we consider the case in which the DQs are included in the set of snapshots (i.e., $V = V^{DQ}$). Then, in section 3.2 we consider the case in which the DQs are not included in the set of snapshots (i.e., $V = V^{no\_DQ}$).

**3.1. Case I ($V = V^{DQ}$).** The standard approach used to prove error estimates in this case is to use the Ritz projection [23, 24, 19, 20]. This is also the standard approach used in the finite element context [39, 14, 34, 26].

We choose $w_r := R_r(u)$ in (3.12), where $R_r(u)$ is the *Ritz projection* of $u$:

$$\big(\nabla(u - R_r(u)), \nabla v_r\big)_{L^2} = 0 \qquad \forall\, v_r \in X^r. \tag{3.17}$$

To emphasize that we are using the Ritz projection, in the remainder of section 3.1 we will use the notation $\eta^{Ritz} = \eta = u - R_r(u)$. Using (3.17), (3.13) becomes

$$(\phi_{r,t}, v_r)_{L^2} + \nu\,(\nabla\phi_r, \nabla v_r)_{L^2} = (\eta_t^{Ritz}, v_r)_{L^2} + \nu\,\underbrace{(\nabla\eta^{Ritz}, \nabla v_r)_{L^2}}_{\to 0}\,, \tag{3.18}$$

where $\eta_t^{Ritz} = u_t - R_r(u_t)$. It is the cancellation of the last term on the RHS of (3.18) that yields optimal error estimates. We let $v_r := \phi_r$ in (3.18), and then we apply the Cauchy–Schwarz inequality to the remaining term on the RHS:

$$\frac{1}{2}\,\frac{d}{dt}\|\phi_r\|_{L^2}^2 + \nu\,\|\nabla\phi_r\|_{L^2}^2 \le \|\eta_t^{Ritz}\|_{L^2}\,\|\phi_r\|_{L^2}. \tag{3.19}$$

We rewrite the first term on the left-hand side (LHS) of (3.19) as $\frac{1}{2}\,\frac{d}{dt}\|\phi_r\|_{L^2}^2 = \|\phi_r\|_{L^2}\,\frac{d}{dt}\|\phi_r\|_{L^2}$ and we apply the Poincaré–Friedrichs inequality to the second term on the LHS of (3.19):

$$\nu\,\|\nabla\phi_r\|_{L^2}^2 \ge C\,\nu\,\|\phi_r\|_{L^2}^2. \tag{3.20}$$

We note that the Poincaré–Friedrichs inequality $C\|v\|_{L^2}^2 \le \|\nabla v\|_{L^2}^2$ holds for every function $v$ in the *continuous* space $H_0^1(\Omega)$, and, in particular, for $\phi_r \in X^r \subset X = H_0^1(\Omega)$ (see (3) in [23]). Thus, the constant $C$ in (3.20) does not depend on $r$. Thus, (3.19) becomes

$$\frac{d}{dt}\|\phi_r\|_{L^2} + C\,\nu\,\|\phi_r\|_{L^2} \le \|\eta_t^{Ritz}\|_{L^2}. \tag{3.21}$$

Using Gronwall's lemma in (3.21), we get for $0 < t \le T$

$$\|\phi_r(t)\|_{L^2} \le e^{-C\,\nu\,t}\,\|\phi_r(0)\|_{L^2} + \int_0^t e^{-C\,\nu\,(t-s)}\,\|\eta_t^{Ritz}(s)\|_{L^2}\,ds. \tag{3.22}$$

Using (3.15), the first term on the RHS of (3.22) can be estimated as $\|\phi_r(0)\|_{L^2} \le$

$\|e(0)\|_{L^2} + \|\eta^{Ritz}(0)\|_{L^2}$. Thus, (3.22) becomes

$$\|\phi_r(t)\|_{L^2} \leq e^{-C\,\nu\,t}\left(\|e(0)\|_{L^2} + \|\eta^{Ritz}(0)\|_{L^2}\right) + \int_0^t e^{-C\,\nu\,(t-s)}\,\|\eta_t^{Ritz}(s)\|_{L^2} ds. \tag{3.23}$$

Applying the triangle inequality, just as in (3.15), we get

$$\begin{aligned}\|e(t)\|_{L^2} \leq \|\eta^{Ritz}(t)\|_{L^2} + e^{-C\,\nu\,t}\left(\|e(0)\|_{L^2} + \|\eta^{Ritz}(0)\|_{L^2}\right)\\ + \int_0^t e^{-C\,\nu\,(t-s)}\,\|\eta_t^{Ritz}(s)\|_{L^2}\,ds.\end{aligned} \tag{3.24}$$

We make the following assumption.

*Assumption* 3.1. We assume that the POD Ritz projection error $\eta^{Ritz}$ satisfies optimal error estimates with respect to $r$ in the $L^2$-norm:

$$\|\eta^{Ritz}\|_{L^2} \leq C\,\|\eta^{proj}\|_{L^2}\,, \tag{3.25}$$

$$\|\eta_t^{Ritz}\|_{L^2} \leq C\,\|\overline{\partial}\eta^{proj}\|_{L^2}\,, \tag{3.26}$$

$$\|\nabla\eta^{Ritz}\|_{L^2} \leq C\,\|\nabla\eta^{proj}\|_{L^2}\,, \tag{3.27}$$

where $\overline{\partial}\eta^{proj}(x,t) := \overline{\partial}u(x,t) - \sum_{j=1}^r (\overline{\partial}u(\cdot,t)\,,\,\varphi_j(\cdot))_{L^2}\,\varphi_j(x)\,.$

Using (3.25) and (3.26) in (3.24), we conclude that the POD error estimate in the $L^2$-norm is optimal. Furthermore, since no reference has been made to the specific way in which the POD basis was calculated, we conclude that the error estimate is optimal with respect to $r$ both for the $L^2$-POD basis and the $H^1$-POD basis.

*Remark* 3.4. Using estimate (3.26) in the integral on the RHS of (3.24) yields a bound that depends on $\|\overline{\partial}\eta^{proj}\|_{L^2}$ instead of $\|\eta^{proj}\|_{L^2}$. Thus, it might seem that, according to Definition 3.1 and Remark 3.2, we do not get optimal error estimates. We emphasize, however, that since both $\|\overline{\partial}\eta^{proj}\|_{L^2}$ and $\|\eta^{proj}\|_{L^2}$ are bounded by $\sum_{j=r+1}^d \lambda_j^{DQ}$ when $V = V^{DQ}$, the error estimate resulting from (3.24) is, in fact, optimal.

**3.1.1. The POD Ritz projection.** In the finite element context, both $\|\eta^{Ritz}\|_{L^2}$ and $\|\eta_t^{Ritz}\|_{L^2}$ are optimal (with respect to the mesh size $h$). In the POD context, however, the optimality (with respect to $r$) is not that clear. In this subsection, we present a theoretical and numerical investigation of the POD Ritz projection.

To the best of the authors' knowledge, the first error estimates for the POD Ritz projection were derived in the pioneering paper of Kunisch and Volkwein [23]. Sharper error estimates were recently proven by Singler [33]. We summarize below the results in [23, 33]. For clarity, we present the $L^2$-POD basis case, and only highlight the differences in the $H^1$-POD basis case.

The main result in [23] regarding the POD Ritz projection is Lemma 3 (see also (10) and (11) in [23]), which, in our notation, states the following:

$$\|\nabla\eta^{Ritz}\|_{L^2} \leq C\,\sqrt{\|S_d\|_2}\,\|\eta^{proj}\|_{L^2}\,. \tag{3.28}$$

For clarity, we have not included in (3.28) the constants that do not depend on $r$. Inequality (3.28) suggests that assumption (3.27) is true (see (3.2) and (3.5)).

We emphasize that Lemma 3 in [23] does not include any bounds for $\|\eta^{Ritz}\|_{L^2}$. This is in clear contrast with the finite element context, in which $\|\eta^{Ritz}\|_{L^2}$ is estimated by the usual duality argument (the Aubin–Nitsche "trick"; see, e.g., [34]). Using a

duality argument, however, is challenging in the POD context, since any auxiliary dual problem would not necessarily inherit the POD approximation property (2.7)–(2.10). To the best of the authors' knowledge, such a duality argument has never been used in a POD context. We emphasize that not being able to use a duality argument in the Ritz projection to get error estimates that are optimal with respect to $r$ has significant consequences in the error analysis. Indeed, in the proof of Theorem 7 in [23] (the error estimate for the backward Euler time discretization), to estimate the $\|\eta^{Ritz}\|_{L^2}$ error component in (27a) and (27b), the authors use the $\|\nabla\eta^{Ritz}\|_{L^2}$ estimate given in Lemma 3 and the Poincaré–Friedrichs inequality given in (3). Since the Poincaré–Friedrichs constant does not depend on $r$, we conclude that $\|\eta^{Ritz}\|_{L^2}$ and $\|\nabla\eta^{Ritz}\|_{L^2}$ scale similarly with respect to $r$. This, in turn, suggests that $\|\eta^{Ritz}\|_{L^2}$ is suboptimal with respect to $r$ (see Definition 3.1 and Remark 3.2). We note that the same approach (i.e., Lemma 4 and the Poincaré–Friedrichs inequality) is used in [23] to estimate the DQ approximation of $\|\eta_t^{Ritz}\|_{L^2}$ (see the two inequalities above (29a)).

To summarize, the analysis in [23] suggests that, when the $L^2$-POD basis is used, assumption (3.27) holds, but assumptions (3.25) and (3.26) do not hold. When the $H^1$-POD basis is used, however, the analysis in [23] shows that assumption (3.27) holds (see Lemma 3 in [23]) and suggests that (3.25) and (3.26) hold as well.

The approach used in Theorem 2 in [33] (see also Lemma 3.2 in [20]) yields error estimates that are sharper than those in (3.28). We note, however, that although sharper, the estimates in [33] cannot be easily used with the definition of optimality of POD error estimates employed in this paper (see Definition 3.1). Furthermore, even if we used the estimates in [33], they would yield the same conclusions regarding the Ritz projection as the conclusions drawn from estimate (3.28).

Next, we present a numerical investigation of Assumption 3.1. We consider the heat equation and the Burgers equation (see section 4 for details regarding the numerical simulations), and both $V = V^{no\_DQ}$ and $V = V^{DQ}$ cases. Since the numerical results for the $H^1$-POD basis are similar to those for the $L^2$-POD basis, we only present the latter. We calculate the ratios $(\mathcal{R}_2^{DQ})_{r,i} = \frac{\|\eta^{Ritz}(\cdot,t_i)\|_{L^2}}{\|\eta^{proj}(\cdot,t_i)\|_{L^2}}$, $(\mathcal{R}_3^{DQ})_{r,i} = \frac{\|\eta_t^{Ritz}\|_{L^2}}{\|\overline{\partial}\eta^{proj}\|_{L^2}}$, and $(\mathcal{R}_4^{DQ})_{r,i} = \frac{\|\nabla\eta^{Ritz}(\cdot,t_i)\|_{L^2}}{\|\nabla\eta^{proj}(\cdot,t_i)\|_{L^2}}$ to investigate assumptions (3.25), (3.26), and (3.27), respectively. These three ratios at $t = 0.5$ are plotted in Figure 6. All three ratios are uniformly bounded with respect to $r$. We note that the ratios display relatively higher values around $r = 20$. We also note that, for $(\mathcal{R}_2^{DQ})_{r,i}$, the highest values are attained at $t = 0$ and $t = 1$ for the heat equation, and at $t = 0$ for the Burgers equation. We believe that this behavior is due to the fact that the solution at $t = 0$ and $t = 1$ is not accurately captured by the first few POD basis functions.

Overall, the plots in Figure 6 provide numerical support for Assumption 3.1 in our setting.

**3.2. Case II ($V = V^{no\_DQ}$).** This approach was used in [10, 33]. The motivation for this approach is the following: In Case I ($V = V^{DQ}$), the first term on the RHS of (3.18), $(\eta_t, v_r)_{L^2}$, yields a term $\|\eta_t\|_{L^2}$ that stays in all the subsequent inequalities, including the final error estimate (3.24). Chapelle, Gariah, and Sainte-Marie proposed in [10] a different approach that eliminated the $(\eta_t, v_r)_{L^2}$ term in (3.18). Thus, instead of using the Ritz projection (as in Case I), they used the $L^2$ projection. That is, they chose $w_r := P_r(u)$, where $P_r(u)$ is the $L^2$ *projection* of $u$, given by

$$\big(u - P_r(u), v_r\big)_{L^2} = 0 \qquad \forall\, v_r \in X^r. \tag{3.29}$$

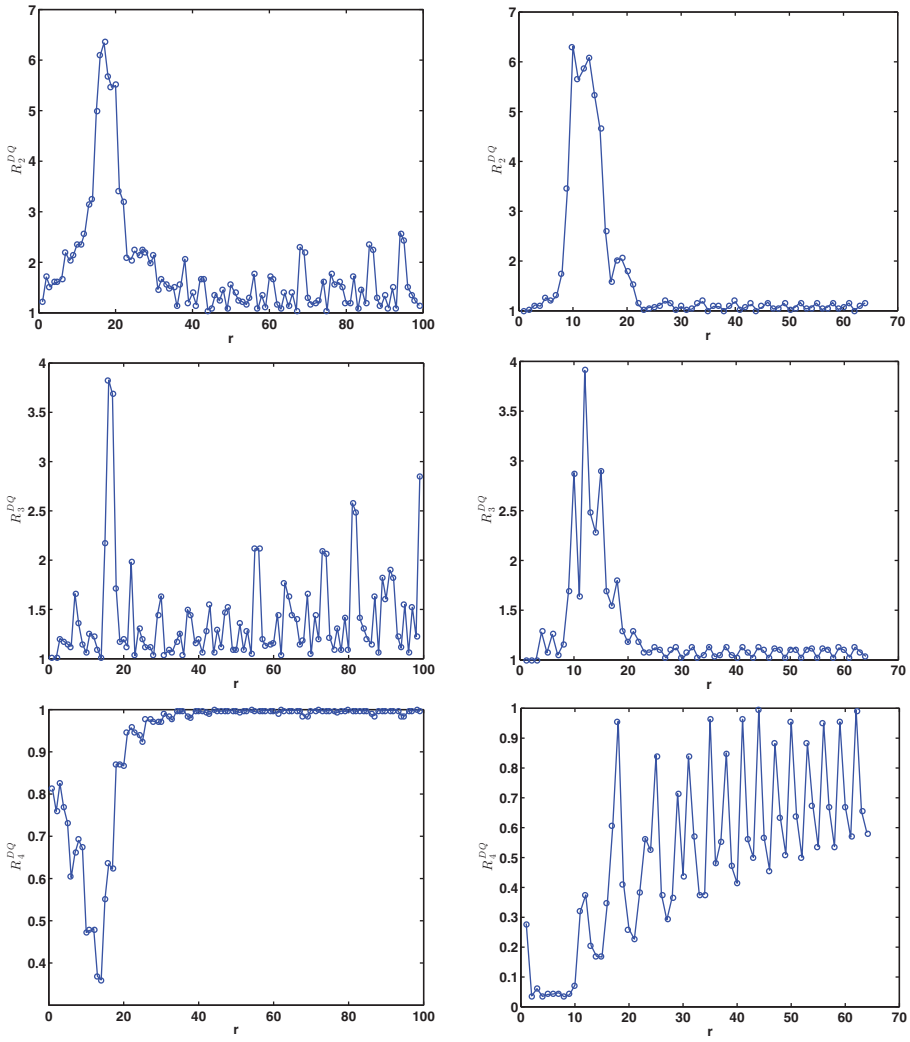


FIG. 6. *Heat equation, $L^2$-POD basis (left column); Burgers equation, $L^2$-POD basis (right column). Plots of the ratios $\mathcal{R}_2^{DQ}$ (top), $\mathcal{R}_3^{DQ}$ (middle), and $\mathcal{R}_4^{DQ}$ (bottom) at $t = 0.5$.*

To emphasize that we are using the $L^2$ projection, in the remainder of section 3.2 we will use the notation $\eta^{L^2} = \eta = u - P_r(u)$.

**3.2.1. The POD $L^2$ projection.** Since the $L^2$ projection plays such an important role in this section, we summarize below some of its properties.

LEMMA 3.3. *If the $L^2$-POD basis is used, then the following estimates hold:*

$$\|\eta^{L^2}\|_{L^2} = \|\eta^{proj}\|_{L^2}\,, \tag{3.30}$$

$$\|\nabla\eta^{L^2}\|_{L^2} \leq C_{inv}^{L^2}(d)\,\|\eta^{proj}\|_{L^2}. \tag{3.31}$$

*If the $H^1$-POD basis is used, then the following estimates hold:*

$$\|\eta^{L^2}\|_{L^2} \leq \|\eta^{proj}\|_{L^2}\,, \tag{3.32}$$

$$\|\nabla\eta^{L^2}\|_{L^2} \leq C_{inv}^{H^1}(d)\,\|\eta^{proj}\|_{L^2}. \tag{3.33}$$

*Proof.* Equality (3.30) follows by noticing that, for the $L^2$-POD basis, the POD $L^2$ projection error $\eta^{L^2}$ is exactly the POD projection error defined in (2.6). Inequality (3.31) follows from (3.30) and Lemma 3.2.

Inequality (3.32) follows along the same lines as Lemma 2 in [23]: Choosing $v_r = \sum_{j=1}^{r} (u, \varphi_j)_{H^1} \varphi_j \in X^r$ and using the Cauchy–Schwarz inequality, we get

$$\|\eta^{L^2}\|_{L^2} = (\eta^{L^2}, u - v_r)_{L^2} + \cancelto{0}{\left(\eta^{L^2}, v_r - \sum_{j=1}^{r} (u, \varphi_j)_{H^1} \varphi_j\right)_{L^2}} \tag{3.34}$$
$$\leq \|\eta^{L^2}\|_{L^2} \|\eta^{proj}\|_{L^2},$$

which, after simplifying by $\|\eta^{proj}\|_{L^2}$, proves (3.32). Inequality (3.33) follows from (3.32) and Lemma 3.2. ∎

*Remark* 3.5. The approach used in Theorem 2 in [33] (see also Lemma 3.2 in [20]) yields error estimates that are sharper than those in (3.31) and (3.33). We did not use this approach, however, for the following two reasons: First, estimates (3.31) and (3.33) are easier to use with the definition of optimality of POD error estimates employed in this paper (see Definition 3.1). Second, although sharper than (3.31) and (3.33), the estimates yielded by Theorem 2 in [33] would not change the conclusions regarding the optimality of the error estimates presented in the remainder of this section (see Remark 3.2).

Next, we show how the error analysis in Case I changes with $w_r = P_r(u)$ as in [10] (see also [33]). Using (3.29), (3.13) becomes

$$(\phi_{r,t}, v_r)_{L^2} + \nu (\nabla \phi_r, \nabla v_r)_{L^2} = \cancelto{0}{(\eta_t^{L^2}, v_r)_{L^2}} + \nu (\nabla \eta^{L^2}, \nabla v_r)_{L^2}, \tag{3.35}$$

where $\eta_t^{L^2} = u_t - P_r(u_t)$. We emphasize that it is the cancellation of the first term on the RHS of (3.35) that yields error estimates that do not require the DQs. We let $v_r := \phi_r$ in (3.35) and we apply the Cauchy–Schwarz inequality:

$$\frac{1}{2} \frac{d}{dt} \|\phi_r\|_{L^2}^2 + \nu \|\nabla \phi_r\|_{L^2}^2 \leq \nu \|\nabla \eta^{L^2}\|_{L^2} \|\nabla \phi_r\|_{L^2}. \tag{3.36}$$

The error analysis can then proceed in several directions.

**3.2.2. Approach II.A.** One approach is to use Young's inequality in (3.36):

$$\frac{1}{2} \frac{d}{dt} \|\phi_r\|_{L^2}^2 + \nu \|\nabla \phi_r\|_{L^2}^2 \leq \frac{\nu}{2} \|\nabla \eta^{L^2}\|_{L^2}^2 + \frac{\nu}{2} \|\nabla \phi_r\|_{L^2}^2, \tag{3.37}$$

which implies

$$\frac{1}{2} \frac{d}{dt} \|\phi_r\|_{L^2}^2 + \frac{\nu}{2} \|\nabla \phi_r\|_{L^2}^2 \leq \frac{\nu}{2} \|\nabla \eta^{L^2}\|_{L^2}^2. \tag{3.38}$$

Noticing that the second term on the LHS of (3.38) is positive, we get

$$\frac{d}{dt} \|\phi_r\|_{L^2}^2 \leq \nu \|\nabla \eta^{L^2}\|_{L^2}^2. \tag{3.39}$$

When the $L^2$-POD basis is used, according to Definition 3.1 and Remark 3.2, inequalities (3.31) and (3.39) imply that Approach II.A will yield error estimates that are suboptimal with respect to $r$. When the $H^1$-POD basis is used, according to Definition 3.1 and Remark 3.2, inequalities (3.33) and (3.39) imply that Approach II.A will yield error estimates that are optimal with respect to $r$.

**3.2.3. Approach II.B.** Another way of continuing from (3.36) is to apply the POD inverse estimates in Lemma 3.2:

$$\|\nabla \phi_r\|_{L^2} \leq C_{inv}(r)\, \|\phi_r\|_{L^2}, \tag{3.40}$$

where $C_{inv}(r) = C_{inv}^{L^2}(r)$ if the $L^2$-POD basis is used and $C_{inv}(r) = C_{inv}^{H^1}(r)$ if the $H^1$-POD basis is used. Using (3.40) in (3.36) yields

$$\frac{1}{2}\,\frac{d}{dt}\|\phi_r\|_{L^2}^2 + \nu\,\|\nabla \phi_r\|_{L^2}^2 \leq C_{inv}(r)\,\nu\,\|\nabla \eta^{L^2}\|_{L^2}\,\|\phi_r\|_{L^2}. \tag{3.41}$$

Dropping $\nu\,\|\nabla \phi_r\|_{L^2}^2$ in (3.41) and simplifying, we get

$$\frac{d}{dt}\|\phi_r\|_{L^2} \leq C_{inv}(r)\,\nu\,\|\nabla \eta^{L^2}\|_{L^2}. \tag{3.42}$$

Comparing estimate (3.42) with estimate (3.39) in Approach II.A, we note that both estimates have $\|\nabla \eta^{L^2}\|_{L^2}$ on the RHS. In addition, estimate (3.42) has $C_{inv}(r)$ on the RHS, which increases the suboptimality with respect to $r$ (see Remark 3.3). Thus, estimate (3.42) suggests that Approach II.B yields estimates that are suboptimal with respect to $r$ when either the $L^2$-POD basis or the $H^1$-POD basis is used.

**3.2.4. Approach II.C.** Since both Approach II.A and Approach II.B yield error estimates that are suboptimal with respect to $r$ in the $L^2$-norm, one can try instead to prove optimal error estimates in the $H^1$-seminorm. To this end, we use the approach in [34] and, instead of choosing $v_r := \phi_r$ in (3.35), we choose $v_r := \phi_{r,t}$:

$$\|\phi_{r,t}\|_{L^2}^2 + \frac{\nu}{2}\,\frac{d}{dt}\|\nabla \phi_r\|_{L^2}^2 \leq \nu\,\|\nabla \eta^{L^2}\|_{L^2}\,\|\nabla \phi_{r,t}\|_{L^2}. \tag{3.43}$$

Applying Young's inequality and the POD inverse estimates in Lemma 3.2 in (3.43), we get

$$\begin{aligned}
\|\phi_{r,t}\|_{L^2}^2 + \frac{\nu}{2}\,\frac{d}{dt}\|\nabla \phi_r\|_{L^2}^2 &\leq \nu\,\|\nabla \eta^{L^2}\|_{L^2}\,\|\nabla \phi_{r,t}\|_{L^2} \\
&\leq \frac{\nu^2}{2}\,C_{inv}(r)^2\,\|\nabla \eta^{L^2}\|_{L^2}^2 + \frac{1}{2\,C_{inv}(r)^2}\,\|\nabla \phi_{r,t}\|_{L^2}^2 \\
&\leq \frac{\nu^2}{2}\,C_{inv}(r)^2\,\|\nabla \eta^{L^2}\|_{L^2}^2 + \frac{1}{2}\,\|\phi_{r,t}\|_{L^2}^2\,, \qquad (3.44)
\end{aligned}$$

where $C_{inv}(r) = C_{inv}^{L^2}(r)$ if the $L^2$-POD basis is used and $C_{inv}(r) = C_{inv}^{H^1}(r)$ if the $H^1$-POD basis is used. Inequality (3.44) implies

$$\frac{d}{dt}\|\nabla \phi_r\|_{L^2}^2 \leq \nu\,C_{inv}(r)^2\,\|\nabla \eta^{L^2}\|_{L^2}^2\,. \tag{3.45}$$

When the $L^2$-POD basis is used, inequalities (3.31) and (3.45), and Remark 3.3 imply that Approach II.C will yield error estimates that are suboptimal with respect to $r$ (see Definition 3.1 and Remark 3.2). When the $H^1$-POD basis is used, inequalities (3.33) and (3.45), and Remark 3.3 imply that Approach II.C will yield error estimates that are suboptimal with respect to $r$.

Since for Case I ($V = V^{DQ}$) in section 3.1 we did not prove error estimates in the $H^1$-norm, for a fair comparison with Approach II.C, we prove these error estimates below. To this end, we let $v_r := \phi_{r,t}$ in (3.18):

$$(\phi_{r,t}, \phi_{r,t})_{L^2} + \nu\,(\nabla \phi_r, \nabla \phi_{r,t})_{L^2} = (\eta_t^{Ritz}, \phi_{r,t})_{L^2}\,. \tag{3.46}$$

TABLE 2
*Theoretical convergence rates for the no_DQ and the DQ cases: $L^2$-POD basis (second and fourth column); $H^1$-POD basis (third and fifth column).*

| | no_DQ $L^2$-POD | no_DQ $H^1$-POD | DQ $L^2$-POD | DQ $H^1$-POD |
|---|---|---|---|---|
| $\mathcal{E}_{C^0(L^2)}$ | suboptimal section 3.2.2 section 3.2.3 | optimal section 3.2.2 | optimal section 3.1 | optimal section 3.1 |
| $\mathcal{E}_{C^0(H^1)}$ | suboptimal section 3.2.4 | suboptimal section 3.2.4 | optimal section 3.2.4 | optimal section 3.2.4 |
| $\mathcal{E}_{L^2(H^1)}$ | optimal section 3.2.5 | optimal section 3.2.5 | optimal | optimal |

Applying Young's inequality on the RHS of (3.46), we get

$$\|\phi_{r,t}\|_{L^2}^2 + \frac{\nu}{2}\frac{d}{dt}\|\nabla\phi_r\|_{L^2}^2 \le \|\eta_t^{Ritz}\|_{L^2}\|\phi_{r,t}\|_{L^2} \le \frac{1}{2}\|\eta_t^{Ritz}\|_{L^2}^2 + \frac{1}{2}\|\phi_{r,t}\|_{L^2}^2, \tag{3.47}$$

which implies

$$\frac{d}{dt}\|\nabla\phi_r\|_{L^2}^2 \le \frac{1}{\nu}\|\eta_t^{Ritz}\|_{L^2}^2. \tag{3.48}$$

Using the triangle inequality $\|\nabla e\|_{L^2} \le \|\nabla\eta^{Ritz}\|_{L^2} + \|\nabla\phi_r\|_{L^2}$, inequality (3.48), and assumptions (3.26)–(3.27), we conclude that, when either the $L^2$-POD basis or the $H^1$-POD basis is used, Case I ($V = V^{DQ}$) in section 3.1 yields error estimates that are optimal with respect to $r$ (see Definition 3.1 and Remark 3.2).

**3.2.5. Approach II.D.** In this section, we derive error estimates in the solution norm (i.e., in the $L^2(0,T;H^1(\Omega))$-norm). Integrating (3.38) from 0 to $T$, we get

$$\|\phi_r(T)\|_{L^2}^2 + \nu\int_0^T \|\nabla\phi_r(s)\|_{L^2}^2\,ds \le \|\phi_r(0)\|_{L^2}^2 + \nu\int_0^T \|\nabla\eta^{L^2}(s)\|_{L^2}^2\,ds. \tag{3.49}$$

Inequality (3.31) (for the $L^2$-POD basis) or inequality (3.33) (for the $H^1$-POD basis) implies that Approach II.D will yield error estimates that are optimal with respect to $r$ (see Definition 3.1 and Remark 3.2). We note that Proposition 3.3 in [10] yields a similar estimate.

The theoretical convergence rates derived in this section are summarized in Table 2.

**4. Numerical results.** The main goal of this section is to numerically investigate the rates of convergence with respect to $r$ of the POD-G-ROM (2.12) in the two cases considered in section 3: Case I ($V = V^{DQ}$) and Case II ($V = V^{no\_DQ}$). We also consider both the $L^2$-POD basis and the $H^1$-POD basis. Although the error analysis in section 3 has been centered around the (linear) heat equation, in this section we consider both the heat equation (section 4.1) and the nonlinear Burgers equation (section 4.2).

Denoting the error at time $t_j$ by $e_j := u_h^r(\cdot,t_j) - u(\cdot,t_j)$, the following error norms are considered: the error in the $C^0(0,T;L^2(\Omega))$-norm, approximated by

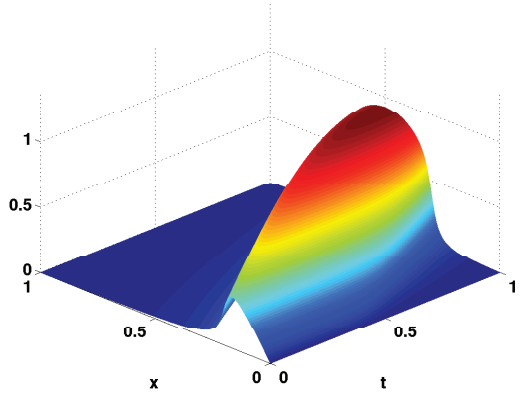


FIG. 7. *Heat equation. Fine resolution finite element solution used to generate the snapshots.*

$\mathcal{E}_{C^0(L^2)} = \max_{0\le j\le N} \|e_j\|_{L^2(\Omega)}$; the error in the $C^0(0,T;H^1(\Omega))$-norm, approximated by $\mathcal{E}_{C^0(H^1)} = \max_{0\le j\le N} \|e_j\|_{H^1(\Omega)}$; and the error in the $L^2(0,T;H^1(\Omega))$-norm, approximated by $\mathcal{E}_{L^2(H^1)} = \sqrt{\frac{1}{N+1}\sum_{0\le j\le N} \|e_j\|^2_{H^1(\Omega)}}$. We also use the following notation:

$$\Lambda_r = \sqrt{\sum_{j=r+1}^{d} \lambda_j} \quad \text{and} \quad \Theta_r = \sqrt{\frac{1}{N+1}\sum_{i=0}^{N} \|u_t(\cdot,t_i) - \sum_{j=1}^{r} (u_t(\cdot,t_i),\varphi_j(\cdot))_{L^2}\varphi_j\|^2_{L^2}}.$$

Definition 3.1, Remark 3.2, and the POD projection error estimates (2.7)–(2.10) yield the following scalings for optimal error estimates: (i) $\mathcal{E}_{C^0(L^2)} = \mathcal{O}(\sqrt{\sum_{j=r+1}^{d} \lambda_j})$ for both the $L^2$-POD basis and the $H^1$-POD basis; and (ii) $\mathcal{E}_{C^0(H^1)} \sim \mathcal{E}_{L^2(H^1)} = \mathcal{O}(\sqrt{\|S_d\|_2 \sum_{j=r+1}^{d} \lambda_j})$ for the $L^2$-POD basis and $\mathcal{E}_{C^0(H^1)} \sim \mathcal{E}_{L^2(H^1)} = \mathcal{O}(\sqrt{\sum_{j=r+1}^{d} \lambda_j})$ for the $H^1$-POD basis.

**4.1. Heat equation.** We consider the one-dimensional heat equation (2.1) with a known exact solution that represents the propagation in time of a steep front:

$$u(x,t) = \sin(\pi x)\left[\frac{1}{\pi}\arctan\left(\frac{c}{25} - c\left(x - \frac{t}{2}\right)^2\right) + \frac{1}{2}\right], \tag{4.1}$$

where $x \in [0,1]$ and $t \in [0,1]$. The constant $c$ in (4.1) controls the steepness of the front. In all the numerical tests in this section, we use the value $c = 100$. The value of the diffusion coefficient used in the heat equation (2.1) is $\nu = 10^{-2}$. Piecewise linear finite elements are used to generate snapshots for the POD-G-ROM (2.12). A mesh size $h = 1/1024$ and the Crank–Nicolson scheme with a time step $\Delta t = 10^{-3}$ are employed for the spatial and temporal discretizations, respectively. The time evolution of the finite element solution is shown in Figure 7. In total, 1001 snapshots are collected and used for generating POD basis functions. The same numerical solver as that used in the finite element approximation is utilized in the POD-G-ROM.

**4.1.1. $L^2$-POD basis.** In this section, we check the rates of convergence with respect to $r$ for the no_DQ and the DQ cases when the $L^2$-POD basis is used. The dimensions of the spaces $V^{no\_DQ}$ and $V^{DQ}$ are $d = 69$ and $d = 100$, respectively.

TABLE 3
*Heat equation, $L^2$-POD basis. Errors in the no_DQ case.*

| $r$ | $\Lambda_r$ | $\Theta_r$ | $\mathcal{E}_{C^0(L^2)}$ | $\mathcal{E}_{C^0(H^1)}$ | $\mathcal{E}_{L^2(H^1)}$ |
|---|---|---|---|---|---|
| 3 | 5.72e-02 | 8.37e-01 | 9.46e-02 | 2.30e+00 | 1.59e+00 |
| 5 | 2.71e-02 | 6.04e-01 | 4.70e-02 | 1.58e+00 | 1.14e+00 |
| 7 | 1.58e-02 | 4.58e-01 | 3.69e-02 | 1.38e+00 | 8.22e-01 |
| 10 | 7.34e-03 | 2.92e-01 | 1.57e-02 | 8.54e-01 | 5.31e-01 |
| 13 | 3.84e-03 | 1.90e-01 | 7.78e-03 | 5.84e-01 | 3.50e-01 |
| 17 | 1.71e-03 | 1.11e-01 | 5.15e-03 | 4.34e-01 | 1.88e-01 |
| 20 | 9.07e-04 | 6.91e-02 | 2.70e-03 | 2.56e-01 | 1.19e-01 |
| 24 | 4.36e-04 | 3.89e-02 | 1.24e-03 | 1.51e-01 | 6.98e-02 |
| 28 | 2.01e-04 | 2.11e-02 | 5.87e-04 | 8.27e-02 | 4.13e-02 |
| 31 | 1.19e-04 | 1.34e-02 | 3.02e-04 | 5.36e-02 | 3.12e-02 |
| 35 | 5.98e-05 | 7.94e-03 | 2.05e-04 | 4.30e-02 | 2.41e-02 |
| 39 | 2.93e-05 | 4.14e-03 | 7.94e-05 | 2.89e-02 | 2.17e-02 |
| 43 | 1.47e-05 | 2.46e-03 | 5.69e-05 | 2.65e-02 | 2.08e-02 |
| 47 | 7.48e-06 | 1.32e-03 | 2.70e-05 | 2.46e-02 | 2.06e-02 |
| 51 | 3.71e-06 | 7.35e-04 | 1.81e-05 | 2.42e-02 | 2.05e-02 |
| 56 | 1.61e-06 | 3.59e-04 | 1.35e-05 | 2.40e-02 | 2.05e-02 |
| 60 | 8.26e-07 | 1.92e-04 | 1.22e-05 | 2.39e-02 | 2.05e-02 |

TABLE 4
*Heat equation, $L^2$-POD basis. Errors in the DQ case.*

| $r$ | $\Lambda_r$ | $\Theta_r$ | $\mathcal{E}_{C^0(L^2)}$ | $\mathcal{E}_{C^0(H^1)}$ | $\mathcal{E}_{L^2(H^1)}$ |
|---|---|---|---|---|---|
| 19 | 5.49e-02 | 7.75e-02 | 7.15e-03 | 4.96e-01 | 3.19e-01 |
| 23 | 2.95e-02 | 4.18e-02 | 2.03e-03 | 1.98e-01 | 1.19e-01 |
| 28 | 1.41e-02 | 2.00e-02 | 6.52e-04 | 7.97e-02 | 4.91e-02 |
| 33 | 6.75e-03 | 9.55e-02 | 2.41e-04 | 3.68e-02 | 2.80e-02 |
| 37 | 3.76e-03 | 5.31e-03 | 8.60e-05 | 2.67e-02 | 2.29e-02 |
| 42 | 1.77e-03 | 2.51e-03 | 2.66e-05 | 2.44e-02 | 2.10e-02 |
| 47 | 8.32e-04 | 1.18e-03 | 1.50e-05 | 2.40e-02 | 2.06e-02 |
| 51 | 4.62e-04 | 6.54e-04 | 1.38e-05 | 2.40e-02 | 2.05e-02 |
| 56 | 2.20e-04 | 3.11e-04 | 1.21e-05 | 2.39e-02 | 2.05e-02 |
| 60 | 1.19e-04 | 1.69e-04 | 1.18e-05 | 2.39e-02 | 2.05e-02 |
| 65 | 5.59e-05 | 7.91e-05 | 1.17e-05 | 2.39e-02 | 2.05e-02 |
| 70 | 2.63e-05 | 3.71e-05 | 1.17e-05 | 2.39e-02 | 2.05e-02 |
| 74 | 1.48e-05 | 2.10e-05 | 1.17e-05 | 2.39e-02 | 2.05e-02 |
| 79 | 6.99e-06 | 9.90e-06 | 1.17e-05 | 2.39e-02 | 2.05e-02 |
| 84 | 3.29e-06 | 4.67e-06 | 1.17e-05 | 2.39e-02 | 2.05e-02 |
| 88 | 1.74e-06 | 2.50e-06 | 1.17e-05 | 2.39e-02 | 2.05e-02 |
| 92 | 9.52e-07 | 1.41e-06 | 1.17e-05 | 2.39e-02 | 2.05e-02 |

The errors are listed in Table 3 (in the no_DQ case) and in Table 4 (in the DQ case). These errors with their linear regression plots are drawn in Figure 8. The convergence rate of the error $\mathcal{E}_{C^0(L^2)}$ is superoptimal in the DQ case and suboptimal in the no_DQ case. This supports the theoretical rates of convergence in Table 2, although the suboptimality in the no_DQ case is mild. The convergence rate of the error $\mathcal{E}_{C^0(H^1)}$ is optimal in the DQ case and strongly suboptimal in the no_DQ case. This again supports the theoretical rates of convergence in Table 2. The convergence rate of the error $\mathcal{E}_{L^2(H^1)}$ is optimal in the DQ case and strongly suboptimal in the no_DQ case. This supports the theoretical rates of convergence in Table 2 for the DQ case, but not for the no_DQ case.

We note that all the errors in Figure 8 reach a plateau. This is due to the fact that, for large $r$ values, the spatial (finite element) discretization error and the time discretization error are the dominating components of the total error.

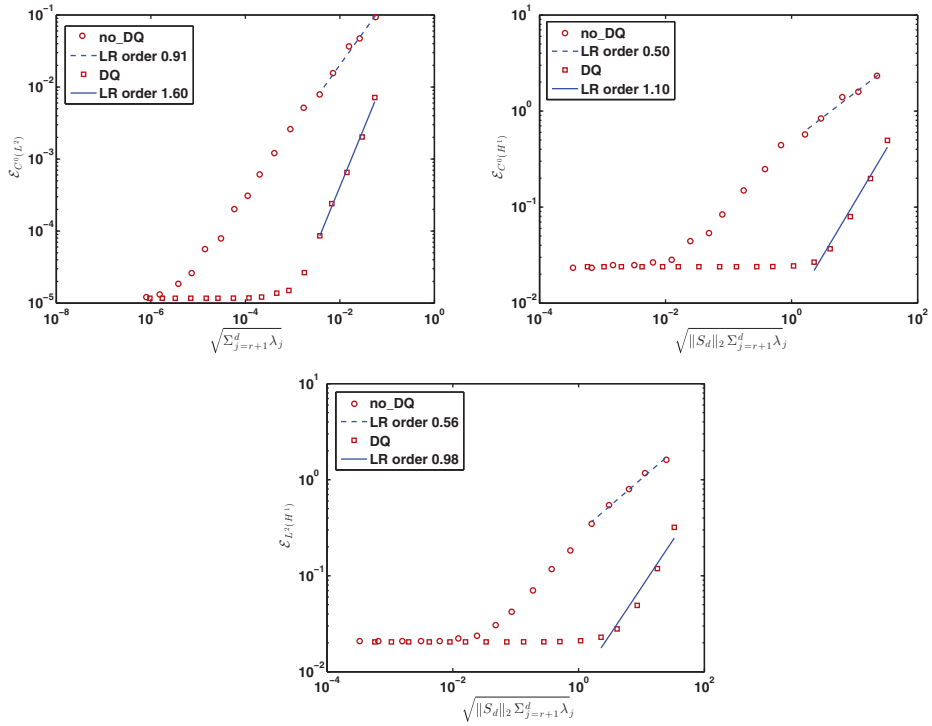


FIG. 8. *Heat equation, $L^2$-POD basis. Plots of errors in $C^0(L^2)$-norm (top, left), $C^0(H^1)$-norm (top, right), and $L^2(H^1)$-norm (bottom).*

We also note that, although the convergence rates for the DQ case in Table 4 are consistently higher than the convergence rates for the no_DQ case in Table 3, to reach a prescribed error tolerance, higher $r$ values are needed in the DQ case than in the no_DQ case (see the $\Lambda_r$ values in the two tables). We emphasize, however, that for other problems, such as those in [11], similar $r$ values are needed in the DQ and no_DQ cases to reach a prescribed tolerance (results not included). We also note that, for a prescribed time derivative error tolerance, similar $r$ values are needed in the DQ and no_DQ cases (see the $\Theta_r$ values in Tables 3 and 4). Finally, we mention that, as $r \to d$, both $\sqrt{\sum_{i=r+1}^{d} \lambda_i^{no\_DQ}}$ and $\sqrt{\sum_{i=r+1}^{d} \lambda_i^{DQ}}$ approach zero exponentially fast.

Overall, the numerical results support the theoretical rates of convergence proved in section 3. We also emphasize that *the convergence rates in the DQ case in all three norms are much higher than (and almost twice as high as) the corresponding convergence rates in the no_DQ case.*

**4.1.2. $H^1$-POD basis.** In this section, we repeat the numerical tests in section 4.1.1, but with the $H^1$-POD basis instead of the $L^2$-POD basis. The dimensions of the spaces $V^{no\_DQ}$ and $V^{DQ}$ are $d = 97$ and $d = 123$, respectively. The errors for different values of $r$ are listed in Table 5 (in the no_DQ case) and in Table 6 (in the DQ case). These errors with their linear regression plots are drawn in Figure 9. The convergence rate of the error $\mathcal{E}_{C^0(L^2)}$ is superoptimal in the DQ case and optimal in the no_DQ case. This supports the theoretical rates of convergence in Table 2. The convergence rate of the error $\mathcal{E}_{C^0(H^1)}$ is superoptimal in the DQ case and suboptimal in the no_DQ case. This supports the theoretical rates of convergence in Table 2. The

TABLE 5
*Heat equation, $H^1$-POD basis. Errors in the no_DQ case.*

| $r$ | $\Lambda_r$ | $\mathcal{E}_{C^0(L^2)}$ | $\mathcal{E}_{C^0(H^1)}$ | $\mathcal{E}_{L^2(H^1)}$ |
|---|---|---|---|---|
| 3 | 1.51e+00 | 1.11e-01 | 2.37e+00 | 1.56e+00 |
| 7 | 8.00e-01 | 4.04e-02 | 1.41e+00 | 8.16e-01 |
| 12 | 3.77e-01 | 1.54e-02 | 8.32e-01 | 3.85e-01 |
| 17 | 1.78e-01 | 6.64e-03 | 4.86e-01 | 1.84e-01 |
| 21 | 1.01e-01 | 2.97e-03 | 2.86e-01 | 1.05e-01 |
| 26 | 4.77e-02 | 1.33e-03 | 1.56e-01 | 5.31e-02 |
| 31 | 2.27e-02 | 5.72e-04 | 8.48e-02 | 3.11e-02 |
| 35 | 1.22e-02 | 3.01e-04 | 5.23e-02 | 2.40e-02 |
| 40 | 5.81e-03 | 1.39e-04 | 3.44e-02 | 2.13e-02 |
| 45 | 2.78e-03 | 6.30e-05 | 2.71e-02 | 2.07e-02 |
| 49 | 1.54e-03 | 2.71e-05 | 2.47e-02 | 2.05e-02 |
| 54 | 7.28e-04 | 1.50e-05 | 2.41e-02 | 2.05e-02 |

TABLE 6
*Heat equation, $H^1$-POD basis. Errors in the DQ case.*

| $r$ | $\Lambda_r$ | $\mathcal{E}_{C^0(L^2)}$ | $\mathcal{E}_{C^0(H^1)}$ | $\mathcal{E}_{L^2(H^1)}$ |
|---|---|---|---|---|
| 32 | 1.50e+00 | 7.85e-02 | 3.02e+00 | 1.18e+00 |
| 38 | 7.27e-01 | 3.97e-04 | 4.75e-02 | 3.64e-02 |
| 43 | 3.87e-01 | 7.65e-05 | 2.49e-02 | 2.13e-02 |
| 49 | 1.80e-01 | 1.54e-05 | 2.40e-02 | 2.06e-02 |
| 54 | 9.29e-02 | 1.29e-05 | 2.39e-02 | 2.05e-02 |
| 59 | 4.76e-02 | 1.20e-05 | 2.39e-02 | 2.05e-02 |
| 64 | 2.43e-02 | 1.18e-05 | 2.39e-02 | 2.05e-02 |
| 69 | 1.24e-02 | 1.17e-05 | 2.39e-02 | 2.05e-02 |
| 75 | 5.55e-03 | 1.17e-05 | 2.39e-02 | 2.05e-02 |
| 80 | 2.77e-03 | 1.17e-05 | 2.39e-02 | 2.05e-02 |
| 85 | 1.38e-03 | 1.17e-05 | 2.39e-02 | 2.05e-02 |
| 90 | 6.90e-04 | 1.17e-05 | 2.39e-02 | 2.05e-02 |

convergence rate of the error $\mathcal{E}_{L^2(H^1)}$ is superoptimal in the DQ case and optimal in the no_DQ case. This supports the theoretical rates of convergence in Table 2.

Overall, the numerical results support the theoretical rates of convergence proved in section 3. As suggested by the error analysis in section 3, the POD-ROM-G for the no_DQ case is more accurate for the $H^1$-POD basis used in this section than for the $L^2$-POD basis used in section 4.1.1. We emphasize, however, that, as in section 4.1.1, *the convergence rates in the DQ case in all three norms are much higher than (and sometimes almost twice as high as) the corresponding convergence rates in the no_DQ case.*

**4.2. Burgers equation.** In this section, we consider the one-dimensional Burgers equation. As mentioned at the beginning of section 4, the error estimates proved in section 3 are valid for the (linear) heat equation, but not necessarily valid for the nonlinear Burgers equation. Nevertheless, to gain some insight into the range of validity of the theoretical developments in section 3, we investigate the convergence rates with respect to $r$ in the no_DQ and the DQ cases for the nonlinear Burgers equation:

$$\text{(4.2)} \qquad \begin{cases} u_t - \nu\, u_{xx} + u\, u_x = f & \text{in } \Omega \times (0,T]\,, \\ u(x,0) = u_0(x) & \text{in } \Omega\,, \\ u(x,t) = g(x,t) & \text{on } \partial\Omega \times (0,T]\,. \end{cases}$$

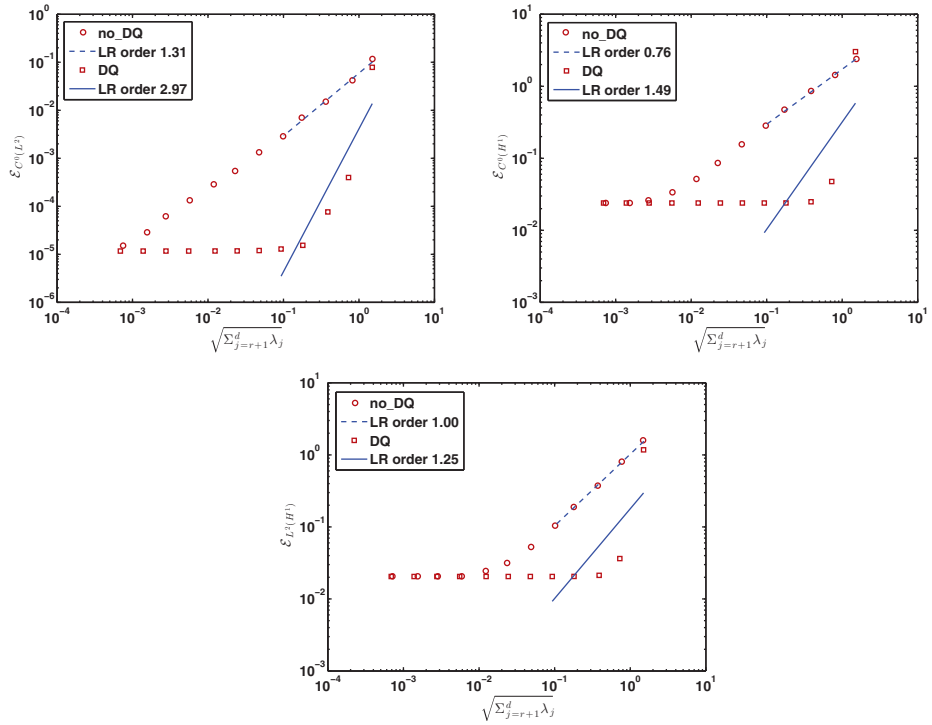


FIG. 9. *Heat equation, $H^1$-POD basis. Plots of errors in $C^0(L^2)$-norm (top, left), $C^0(H^1)$-norm (top, right), and $L^2(H^1)$-norm (bottom).*

The initial condition is

$$u_0(x) = \begin{cases} 1 & \text{if } x \in \left(0, \frac{1}{2}\right] , \\ 0 & \text{if } x \in \left(\frac{1}{2}, 1\right) , \end{cases} \tag{4.3}$$

which is similar to that used in [23]. We note that, although used in other numerical investigations, the initial condition (4.3) does not fit into the classical theoretical framework of section 3 (since it is discontinuous). The diffusion parameter is $\nu = 10^{-2}$, the forcing term is $f = 0$, $\Omega = [0, 1]$, and $T = 1$. The boundary conditions are homogeneous Dirichlet, that is, $u(0, t) = u(1, t) = 0$ for all $t \in [0, 1]$.

To generate snapshots, we use piecewise linear finite elements with mesh size $h = 1/1024$ and the backward Euler method with a time step $\Delta t = 10^{-4}$, and save data at each time instance. The time evolution of the finite element solution is shown in Figure 10. All snapshots are used for the POD basis generation and the same numerical solver is used in the POD-G-ROM.

**4.2.1. $L^2$-POD basis.** In this section, we check the rates of convergence with respect to $r$ for the no_DQ and the DQ cases. The dimensions of the spaces $V^{no\_DQ}$ and $V^{DQ}$ are $d = 49$ and $d = 65$, respectively. Since the exact solution of the Burgers equation (4.2) with the initial condition (4.3) is not known, we consider the errors between the POD-G-ROM results and the snapshots. The errors are listed in Table 7 (in the no_DQ case) and in Table 8 (in the DQ case). These errors with their linear regression plots are drawn in Figure 11. The convergence rate of the error $\mathcal{E}_{C^0(L^2)}$

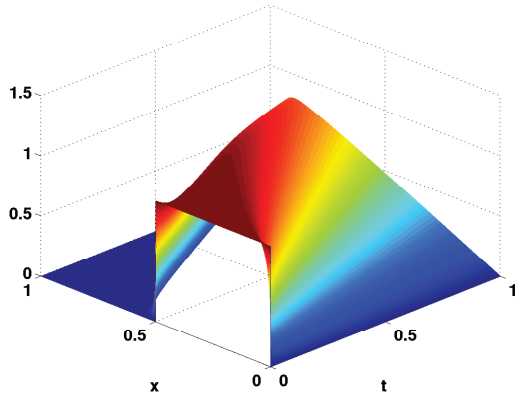


FIG. 10. *Burgers equation. Fine resolution finite element solution used to generate the snapshots.*

TABLE 7
*Burgers equation, $L^2$-POD basis. Errors in the no_DQ case.*

| $r$ | $\Lambda_r$ | $\Theta_r$ | $\mathcal{E}_{C^0(L^2)}$ | $\mathcal{E}_{C^0(H^1)}$ | $\mathcal{E}_{L^2(H^1)}$ |
|---|---|---|---|---|---|
| 3 | 8.74e-02 | 2.77e+00 | 2.38e-01 | 4.48e+01 | 2.34e+00 |
| 5 | 3.95e-02 | 2.62e+00 | 1.60e-01 | 4.44e+01 | 1.76e+00 |
| 7 | 1.97e-02 | 2.52e+00 | 1.17e-01 | 4.37e+01 | 1.24e+00 |
| 9 | 1.02e-02 | 2.44e+00 | 9.05e-02 | 4.28e+01 | 8.94e-01 |
| 11 | 5.47e-03 | 2.37e+00 | 7.01e-02 | 4.14e+01 | 6.84e-01 |
| 13 | 3.06e-03 | 2.26e+00 | 5.37e-02 | 3.91e+01 | 5.48e-01 |
| 16 | 1.36e-03 | 2.03e+00 | 3.44e-02 | 3.32e+01 | 4.06e-01 |
| 19 | 6.39e-04 | 1.63e+00 | 1.91e-02 | 2.32e+01 | 2.76e-01 |
| 22 | 3.05e-04 | 1.08e+00 | 8.13e-03 | 1.14e+01 | 1.49e-01 |
| 24 | 1.86e-04 | 7.31e-01 | 4.32e-03 | 5.86e+00 | 8.81e-02 |
| 27 | 8.83e-05 | 3.73e-01 | 2.54e-03 | 2.40e+00 | 3.89e-02 |
| 30 | 4.18e-05 | 1.82e-01 | 1.21e-03 | 1.06e+00 | 1.77e-02 |
| 33 | 1.96e-05 | 8.64e-02 | 4.96e-04 | 4.17e-01 | 8.22e-03 |
| 35 | 1.18e-05 | 5.22e-02 | 2.58e-04 | 2.64e-01 | 4.94e-03 |
| 38 | 5.48e-06 | 2.43e-02 | 1.20e-04 | 1.35e-01 | 2.31e-03 |

TABLE 8
*Burgers equation, $L^2$-POD basis. Errors in the DQ case.*

| $r$ | $\Lambda_r$ | $\Theta_r$ | $\mathcal{E}_{C^0(L^2)}$ | $\mathcal{E}_{C^0(H^1)}$ | $\mathcal{E}_{L^2(H^1)}$ |
|---|---|---|---|---|---|
| 18 | 8.55e-02 | 1.21e-01 | 6.83e-03 | 5.60e-01 | 2.82e-01 |
| 21 | 4.56e-02 | 6.45e-02 | 2.99e-03 | 2.49e-01 | 1.37e-01 |
| 24 | 2.39e-02 | 3.38e-02 | 1.31e-03 | 1.23e-01 | 6.72e-02 |
| 28 | 9.81e-03 | 1.39e-02 | 4.29e-04 | 4.73e-02 | 2.57e-02 |
| 31 | 4.97e-03 | 7.03e-03 | 1.88e-04 | 2.28e-02 | 1.25e-02 |
| 34 | 2.49e-03 | 3.52e-03 | 8.27e-05 | 1.10e-02 | 6.10e-03 |
| 37 | 1.23e-03 | 1.73e-03 | 3.63e-05 | 5.24e-03 | 2.96e-03 |
| 39 | 7.62e-04 | 1.08e-03 | 2.10e-05 | 3.19e-03 | 1.82e-03 |
| 42 | 3.69e-04 | 5.22e-04 | 9.21e-06 | 1.51e-03 | 8.72e-04 |
| 45 | 1.77e-04 | 2.50e-04 | 4.03e-06 | 7.07e-04 | 4.15e-04 |
| 48 | 8.35e-05 | 1.18e-04 | 1.76e-06 | 3.29e-04 | 1.96e-04 |
| 51 | 3.90e-05 | 5.52e-05 | 7.62e-07 | 1.52e-04 | 9.15e-05 |
| 53 | 2.33e-05 | 3.30e-05 | 4.35e-07 | 9.02e-05 | 5.47e-05 |
| 56 | 1.06e-05 | 1.51e-05 | 1.86e-07 | 4.09e-05 | 2.51e-05 |
| 59 | 4.74e-06 | 6.81e-06 | 7.94e-08 | 1.84e-05 | 1.14e-05 |

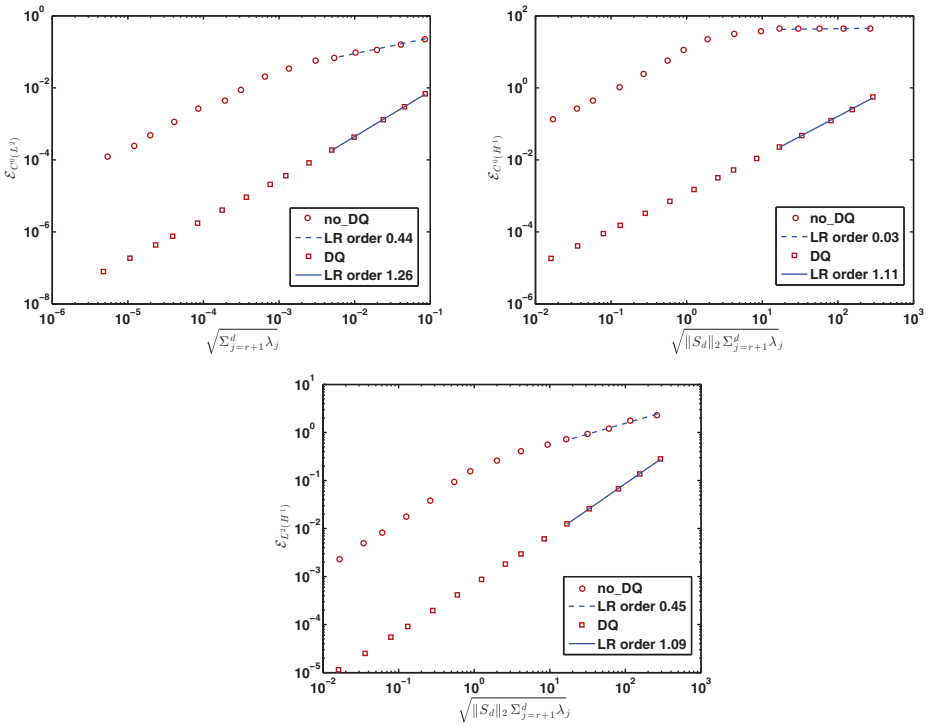


FIG. 11. *Burgers equation, $L^2$-POD basis. Plots of errors in $C^0(L^2)$-norm (top, left), $C^0(H^1)$-norm (top, right), and $L^2(H^1)$-norm (bottom).*

is superoptimal in the DQ case and strongly suboptimal in the no_DQ case. This clearly supports the theoretical rates of convergence in Table 2. The convergence rate of the error $\mathcal{E}_{C^0(H^1)}$ is optimal in the DQ case and extremely suboptimal in the no_DQ case. This strongly supports the theoretical rates of convergence in Table 2. Finally, the convergence rate of the error $\mathcal{E}_{L^2(H^1)}$ is optimal in the DQ case and strongly suboptimal in the no_DQ case. This supports the theoretical rates of convergence in Table 2 for the DQ case, but not for the no_DQ case.

Overall, the numerical results clearly support the theoretical rates of convergence proved in section 3. We also emphasize that *the convergence rates in the DQ case in all three norms are much higher than (and generally three times as high as) the corresponding convergence rates in the no_DQ case.*

**4.2.2. $H^1$-POD basis.** In this section, we repeat the numerical tests in section 4.2.1, but with the $H^1$-POD basis instead of the $L^2$-POD basis. The dimensions of the spaces $V^{no\_DQ}$ and $V^{DQ}$ are $d = 66$ and $d = 61$, respectively. The errors for different values of $r$ are listed in Table 9 (in the no_DQ case) and in Table 10 (in the DQ case). These errors with their linear regression plots are drawn in Figure 12. The convergence rate of the error $\mathcal{E}_{C^0(L^2)}$ is superoptimal in the DQ case and optimal in the no_DQ case. This supports the theoretical rates of convergence in Table 2. The convergence rate of the error $\mathcal{E}_{C^0(H^1)}$ is superoptimal in the DQ case and suboptimal in the no_DQ case. This again supports the theoretical rates of convergence in Table 2. The convergence rate of the error $\mathcal{E}_{L^2(H^1)}$ is superoptimal in the DQ case and optimal in the no_DQ case. This supports the theoretical rates of convergence in Table 2.

TABLE 9
*Burgers equation, $H^1$-POD basis. Errors in the no_DQ case.*

| $r$ | $\Lambda_r$ | $\mathcal{E}_{C^0(L^2)}$ | $\mathcal{E}_{C^0(H^1)}$ | $\mathcal{E}_{L^2(H^1)}$ |
|---|---|---|---|---|
| 3 | 2.22e+00 | 2.07e-01 | 4.45e+01 | 2.43e+00 |
| 7 | 1.13e+00 | 9.31e-02 | 4.23e+01 | 1.25e+00 |
| 11 | 5.26e-01 | 2.94e-02 | 2.65e+01 | 6.35e-01 |
| 14 | 2.77e-01 | 1.23e-02 | 1.11e+01 | 3.35e-01 |
| 17 | 1.44e-01 | 6.30e-03 | 5.65e+00 | 1.71e-01 |
| 20 | 7.44e-02 | 3.22e-03 | 2.85e+00 | 8.72e-02 |
| 24 | 3.09e-02 | 1.14e-03 | 9.48e-01 | 3.56e-02 |
| 27 | 1.57e-02 | 4.68e-04 | 4.41e-01 | 1.79e-02 |
| 30 | 7.90e-03 | 2.19e-04 | 2.51e-01 | 8.93e-03 |
| 33 | 3.95e-03 | 1.21e-04 | 1.23e-01 | 4.43e-03 |
| 36 | 1.96e-03 | 5.78e-05 | 5.54e-02 | 2.18e-03 |
| 39 | 9.58e-04 | 2.51e-05 | 2.31e-02 | 1.06e-03 |

TABLE 10
*Burgers equation, $H^1$-POD basis. Errors in the DQ case.*

| $r$ | $\Lambda_r$ | $\mathcal{E}_{C^0(L^2)}$ | $\mathcal{E}_{C^0(H^1)}$ | $\mathcal{E}_{L^2(H^1)}$ |
|---|---|---|---|---|
| 26 | 2.24e+00 | 7.07e-01 | 7.50e+01 | 5.00e+00 |
| 30 | 1.04e+00 | 4.09e-04 | 4.56e-02 | 2.37e-02 |
| 33 | 5.65e-01 | 1.65e-04 | 2.04e-02 | 1.08e-02 |
| 36 | 3.01e-01 | 6.95e-05 | 9.39e-03 | 5.10e-03 |
| 40 | 1.26e-01 | 2.23e-05 | 3.37e-03 | 1.88e-03 |
| 43 | 6.46e-02 | 9.55e-06 | 1.56e-03 | 8.84e-04 |
| 46 | 3.26e-02 | 4.10e-06 | 7.17e-04 | 4.14e-04 |
| 49 | 1.62e-02 | 1.75e-06 | 3.28e-04 | 1.92e-04 |
| 52 | 7.90e-03 | 7.43e-07 | 1.48e-04 | 8.84e-05 |
| 54 | 4.83e-03 | 4.19e-07 | 8.72e-05 | 5.24e-05 |
| 57 | 2.19e-03 | 2.58e-07 | 5.33e-05 | 2.54e-05 |
| 59 | 1.14e-03 | 4.15e-07 | 9.13e-05 | 1.80e-05 |

Overall, the numerical results support the theoretical rates of convergence proved in section 3. As suggested by the error analysis in section 3, the convergence rates of the POD-G-ROM for the no_DQ case are higher for the $H^1$-POD basis used in this section than for the $L^2$-POD basis used in section 4.2.1. We emphasize, however, that, as in section 4.2.1 *the convergence rates in the DQ case in all three norms are much higher than (and almost three times as high as) the corresponding convergence rates in the no_DQ case.*

**5. Conclusions.** The effect of using or not using the snapshot DQs in the generation of the POD basis (the DQ and the no_DQ cases, respectively) was investigated theoretically and numerically for both the $L^2$-POD basis and the $H^1$-POD basis. The criterion used in this theoretical and numerical investigation was the optimality of the rate of convergence with respect to $r$ of the POD-G-ROM solution to the exact solution, where $r$ is the number of POD basis functions used in the POD-G-ROM.

The error estimates in section 3 yielded the following conclusions: In the DQ case, the error estimates were optimal in all norms (i.e., the $C^0(L^2)$-norm, the $C^0(H^1)$-norm, and the $L^2(H^1)$-norm) for both the $L^2$-POD basis and the $H^1$-POD basis. In the no_DQ case, however, the error estimates were suboptimal in the $C^0(L^2)$-norm for the $L^2$-POD basis and in the $C^0(H^1)$-norm for both the $L^2$-POD basis and the $H^1$-POD basis.

The numerical results in section 4 for the (linear) heat equation and the (nonlinear) Burgers equation confirmed the conclusions suggested by the theoretical error

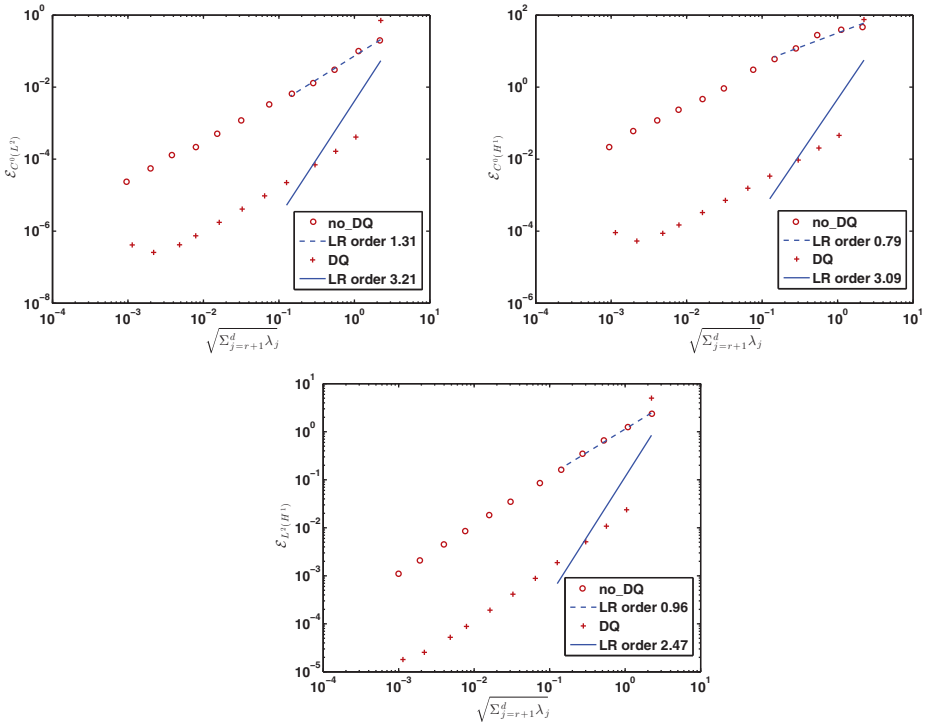


FIG. 12. *Burgers equation, $H^1$-POD basis. Plots of errors in $C^0(L^2)$-norm (top, left), $C^0(H^1)$-norm (top, right), and $L^2(H^1)$-norm (bottom).*

estimates in section 3. We emphasize that, for both the heat equation and the Burgers equation, and for all norms and bases considered, *the convergence rates for the DQ case were much higher than (and usually at least twice as high as) the corresponding convergence rates for the no_DQ case.*

The theoretical error estimates in section 3 and the numerical results in section 4 strongly suggest the following conjecture: *"The snapshot DQs should be used in the generation of the POD basis in order to achieve optimal pointwise-in-time rates of convergence with respect to $r$, the number of POD basis functions utilized in the POD-G-ROM.* We also conjecture that using the snapshot DQs in the generation of the POD basis could alleviate some of the degrading of convergence with respect to $r$ seen in, e.g., [29, 9, 3, 8]. We intend to investigate this conjecture in a future study.

To keep the length of the paper within reasonable limits, for each of the two cases investigated (i.e., DQ and no_DQ), we considered two subcases (i.e., using the $L^2$-POD basis or the $H^1$-POD basis). One could, however, consider several additional options. For example, one could replace the optimality of the error estimates given in Definition 3.1 with the alternative definition proposed in Remark 3.1. One could also consider variable parameters $M$, $\Delta t$, and $d$ (these parameters were fixed in the present study; see Remark 2.2). Finally, one could try to use the recent sharper error estimates for the POD projection error proved in [33]. Each of these options should be investigated both theoretically and numerically to determine whether the conclusions regarding the optimality of the POD error estimates in the two cases (i.e., DQ and no_DQ) would change. We plan to pursue some of these research directions in future studies.

**Acknowledgment.** We thank the referees for their constructive comments and suggestions, which significantly improved the manuscript.

REFERENCES

[1] D. Amsallem and U. Hetmaniuk, *Error estimates for Galerkin reduced-order models of the semi-discrete wave equation*, ESAIM Math. Model. Numer. Anal., 48 (2014), pp. 135–163.
[2] J. A. Atwell and B. B. King, *Reduced order controllers for spatially distributed systems via proper orthogonal decomposition*, SIAM J. Sci. Comput., 26 (2004), pp. 128–151.
[3] J. Baiges, R. Codina, and S. Idelsohn, *Explicit reduced order models for the stabilized finite element approximation of the incompressible Navier-Stokes equations*, Internat. J. Numer. Methods Fluids, 72 (2013), pp. 1219–1243.
[4] M. J. Balajewicz, E. H. Dowell, and B. R. Noack, *Low-dimensional modelling of high-Reynolds-number shear flows incorporating constraints from the Navier–Stokes equation*, J. Fluid Mech., 729 (2013), pp. 285–308.
[5] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera, *An 'empirical interpolation' method: Application to efficient reduced-basis discretization of partial differential equations*, C. R. Acad. Sci. Paris, Sér. I Math., 339 (2004), pp. 667–672.
[6] T. Bui-Thanh, K. Willcox, and O. Ghattas, *Model reduction for large-scale systems with high-dimensional parametric input space*, SIAM J. Sci. Comput., 30 (2008), pp. 3270–3288.
[7] J. Burkardt, M. Gunzburger, and H. Lee, *POD and CVT-based reduced-order modeling of Navier–Stokes flows*, Comput. Methods Appl. Mech. Engrg., 196 (2006), pp. 337–355.
[8] A. Caiazzo, T. Iliescu, V. John, and S. Schyschlowa, *A numerical investigation of velocity-pressure reduced order models for incompressible flows*, J. Comput. Phys., 259 (2014), pp. 598–616.
[9] K. Carlberg, C. Bou-Mosleh, and C. Farhat, *Efficient non-linear model reduction via a least-squares Petrov–Galerkin projection and compressive tensor approximations*, Internat. J. Numer. Methods Engrg., 86 (2011), pp. 155–181.
[10] 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.
[11] S. Chaturantabut and D. C. Sorensen, *Nonlinear model reduction via discrete empirical interpolation*, SIAM J. Sci. Comput., 32 (2010), pp. 2737–2764.
[12] 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.
[13] M. Drohmann, B. Haasdonk, and M. Ohlberger, *Reduced basis approximation for nonlinear parametrized evolution equations based on empirical operator interpolation*, SIAM J. Sci. Comput., 34 (2012), pp. A937–A969.
[14] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer Ser. Comput. Math. 5, Springer-Verlag, Berlin, 1986.
[15] S. Herkt, M. Hinze, and R. Pinnau, *Convergence analysis of Galerkin POD for linear second order evolution equations*, Electron. Trans. Numer. Anal., 40 (2013), pp. 321–337.
[16] P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge, Cambridge, 1996.
[17] D. Hömberg and S. Volkwein, *Control of laser surface hardening by a reduced-order approach using proper orthogonal decomposition*, Math. Comput. Model., 38 (2003), pp. 1003–1028.
[18] C. Homescu, L. R. Petzold, and R. Serban, *Error estimation for reduced-order models of dynamical systems*, SIAM Rev., 49 (2007), pp. 277–299.
[19] T. Iliescu and Z. Wang, *Variational multiscale proper orthogonal decomposition: Convection-dominated convection-diffusion-reaction equations*, Math. Comp., 82 (2013), pp. 1357–1378.
[20] T. Iliescu and Z. Wang, *Variational multiscale proper orthogonal decomposition: Navier-Stokes equations*, Numer. Methods Partial Differential Equations, 30 (2014), pp. 641–663.
[21] 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.
[22] K. Kunisch and S. Volkwein, *Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition*, J. Optim. Theory Appl., 102 (1999), pp. 345–371.
[23] K. Kunisch and S. Volkwein, *Galerkin proper orthogonal decomposition methods for parabolic problems*, Numer. Math., 90 (2001), pp. 117–148.

[24] 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.
[25] A. Labovschii, *A defect correction method for the time-dependent Navier-Stokes equations*, Numer. Methods Partial Differential Equations, 25 (2009), pp. 1–25.
[26] W. J. Layton, *Introduction to the Numerical Analysis of Incompressible Viscous Flows*, Comput. Sci. Eng. 6, SIAM, Philadelphia, 2008.
[27] Z. D. Luo, J. Chen, P. Sun, and X. Z. Yang, *Finite element formulation based on proper orthogonal decomposition for parabolic equations*, Sci. China Ser. A, 52 (2009), pp. 585–596.
[28] 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.
[29] C. W. Rowley, T. Colonius, and R. M. Murray, *Model reduction for compressible flows using POD and Galerkin projection*, Phys. D, 189 (2004), pp. 115–129.
[30] G. Rozza, D. B. P. Huynh, and A. T. Patera, *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations*, Arch. Comput. Methods Engrg., 15 (2008), pp. 229–275.
[31] E. Sachs and M. Schu, *A priori error estimates for reduced order models in finance*, ESAIM Math. Model. Numer. Anal., 47 (2013), pp. 449–469.
[32] L. Shan, H. Zheng, and W. J. Layton, *A decoupling method with different subdomain time steps for the nonstationary Stokes–Darcy model*, Numer. Methods Partial Differential Equations, 29 (2013), pp. 549–583.
[33] 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.
[34] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer Verlag, Berlin, 2006.
[35] K. Urban and A. T. Patera, *A new error bound for reduced basis approximation of parabolic partial differential equations*, C. R. Acad. Sci. Paris Sér. I Math., 350 (2012), pp. 203–207.
[36] K. Veroy and A. T. Patera, *Certified real-time solution of the parametrized steady incompressible Navier–Stokes equations: Rigorous reduced-basis a posteriori error bounds*, Internat. J. Numer. Methods Fluids, 47 (2005), pp. 773–788.
[37] S. Volkwein, *Model Reduction Using Proper Orthogonal Decomposition*, Lecture Notes, Faculty of Mathematics and Statistics, University of Konstanz, Konstanz, Germany, (2011).
[38] Z. Wang, I. Akhtar, J. Borggaard, and T. Iliescu, *Proper orthogonal decomposition closure models for turbulent flows: A numerical comparison*, Comput. Methods Appl. Mech. Engrg., 237–240 (2012), pp. 10–26.
[39] M. F. Wheeler, *A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations*, SIAM J. Numer. Anal., 10 (1973), pp. 723–759.