Mathematical Models and Methods in Applied Sciences

A SPACE-TIME *hp*-INTERPOLATION-BASED CERTIFIED REDUCED BASIS METHOD FOR BURGERS' EQUATION

MASAYUKI YANO

Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
myano@mit.edu

ANTHONY T. PATERA

Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
patera@mit.edu

KARSTEN URBAN

Institute for Numerial Mathematics, Ulm University, Helmholtzstrasse 18, 89081, Ulm, Germany
karsten.urban@uni-ulm.de

Received (Day Month Year)
Revised (Day Month Year)
Communicated by (xxxxxxxxxx)

We present a space-time interpolation-based certified reduced basis method for Burgers' equation over the spatial interval $(0, 1)$ and the temporal interval $(0, T]$ parametrized with respect to the Peclet number. We first introduce a Petrov-Galerkin space-time finite element discretization which enjoys a favorable inf-sup constant that decreases slowly with Peclet number and final time T. We then consider an *hp* interpolation-based space-time reduced basis approximation and associated Brezzi-Rappaz-Raviart *a posteriori* error bounds. We describe computational *offline-online* decomposition procedures for the three key ingredients of the error bounds: the dual norm of the residual, a lower bound for the inf-sup constant, and the space-time Sobolev embedding constant. Numerical results demonstrate that our space-time formulation provides improved stability constants compared to classical L^2-error estimates; the error bounds remain sharp over a wide range of Peclet numbers and long integration times T, in marked contrast to the exponentially growing estimate of the classical formulation for high Peclet number cases.

Keywords: space-time variational formulation, parametrized parabolic equations, quadratic noninearity, reduced basis, Brezzi-Rappaz-Raviart theory, *a posteriori* error bounds, inf-sup constant, Burgers' equation

AMS Subject Classification: 22E46, 53C35, 57S20

1. Introduction

In this paper, we develop a certified reduced basis method for the parametrized unsteady Burgers' equation. Classically, parametrized parabolic partial differential equations (PDEs) are treated by collecting solution snapshots in the parameter-time space and by constructing the reduced basis space using the proper orthogonal decomposition of the snapshots.[5,6,11,8] Such a formulation enables rapid approximation of parametrized PDEs by an *offline-online* computational decomposition, and the reduced basis solution converges exponentially to the truth finite element solution for sufficiently regular problems. However, the quality of the associated L^2-in-time *a posteriori* error bound relies on the coercivity of the spatial operator. If the spatial operator is noncoercive, the formulation suffers from exponential temporal instability, producing error bounds that grow exponentially in time, and rendering the bounds meaningless for long-time integration. In particular, limited applicability of the classical *a posteriori* error bounding technique to unsteady Burgers' and Boussinesq equations is documented by Nguyen *et al.*[11] and Knezevic *et al.*[8], respectively.

In order to overcome the instability of the classical L^2-in-time error-bound formulation, we follow the space-time approach recently devised by Urban and Patera[15,14]: we consider a space-time variational formulation and associated finite element approximation that produces a favorable inf-sup stability constant; we then incorporate the space-time truth within a space-time reduced basis approach. The approach is inspired by the recent work on the space-time Petrov-Galerkin formulation by Schwab and Stevenson[13].

The main contribution of this work is twofold. First is the application of the space-time finite-element and reduced-basis approach to the unsteady Burgers' equation with quadratic nonlinearity. The formulation results in Crank-Nicolson-like time-marching procedure but benefits from full space-time variational interpretation and favorable inf-sup stability constant. The second contribution is the application of the Brezzi-Rappaz-Raviart theory to the space-time formulation to construct an error bound in the case of a quadratic nonlinearity. Particular attention is given to the development of an efficient computation procedure that permits *offline-online* decomposition for the three key ingredients of the theory within the space-time framework: the dual norm of the residual, an inf-sup lower bound, and the Sobolev embedding constant.

This paper is organized as follows. Section 2 reviews the spaces and forms used throughout this paper and introduces a space-time Petrov-Galerkin variational formulation and associated finite element approximation of the Burgers' equation. Section 3 first presents an *hp* interpolation-based reduced basis approximation and then an associated *a posteriori* error estimate based on the Brezzi-Rappaz-Raviart theory. We describe the calculation of the dual-norm of the residual, an inf-sup lower bound, and the space-time Sobolev embedding constant, paying particular attention to the *offline-online* computational decomposition. Finally, Section 4 considers

two examples of Burgers' problems and demonstrates that the new space-time error bound provides a meaningful error estimate even for noncoercive cases for which the classical estimate fails. We also demonstrate that the *hp* interpolation method provides certified solutions over a wide range of parameters using a reasonable number of points. Although we consider a single-parameter, one-dimensional Burgers' equation in order to simplify the presentation and facilitate numerical tests, the method extends to the multi-dimensional incompressible Navier-Stokes equations and several parameters as will be considered in future work.[17]

2. Truth Solution

2.1. *Governing Equation*

This work considers a parametrized, unsteady, one-dimensional Burgers' equation of the form

$$\frac{\partial \tilde{u}}{\partial \tilde{t}} + \frac{\partial}{\partial x}\left(\frac{1}{2}\tilde{u}^2\right) - \frac{1}{\mathrm{Pe}}\frac{\partial^2 \tilde{u}}{\partial x^2} = g(x), \quad x \in \Omega,\, \tilde{t} \in \tilde{I}, \tag{2.1}$$

where $\tilde{u}$ is the state variable, Pe is the Peclet number, g is the forcing term, $\Omega \equiv (0,1)$ is the unit one-dimensional domain, and $I \equiv (0, \tilde{T}]$ is the temporal interval with $\tilde{T}$ denoting the final time of interest. We impose homogeneous Dirichlet boundary conditions,

$$\tilde{u}(0,t) = \tilde{u}(1,t) = 0, \quad \forall t \in I,$$

and set the initial condition to

$$\tilde{u}(x,0) = 0, \quad \forall x \in \Omega.$$

Setting $t = \tilde{t}/\mathrm{Pe}$ and $u = \mathrm{Pe} \cdot \tilde{u}$, (2.1) simplifies to

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}u^2\right) - \frac{\partial^2 u}{\partial x^2} = \mathrm{Pe}^2 \cdot g(x), \quad x \in \Omega,\, t \in I. \tag{2.2}$$

Note that the transformation makes the left-hand side of the equation independent of the parameter Pe. The homogeneous boundary conditions and the initial condition are unaltered by the transformation. Moreover, note that $T = \mathcal{O}(1)$ represents a long time integration from $t = 0$ to $\tilde{T} = \mathcal{O}(\mathrm{Pe})$ based on the convection time scale. From hereon, we will exclusively work with this transformed form of the Burgers' equation, (2.2).

2.2. *Spaces and Forms*

Let us now define a few spaces and forms that are used throughout this paper.[12] The standard $L^2(\Omega)$-Hilbert space over $\Omega \equiv (0,1)$ is equipped with an inner product $(\psi,\phi)_{L^2(\Omega)} \equiv \int_\Omega \psi\phi dx$ and induced norm $\|\psi\|_{L^2(\Omega)} \equiv \sqrt{(\psi,\psi)_{L^2(\Omega)}}$ for functions $\{\psi : \|\psi\|_{L^2(\Omega)} < \infty\}$; for convenience, we set $H \equiv L^2(\Omega)$. The space V is equipped with an inner product $(\psi,\phi)_V \equiv \int_\Omega \frac{\partial\psi}{\partial x}\frac{\partial\phi}{\partial x} dx$ and induced norm $\|\psi\|_V \equiv \sqrt{(\psi,\psi)_V}$

for functions $H^1_0(\Omega)$. We define the Gelfand triple (V, H, V') and associated duality paring $\langle \cdot, \cdot \rangle_{V' \times V}$. Here the norm of $\ell \in V'$ is defined by $\|\ell\|_{V'} \equiv \sup_{\phi \in V} \frac{\langle \ell, \phi \rangle_{V' \times V}}{\|\phi\|_V}$, which is identical to $\|R\ell\|_V$ where the Riesz operator $R : V' \to V$ satisfies, for each $\ell \in V'$, $(R\ell, \phi)_V = \langle \ell, \phi \rangle_{V' \times V}$, $\forall \phi \in V$.

Let us now define Bochner spaces, which play key roles in our space-time formulation. The space $L^2(I; V)$ is equipped with an inner product

$$(w, v)_{L^2(I;V)} \equiv \int_I (w(t), v(t))_V dt$$

and induced norm $\|w\|_{L^2(I;V)} \equiv \sqrt{(w, w)_{L^2(I;V)}}$. The dual space $L^2(I; V')$ is equipped with an inner product

$$(w, v)_{L^2(I;V')} \equiv \int_I (Rw(t), Rv(t))_V dt = \int_I (w(t), v(t))_{V'} dt$$

and induced norm $\|w\|_{L^2(I;V')} \equiv \sqrt{(w, w)_{L^2(I;V')}}$, where $R : V' \to V$ is the aforementioned Riesz operator. The space $H^1_{(0)}(I; V')$ is equipped with an inner product $(w, v)_{H^1(I;V')} \equiv (\dot{w}, \dot{v})_{L^2(I;V')}$ and induced norm $\|w\|_{H^1(I;V')} \equiv \sqrt{(w, w)_{H^1(I;V')}}$ for functions $\{w : \|w\|_{H^1(I;V')} < \infty, w(0) = 0\}$; here $\dot{w} \equiv \frac{\partial w}{\partial t}$ denotes the temporal derivative of w. The trial space for our space-time Burgers' formulation is

$$\mathcal{X} \equiv L^2(I; V) \cap H^1_{(0)}(I; V')$$

equipped with an inner product

$$(w, v)_{\mathcal{X}} \equiv (w, v)_{H^1(I;V')} + (w, v)_{L^2(I;V)} + (w(T), v(T))_H$$

and induced norm $\|w\|_{\mathcal{X}} \equiv \sqrt{(w, w)_{\mathcal{X}}}$, as introduced by Urban and Patera for (linear) advection-diffusion equations.[15,14] Note that $\|w\|^2_{\mathcal{X}} = \|w\|^2_{H^1(I;V')} + \|w\|^2_{L^2(I;V)} + \|w(T)\|^2_H$. The test space is $\mathcal{Y} \equiv L^2(I; V)$ equipped with inner product and norm in $L^2(I; V)$.

Having defined spaces, we are ready to express the governing equation (2.2) in a weak form. We may seek a solution to the Burgers' equation expressed in a semi-weak form: find $\psi \in C^1_{(0)}(I; L^2(\Omega)) \cap L^2(I; V)$ such that[12]

$$(\dot{\psi}(t), \phi)_H + a(\psi(t), \phi) + b(\psi(t), \psi(t), \phi) = f(\phi; \mathrm{Pe}), \quad \forall \phi \in V, \forall t \in I,$$

where C^p is the space of functions with continuous p-th derivative, and $C^p_{(0)}$ is the subspace of C^p that consists of functions that satisfy the zero initial condition. The bilinear form $a(\cdot, \cdot)$, the trilinear form $b(\cdot, \cdot, \cdot)$, and the parametrized linear form $f(\cdot; \mathrm{Pe})$ are given by

$$\begin{aligned} a(\psi, \phi) &\equiv \int_\Omega \frac{\partial \psi}{\partial x} \frac{\partial \phi}{\partial x} dx, \quad \forall \psi, \phi \in V, \\ b(\psi, \zeta, \phi) &\equiv -\frac{1}{2} \int_\Omega \psi \zeta \frac{d\phi}{dx} dx, \quad \forall \psi, \zeta, \phi \in V, \\ f(\phi; \mathrm{Pe}) &\equiv \mathrm{Pe}^2 \cdot \langle g, \phi \rangle_{V' \times V}, \quad \forall \phi \in V. \end{aligned}$$

Note that the trilinear form $b(\cdot,\cdot,\cdot)$ is symmetric in the first two arguments. By choosing $\mu = \mathrm{Pe}^2$, we may express the linear form as a linear function of the parameter μ, i.e.

$$f(\phi;\mu) \equiv \mu \cdot \langle g, \phi\rangle_{V'\times V}.$$

Thus, our linear form permits a so-called affine decomposition with respect to the parameter μ. (We note that the certified reduced-basis formulation presented in this work readily treats any f that is affine in a function of parameter μ though the work is presented for the simple single-parameter case above.)

More generally, we may seek the solution to the Burgers' equation in the space-time space $\mathcal{X}$. We have the following space-time weak statement: find $u \in \mathcal{X}$ such that

$$\mathcal{G}(u, v;\mu) = 0, \quad \forall v \in \mathcal{Y}, \tag{2.3}$$

where the semilinear form $\mathcal{G}(\,\cdot\,,\,\cdot\,;\mu)$ is given by

$$\mathcal{G}(w, v;\mu) = \mathcal{M}(\dot{w}, v) + \mathcal{A}(w, v) + \mathcal{B}(w, w, v) - \mathcal{F}(v;\mu), \quad \forall w \in \mathcal{X},\, \forall v \in \mathcal{Y}, \tag{2.4}$$

with the space-time forms

$$\begin{aligned}
\mathcal{M}(\dot{w}, v) &\equiv \int_I \langle \dot{w}(t), v(t)\rangle_{V'\times V} dt, \quad \forall w \in \mathcal{X}, \forall v \in \mathcal{Y},\\
\mathcal{A}(w, v) &\equiv \int_I a(w(t), v(t)) dt, \quad \forall w \in \mathcal{X}, \forall v \in \mathcal{Y},\\
\mathcal{B}(w, z, v) &\equiv \int_I b(w(t), z(t), v(t)) dt, \quad \forall w, z \in \mathcal{X}, \forall v \in \mathcal{Y},\\
\mathcal{F}(v;\mu) &\equiv \mu \cdot \int_I \langle g, v(t)\rangle_{V'\times V} dt, \quad \forall v \in \mathcal{Y}.
\end{aligned}$$

Note that the trilinear form $\mathcal{B}(\,\cdot\,,\,\cdot\,,\,\cdot\,)$ inherits the symmetry with respect to the first two arguments. Furthermore, we will denote the Fréchet derivative bilinear form associated with $\mathcal{G}$ by $\partial\mathcal{G}$, i.e.

$$\partial\mathcal{G}(w, z, v) = \mathcal{M}(\dot{w}, v) + \mathcal{A}(w, v) + 2\mathcal{B}(w, z, v), \quad \forall w, z \in \mathcal{X},\, \forall v \in \mathcal{Y},$$

where $z \in \mathcal{X}$ is the linearization point.

Let us note a few important properties of our unsteady Burgers' problem. First, our space-time linear form $\mathcal{F}$ permits trivial affine-decomposition, i.e. $\mathcal{F}(v;\mu) = \mu\mathcal{F}_0(v)$ where $\mathcal{F}_0 = \int_I \langle g, v(t)\rangle_{V'\times V} dt$. Second, our trilinear form is bounded by

$$|\mathcal{B}(w, z, v)| \equiv \left| \int_I \int_\Omega -\frac{1}{2} wz \frac{\partial v}{\partial x} dx dt \right| \le \frac{1}{2}\rho^2 \|w\|_{\mathcal{X}} \|z\|_{\mathcal{X}} \|v\|_{\mathcal{Y}}, \quad \forall w, z \in \mathcal{X},\, \forall v \in \mathcal{Y},$$

where ρ is the L^4-$\mathcal{X}$ embedding constant

$$\rho \equiv \sup_{w\in\mathcal{X}} \frac{\|w\|_{L^4(I;L^4(\Omega))}}{\|w\|_{\mathcal{X}}}.$$

Recall that the L^p norm is defined as $\|w\|_{L^p(I;L^p(\Omega))} \equiv \left(\int_I \int_\Omega w^p dx dt\right)^{1/p}$. This second property plays an important role in the application of the Brezzi-Rappaz-Raviart theory to construct an *a posteriori* error bound. We note that the existence and boundedness of ρ is not trivial; we have a numerical demonstration in $\Omega \subset \mathbb{R}^1$ and $\Omega \subset \mathbb{R}^2$ but an indication that alternative norms will be required in $\mathbb{R}^3$.[9] Although we consider only Burgers' equation in this paper, we can readily extend the formulation to any quadratically nonlinear equation which satisfies suitable hypotheses on the forms, as implicitly verified above for Burgers'. (We can also consider non-time-invariant operators subject to the usual affine restrictions.)

2.3. *Petrov-Galerkin Finite Element Approximation*

In order to find a discrete approximation to the true solution $u \in \mathcal{X}$, let us introduce finite dimensional subspaces $\mathcal{X}_\delta \subset \mathcal{X}$ and $\mathcal{Y}_\delta \subset \mathcal{Y}$. The notation used in this section closely follows that of Urban and Patera.[14] We denote the triangulations of the temporal interval and spatial domain by $\mathcal{T}_{\Delta t}^{\text{time}}$ and $\mathcal{T}_h^{\text{space}}$, respectively. In particular, $\mathcal{T}_{\Delta t}^{\text{time}}$ consists of non-overlapping intervals $I^k = (t^{k-1}, t^k]$ of length $\Delta t^k \equiv t^k - t^{k-1}$, $k = 1, \ldots, K$, with $t^0 = 0$ and $t^K = T$; here $\max_k(\Delta t^k)/T \leq \Delta t$ and the family $\{\mathcal{T}_{\Delta t}\}_{\Delta t \in (0,1]}$ is assumed to be quasi-uniform. Similarly, $\mathcal{T}_h^{\text{space}}$ consists of $\mathcal{N}+1$ elements with $\max_{\kappa \in \mathcal{T}_h} \text{diam}(\kappa) \leq h$, belonging to a quasi-uniform family of meshes. We now introduce a temporal trial space $S_{\Delta t}$, a temporal test space $Q_{\Delta t}$, and a spatial approximation space V_h defined by

$$
\begin{aligned}
S_{\Delta t} &\equiv \{v \in H^1_{(0)}(I) : v|_{I^k} \in \mathbb{P}^1(I^k),\ k = 1, \ldots, K\},\\
Q_{\Delta t} &\equiv \{v \in L^2(I) : v|_{I^k} \in \mathbb{P}^0(I^k),\ k = 1, \ldots, K\},\\
V_h &\equiv \{v \in H^1_0(\Omega) : v|_\kappa \in \mathbb{P}^1(\kappa), \kappa \in \mathcal{T}_h\}.
\end{aligned}
$$

Our space-time finite element trial and test spaces are given by

$$
\mathcal{X}_\delta = S_{\Delta t} \otimes V_h \quad \text{and} \quad \mathcal{Y}_\delta = Q_{\Delta t} \otimes V_h,
$$

respectively, where $\delta \equiv (\Delta t, h)$ is the characteristic scale of our space-time discretization. Furthermore, we equip the space $\mathcal{X}_\delta$ with a mesh-dependent inner product[15,14]

$$
(w, v)_{\mathcal{X}_\delta} \equiv (w, v)_{H^1(I;V')} + (\bar{w}, \bar{v})_{L^2(I;V)} + (w(T), v(T))_H;
$$

here $\bar{w} \in \mathcal{Y}_\delta$ is a temporally piecewise constant function whose value over I^k is the temporal average of the function $w \in \mathcal{X}_\delta$, i.e.

$$
\bar{w}^k \equiv \frac{1}{\Delta t^k} \int_{I^k} w dt, \quad k = 1, \ldots, K.
$$

We also introduce an associated induced norm $\|w\|^2_{\mathcal{X}_\delta} = (w, w)_{\mathcal{X}_\delta}$. The choice of this mesh-dependent norm is motivated by the fact that the norm provides the unity inf-sup and continuity constant for the Petrov-Galerkin finite element discretization of the heat equation.[15,14] In other words, $\|\cdot\|_{\mathcal{X}_\delta}$ is the natural norm associated with

the Petrov-Galerkin finite element discretization of the heat equation. The space $\mathcal{Y}_\delta$ inherits the inner product and induced norm from the space $\mathcal{Y}$; i.e., $(w,v)_{\mathcal{Y}_\delta} \equiv (w,v)_{\mathcal{Y}}$ and $\|w\|_{\mathcal{Y}_\delta} \equiv \|w\|_{\mathcal{Y}}$.

We have the following space-time finite element approximation to the Burgers' equation (2.3): find $u_\delta \in \mathcal{X}_\delta$ such that

$$\mathcal{G}(u_\delta, v_\delta; \mu) = 0, \quad \forall v_\delta \in \mathcal{Y}_\delta. \tag{2.5}$$

The well-posedness of the space-time finite element formulation will be verified *a posteriori* using the Brezzi-Rappaz-Raviart theory. The temporal integration required for the evaluation of the source term $\mathcal{F}$ is performed using the trapezoidal rule.

2.4. *Algebraic Forms and Time-Marching Interpretation*

In this subsection, we construct algebraic forms of temporal, spatial, and space-time operators required for the computation of our finite element approximation, various norms, and inf-sup constants. In addition, we demonstrate that our Petrov-Galerkin finite element formulation can in fact be written as a time-stepping scheme for a particular set of trial and test basis functions.

Throughout this section, we will use standard hat-functions σ^k with the node at t^k, $k = 1, \ldots, K$, as our basis functions for $S_{\Delta t}$; note that $\text{supp}(\sigma^k) = I^k \cup I^{k+1}$ (except for σ^K, which is truncated to have $\text{supp}(\sigma^K) = I^K$). We further choose characteristic functions $\tau^k = \chi_{I^k}$ as our basis functions for $Q_{\Delta t}$. Finally, let ϕ_i, $i = 1, \ldots, \mathcal{N}$, be standard hat-functions for V_h. With the specified basis, we can express a space-time trial function $w_\delta \in \mathcal{X}_\delta$ in terms of basis coefficients $\{w_i^k\}_{i=1,\ldots,\mathcal{N}}^{k=1,\ldots,K}$ as $w_\delta = \sum_{k=1}^K \sum_{i=1}^{\mathcal{N}} w_i^k \sigma^k \otimes \phi_i$; similarly a trial function $v_\delta \in \mathcal{Y}_\delta$ may be expressed as $v_\delta = \sum_{k=1}^K \sum_{i=1}^{\mathcal{N}} v_i^k \tau^k \otimes \phi_i$. The following sections introduce temporal, spatial, and space-time matrices and their explicit expressions that facilitate evaluation of the residual, norms, and inf-sup constants in the subsequent sections.

2.4.1. *Temporal Operators*

First, let us form temporal matrices required for the evaluation of the Petrov-Galerkin finite element semilinear form. We will explicitly determine the entries of the matrices (i.e. the inner products) for our particular choice of basis functions, which are later required to construct a time-marching interpretation. The Petrov-Galerkin temporal matrices $\mathbf{M}_h^{\text{time}} \in \mathbb{R}^{K\times K}$ and $\dot{\mathbf{M}}_h^{\text{time}} \in \mathbb{R}^{K\times K}$ are given by

$$(\dot{\mathbf{M}}_{\Delta t}^{\text{time}})_{lk} = (\dot{\sigma}^k, \tau^l)_{L^2(I)} = \delta_{k,l} - \delta_{k+1,l}$$
$$(\mathbf{M}_{\Delta t}^{\text{time}})_{lk} = (\sigma^k, \tau^l)_{L^2(I)} = \frac{\Delta t^l}{2}(\delta_{k,l} + \delta_{k+1,l}),$$

where $\delta_{k,l}$ is the Kronecker delta. Note that, with our particular choice of basis functions for $S_{\Delta t}$ and $Q_{\Delta t}$, the matrices are lower bidiagonal. The triple product

resulting from the trilinear form evaluates to

$$(\sigma^k \sigma^m, \tau^l)_{L^2(I)} = \frac{\Delta t^l}{6} \left(2\delta_{k,l}\delta_{m,l} + \delta_{k,l}\delta_{m+1,l} + \delta_{k+1,l}\delta_{m,l} + 2\delta_{k+1,l}\delta_{m+1,l}\right) \quad \text{(no sum implied on } l\text{)}.$$

In addition, evaluation of the $\mathcal{X}_\delta$ inner product requires matrices $\dot{\mathbf{M}}^S_{\Delta t} \in \mathbb{R}^{K\times K}$ and $\overline{\mathbf{M}}^S_{\Delta t} \in \mathbb{R}^{K\times K}$ associated with $S_{\Delta t}$ given by

$$(\dot{\mathbf{M}}^S_{\Delta t})_{lk} = (\dot{\sigma}^k, \dot{\sigma}^l)_{L^2(I)} = -\frac{1}{\Delta t^l}\delta_{k+1,l} + \left(\frac{1}{\Delta t^l} + \frac{1}{\Delta t^{l+1}}\right)\delta_{k,l} - \frac{1}{\Delta t^{l+1}}\delta_{k-1,l}$$

$$(\overline{\mathbf{M}}^S_{\Delta t})_{lk} = (\bar{\sigma}^k, \bar{\sigma}^l)_{L^2(I)} = \frac{\Delta t^l}{4}\delta_{k+1,l} + \frac{\Delta t^l + \Delta t^{l+1}}{4}\delta_{k,l} + \frac{\Delta t^{l+1}}{4}\delta_{k-1,l}$$

with an interpretation $1/\Delta t^{K+1} = 0$ for $\dot{\mathbf{M}}^S_{\Delta t}$ and $\Delta t^{K+1} = 0$ for $\overline{\mathbf{M}}^S_{\Delta t}$. Because the support of the basis functions are unaltered by differentiation or the averaging operation, both $\dot{\mathbf{M}}^S_{\Delta t}$ and $\overline{\mathbf{M}}^S_{\Delta t}$ are tridiagonal. Finally, the evaluation of the $\mathcal{Y}$ inner product requires a matrix $\mathbf{M}^Q_{\Delta t} \in \mathbb{R}^{K\times K}$ associated with $Q_{\Delta t}$ given by

$$(\mathbf{M}^Q_{\Delta t})_{lk} = (\tau^k, \tau^l)_{L^2(I)} = \Delta t^l \delta_{k,l}.$$

Because τ^k, $k = 1, \ldots, K$, have element-wise compact support, $\mathbf{M}^Q_{\Delta t}$ is a diagonal matrix.

2.4.2. *Spatial Operators*

The spatial matrices $\mathbf{M}^{\text{space}}_h \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$ and $\mathbf{A}^{\text{space}}_h \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$ associated with the $L^2(\Omega)$ inner product and the bilinear form $a(\,\cdot\,,\,\cdot\,)$ are given by

$$(\mathbf{M}^{\text{space}}_h)_{ji} = (\phi_i, \phi_j)_H \quad \text{and} \quad (\mathbf{A}^{\text{space}}_h)_{ji} = a(\phi_i, \phi_j)\ ;$$

we omit explicit forms since these matrices are standard. To simplify the notation, let us denote the spatial basis coefficients at time t^k by the vector $\boldsymbol{w}^k \in \mathbb{R}^{\mathcal{N}}$, i.e. the j-th entry of $\boldsymbol{w}^k$ is $(\boldsymbol{w}^k)_j = w^k_j$. The vector $\boldsymbol{z}^m \in \mathbb{R}^{\mathcal{N}}$ is defined similarly. Then, we can express the action of the quadratic term in terms of a function $\boldsymbol{b}^{\text{space}}_h : \mathbb{R}^{\mathcal{N}} \times \mathbb{R}^{\mathcal{N}} \to \mathbb{R}^{\mathcal{N}}$, the j-th component of whose output is given by

$$(\boldsymbol{b}^{\text{space}}_h(\boldsymbol{w}^k, \boldsymbol{z}^m))_j = \sum_{i,n=1}^{\mathcal{N}} w^k_i z^m_n b(\phi_i, \phi_n, \phi_j).$$

2.4.3. *Space-Time Operators: Burgers' Equation*

Combining the expressions for the temporal inner products and the spatial operators, the space-time forms evaluated against the test function $\tau^l \otimes \phi_j$ may be

expressed as

$$\begin{aligned}
\mathcal{M}(\dot{w}_\delta, \tau^l \otimes \phi_j) &= \sum_{k=1}^{K}\sum_{i=1}^{\mathcal{N}} w_i^k (\dot{\sigma}^k, \tau^l)_{L^2(I)} (\phi_i, \phi_j)_H = (\mathbf{M}_h^{\text{space}}(\boldsymbol{w}^l - \boldsymbol{w}^{l-1}))_j \\
\mathcal{A}(w_\delta, \tau^l \otimes \phi_j) &= \sum_{k=1}^{K}\sum_{i=1}^{\mathcal{N}} w_i^k (\sigma^k, \tau^l)_{L^2(I)} a(\phi_i, \phi_j) = \frac{\Delta t^l}{2}\left(\mathbf{A}_h^{\text{space}}(\boldsymbol{w}^l + \boldsymbol{w}^{l-1})\right)_j \\
\mathcal{B}(w_\delta, z_\delta, \tau^l \otimes \phi_j) &= \sum_{k,m=1}^{K}\sum_{i,n=1}^{\mathcal{N}} w_i^k z_n^m (\sigma^k \sigma^m, \tau^l)_{L^2(I)} b(\phi_i, \phi_n, \phi_j) \\
&= \sum_{i,n=1}^{\mathcal{N}} \frac{\Delta t^l}{6}\big(2w_i^l z_n^l b(\phi_i, \phi_n, \phi_j) + w_i^l z_n^{l-1} b(\phi_i, \phi_n, \phi_j) \\
&\qquad + w_i^{l-1} z_n^l b(\phi_i, \phi_n, \phi_j) + 2w_i^{l-1} z_n^{l-1} b(\phi_i, \phi_n, \phi_j)\big) \\
&= \frac{\Delta t^l}{6}\big(2\boldsymbol{b}_h^{\text{space}}(\boldsymbol{w}^l, \boldsymbol{z}^l) + \boldsymbol{b}_h^{\text{space}}(\boldsymbol{w}^l, \boldsymbol{z}^{l-1}) \\
&\qquad + \boldsymbol{b}_h^{\text{space}}(\boldsymbol{w}^{l-1}, \boldsymbol{z}^l) + 2\boldsymbol{b}_h^{\text{space}}(\boldsymbol{w}^{l-1}, \boldsymbol{z}^{l-1})\big)_j .
\end{aligned}$$

The trilinear form further simplifies when the first two arguments are the same, as in the case for the semilinear form of the Burgers' equation, (2.4); i.e.

$$\mathcal{B}(w_\delta, w_\delta, \tau^l \otimes \phi_j) = \frac{\Delta t^l}{3}\left(\boldsymbol{b}_h^{\text{space}}(\boldsymbol{w}^l, \boldsymbol{w}^l) + \boldsymbol{b}_h^{\text{space}}(\boldsymbol{w}^l, \boldsymbol{w}^{l-1}) + \boldsymbol{b}_h^{\text{space}}(\boldsymbol{w}^{l-1}, \boldsymbol{w}^{l-1})\right)_j .$$

In addition, the integration of the forcing function using the trapezoidal rule results in

$$\begin{aligned}
\mathcal{F}(\tau^l \otimes \phi_j; \mu) &\equiv \mu \cdot \int_I \langle g_0(t), \tau^l \otimes \phi_j \rangle_{V' \times V} dt \approx \Delta t^l \mu \cdot \frac{1}{2}\langle g_0(t^l) + g_0(t^{l-1}), \phi_j \rangle_{V' \times V} \\
&= \Delta t^l \mu \frac{1}{2}(\boldsymbol{g}_{0,h}^l + \boldsymbol{g}_{0,h}^{l-1})_j,
\end{aligned}$$

where $\boldsymbol{g}^l \in \mathbb{R}^{\mathcal{N}}$ with $(\boldsymbol{g}_h^l)_j = \langle g(t^l), \phi_j \rangle_{V' \times V}$. Combining the expressions for our particular choice of the Petrov-Galerkin test functions, the finite element residual statement, (2.5), may be simplified to

$$\begin{aligned}
&\frac{1}{\Delta t^l}\mathbf{M}_h^{\text{space}}(\boldsymbol{w}^l - \boldsymbol{w}^{l-1}) + \frac{1}{2}\mathbf{A}_h^{\text{space}}(\boldsymbol{w}^l + \boldsymbol{w}^{l-1}) \\
&\quad + \frac{1}{3}\left(\boldsymbol{b}_h^{\text{space}}(\boldsymbol{w}^l, \boldsymbol{w}^l) + \boldsymbol{b}_h^{\text{space}}(\boldsymbol{w}^l, \boldsymbol{w}^{l-1}) + \boldsymbol{b}_h^{\text{space}}(\boldsymbol{w}^{l-1}, \boldsymbol{w}^{l-1})\right) \\
&\quad - \mu\frac{1}{2}(\boldsymbol{g}_{0,h}^l + \boldsymbol{g}_{0,h}^{l-1}) = 0 \quad \text{in } \mathbb{R}^{\mathcal{N}},
\end{aligned}$$

for $l = 1, \ldots, K$, with $\boldsymbol{w}^0 = 0$. Note that the treatment of the linear terms are identical to that of the Crank-Nicolson time stepping, whereas the quadratic term results in a different form. In any event, the Petrov-Galerkin space-time formulation admits a time-marching interpretation; the solution can be obtained by sequentially solving K systems of nonlinear equations, each having $\mathbb{R}^{\mathcal{N}}$ unknowns; thus, the computational cost is equivalent to that of the Crank-Nicolson scheme.

2.4.4. *Space-Time Operators: $\mathcal{X}_\delta$ and $\mathcal{Y}_\delta$ Inner Products*

Combining the temporal matrices with the spatial matrices introduced in Section 2.3, we can express the matrix associated with the $\mathcal{X}_\delta$ inner product, $\mathbf{X} \in \mathbb{R}^{(K\cdot\mathcal{N})\times(K\cdot\mathcal{N})}$, as

$$\mathbf{X} = \dot{\mathbf{M}}^S_{\Delta t} \otimes \left(\mathbf{M}^{\text{space}}_h (\mathbf{A}^{\text{space}}_h)^{-1} \mathbf{M}^{\text{space}}_h\right) + \overline{\mathbf{M}}^S_{\Delta t} \otimes \mathbf{A}^{\text{space}}_h + \text{diag}(\boldsymbol{e}_K) \otimes \mathbf{M}^{\text{space}}_h,$$

where $\text{diag}(\boldsymbol{e}_K)$ is a $K \times K$ matrix with one at the (K, K) entry and zero elsewhere. Note that $\mathbf{X}$ is block-tridiagonal. The norm induced by the $\overline{\mathbf{M}}^S_{\Delta t} \otimes \mathbf{A}^{\text{space}}_h$ part of the $\mathbf{X}$ matrix is identical to the usual norm for the Crank-Nicolson scheme, i.e.

$$\begin{aligned}\{w^k_i\}^T (\overline{\mathbf{M}}^S_{\Delta t} \otimes \mathbf{A}^{\text{space}}_h) \{w^k_i\} &= \|w_\delta\|^2_{\text{CN}} \\ &\equiv \sum_{k=1}^{K} \left(\frac{1}{2}(\boldsymbol{w}^k + \boldsymbol{w}^{k-1})\right)^T \mathbf{A}^{\text{space}}_h \left(\frac{1}{2}(\boldsymbol{w}^k + \boldsymbol{w}^{k-1})\right),\end{aligned}$$

where $\{w^k_i\} \in \mathbb{R}^{K\cdot\mathcal{N}}$ is a vector of space-time basis coefficients for w_δ. The identity — together with the equivalence of our space-time Petrov-Galerkin formulation with the Crank-Nicolson scheme for linear problems — suggests that the inclusion of the averaging operator in our $\mathcal{X}_\delta$ norm is rather natural for the particular scheme we consider. Similarly, the matrix associated with the $\mathcal{Y}_\delta$ inner product, $\mathbf{Y} \in \mathbb{R}^{(K\cdot\mathcal{N})\times(K\cdot\mathcal{N})}$, is given by

$$\mathbf{Y} = \mathbf{M}^Q_{\Delta t} \otimes \mathbf{A}^{\text{space}}_h.$$

The matrix $\mathbf{Y}$ is block diagonal because $\mathbf{M}^Q_{\Delta t}$ is diagonal.

Remark 2.1. As noted previously, the $\mathcal{X}_\delta$ trial norm is in fact the natural norm associated with the Petrov-Galerkin discretization of the heat equation with the $\mathcal{Y}_\delta$ test norm. Thus, it follows that

$$\mathbf{X} = \mathbf{G}^T_{\text{heat}} \mathbf{Y}^{-1} \mathbf{G}_{\text{heat}} \ ,$$

where $\mathbf{G}_{\text{heat}} \in \mathbb{R}^{(K\cdot\mathcal{N})\times(K\cdot\mathcal{N})}$ is the space-time matrix associated with the heat equation, $\mathbf{G}_{\text{heat}} = \dot{\mathbf{M}}^{\text{time}}_{\Delta t} \otimes \mathbf{M}^{\text{space}}_h + \mathbf{M}^{\text{time}}_{\Delta t} \otimes \mathbf{A}^{\text{space}}_h$. This decomposition of the block-tridiagonal matrix $\mathbf{X}$ allows a computationally efficient application of $\mathbf{X}^{-1} = \mathbf{G}^{-1}_{\text{heat}} \mathbf{Y} \mathbf{G}^{-T}_{\text{heat}}$ by a three-step procedure: 1) we first solve the space-time linear system backward in time, i.e. the adjoint solve $\mathbf{G}^{-T}_{\text{heat}}$; we then apply the block-diagonal operator $\mathbf{Y}$; and we finally solve the space-time linear system forward in time ($\mathbf{G}^{-1}_{\text{heat}}$). The efficient application of $\mathbf{X}^{-1}$ plays an important role in our space-time formulation, including in the computation of the L^4-$\mathcal{X}_\delta$ embedding constant (as we will see in Section 3.2.3) and the adjoint residual (which would be required for the primal-dual formulation of output error bounds).

3. Certified Space-Time Reduced-Basis Approximation

3.1. *N_μ-p Interpolation-Based Approximation*

Here, we introduce a simple reduced-basis approximation procedure based on interpolation (rather than projection). We choose interpolation as it is less expensive than projection, sufficiently accurate in one parameter dimension, and also facilitates the construction of an inf-sup lower bound as we will show in Section 3.2.2. We note that interpolation-based model reduction techniques have been employed previously.[1]

We consider an *hp*-decomposition (or, more specifically, N_μ-p decomposition) of the parameter domain $\mathcal{D}$ as considered in Eftang *et al.*[4] In particular, we partition $\mathcal{D} \subset \mathbb{R}^1$ into N_μ subdomains, $\mathcal{D}_j = [\mu_j^L, \mu_j^U]$, $j = 1, \ldots, N_\mu$, and approximate the solution variation over each subdomain using a degree-p polynomial. On each $\mathcal{D}_j$, we use $p+1$ Chebyshev-Lobatto nodes $\mu_{j,k}$, $k = 1, \ldots, p+1$, defined by

$$\frac{\mu_{j,k} - \mu_j^L}{\mu_j^U - \mu_j^L} = \frac{1}{2} - \frac{1}{2}\cos\left(\frac{k-1}{p}\pi\right)$$

as the interpolation points. At each interpolation point, we obtain the truth solution $u_{\delta,j,k} \equiv u_\delta(\mu_{j,k}) \in \mathcal{X}_{\delta,j}$ by solving the finite element approximation (2.5) for the space-time finite element trial-test space pair $\mathcal{X}_{\delta,j}$ and $\mathcal{Y}_{\delta,j}$. Note that we employ a different finite element space pair $(\mathcal{X}_{\delta,j}, \mathcal{Y}_{\delta,j})$ (induced by $\mathcal{T}_{\Delta t,j}^{\text{time}}$ and $\mathcal{T}_{h,j}^{\text{space}}$) for each parameter domain $\mathcal{D}_j$; each space pair is tailored toward resolving the solution encountered over the associated parameter range.

We now construct our reduced basis approximation to $u_\delta = u_\delta(\mu)$ by a direct sum of N_μ polynomials

$$\tilde{u}_\delta^p = \bigoplus_{j=1}^{N_\mu} \tilde{u}_{\delta,j}^p,$$

where $\tilde{u}_{\delta,j}^p$ is a degree-p polynomial over $\mu \in \mathcal{D}_j$ given by

$$\tilde{u}_{\delta,j}^p(\mu) = \sum_{k=1}^{p+1} u_{\delta,j,k}\psi_k^p(\mu) \tag{3.1}$$

for $j = 1, \ldots, N_\mu$. Here ψ_k^p, $k = 1, \ldots, K$, is the degree-p polynomial with the interpolation property through Chebyshev points, i.e. $\psi_k^p \in \mathbb{P}^p(\mathcal{D}_j)$ such that $\psi_k^p(x_l) = \delta_{k,l}$, $k, l = 1, \ldots, p+1$. Given μ, we simply identify the parameter domain to which μ belongs and then evaluate (3.1); the N_μ-p strategy ensures that the polynomial degree (and hence the cost associated with the interpolation) is relatively small. Note that, unlike in the classical time-marching formulation,[6,11,8] the computational cost of constructing the reduced-basis approximation using our space-time formulation is independent of the number of time steps, K. In this work, we do not assess the relative approximation properties of classical time-marching formulation (e.g. POD-Greedy) and our N_μ-p interpolation method.

3.2. *Brezzi-Rappaz-Raviart Theory*

We now construct error bounds for our N_μ-p interpolation-based reduced basis approximation. Our *a posteriori* error bound for the Burgers' equation is a straightforward application of the Brezzi-Rappaz-Raviart (BRR) theory.[2] The bound procedure separately applies to each of the N_μ parameter subdomains; hence, in this section, we focus on the error certification for a *single* subdomain, and, to avoid the notational clutter, we accordingly drop the subscript j for the subdomain designation: we indicate the "working" subdomain by $\mathcal{D}_{\text{work}}$ and denote, for example, the trial space associated with $\mathcal{D}_{\text{work}}$ by $\mathcal{X}_\delta$ (instead of $\mathcal{X}_{\delta,j}$) and the reduced basis approximation by $\tilde{u}^p_\delta$ (instead of $\tilde{u}^p_{\delta,j}$).

The following proposition states the main results of the theory; detailed proof for a general case is provided in the original paper[2] and for quadratic nonlinearity is presented by Veroy and Patera.[16]

Proposition 3.1. *Let $\tilde{u}^p_\delta(\mu)$ be the reduced basis approximation* (3.1)*. We define the dual norm of the residual, the inf-sup constant, and the L^4-$\mathcal{X}_\delta$ Sobolev embedding constant by*

$$\epsilon^p_\delta(\mu) \equiv \sup_{v\in\mathcal{Y}_\delta} \frac{\mathcal{G}(\tilde{u}^p_\delta(\mu), v; \mu)}{\|v\|_{\mathcal{Y}_\delta}},$$
$$\beta^p_\delta(\mu) \equiv \inf_{w\in\mathcal{X}_\delta} \sup_{v\in\mathcal{Y}_\delta} \frac{\partial\mathcal{G}(w, \tilde{u}^p_\delta(\mu), v)}{\|w\|_{\mathcal{X}_\delta}\|v\|_{\mathcal{Y}_\delta}},$$
$$\rho_\delta \equiv \sup_{w\in\mathcal{X}_\delta} \frac{\|w\|_{L^4(I;L^4(\Omega))}}{\|w\|_{\mathcal{X}_\delta}}.$$

In addition, let $\beta^p_{\delta,LB}(\mu)$ be a lower bound of $\beta^p_\delta(\mu)$, i.e. $\beta^p_{\delta,LB}(\mu) \le \beta^p_\delta(\mu)$, $\forall\mu\in\mathcal{D}$. Let the proximity indicator be $\tau^p_\delta(\mu) \equiv 2\rho^2_\delta\epsilon^p_\delta(\mu)/(\beta^p_{\delta,LB}(\mu))^2$. Then, for $\tau^p_\delta(\mu) < 1$, there exists a unique solution $u_\delta(\mu) \in \mathfrak{B}(\tilde{u}^p_\delta(\mu), \beta^p_\delta(\mu)/\rho^2_\delta)$ to the finite element problem (2.5)*, where $\mathfrak{B}(z,r) \equiv \{x\in\mathcal{X}_\delta : \|x-z\|_{\mathcal{X}_\delta} < r\}$. Furthermore, $\|u_\delta(\mu) - \tilde{u}^p_\delta(\mu)\|_{\mathcal{X}_\delta} \le \Delta^p_\delta(\mu)$ where*

$$\Delta^p_\delta(\mu) \equiv \frac{\beta^p_{\delta,LB}(\mu)}{\rho^2_\delta}\left(1 - \sqrt{1-\tau^p_\delta(\mu)}\right).$$

Proof. Proof is provided in, for example, Veroy and Patera.[16] □

The following subsections detail the computation of the three key ingredients of the BRR theory: the dual norm of residual $\epsilon^p_\delta(\mu)$; the inf-sup constant $\beta^p_{\delta,\text{LB}}(\mu)$; and the L^4-$\mathcal{X}_\delta$ Sobolev embedding constant ρ_δ. In particular, we will present efficient means of computing these variables in the space-time context that permits an *offline-online* decomposition.

3.2.1. *Residual Evaluation*

Here, we briefly review a technique for efficiently computing the dual norm of the residual in the online stage, the technique originally presented by Veroy *et al.*[16] in the space-only context. We first note that $\epsilon_\delta^p(\mu) \equiv \|\mathcal{G}(\tilde{u}_\delta^p(\mu),\,\cdot\,;\mu)\|_{(\mathcal{Y}_\delta)'} = \|\hat{e}^p\|_{\mathcal{Y}_\delta}$, where the Riesz representation of the residual is given by $\hat{e}^p \equiv R\mathcal{G}(\tilde{u}_\delta^p(\mu),\,\cdot\,;\mu) \in \mathcal{Y}_\delta$ and satisfies

$$\begin{aligned}(\hat{e}^p, v)_{\mathcal{Y}_\delta} &= \mathcal{G}(\tilde{u}_\delta^p(\mu), v; \mu) \\ &= \sum_{k=1}^{p+1} \psi_k^p(\mu)\left[\mathcal{M}(\dot{u}_{\delta,k}, v) + \mathcal{A}(u_{\delta,k}, v)\right] \\ &\quad + \sum_{k,l=1}^{p+1} \psi_k^p(\mu)\psi_l^p(\mu)\mathcal{B}(u_{\delta,k}, u_{\delta,l}, v) - \mu\cdot\mathcal{F}_0(v), \quad \forall v \in \mathcal{Y}_\delta.\end{aligned}$$

Let us introduce (pieces of) Riesz representations χ^0, $\{\chi_k^1\}_{k=1}^{p+1}$, and $\{\chi_{kl}^2\}_{k,l=1}^{p+1}$ of the residual contribution from the linear, bilinear, and trilinear form, respectively, for the snapshots according to

$$(\chi^0, v)_{\mathcal{Y}_\delta} = \mathcal{F}_0(v), \quad \forall v \in \mathcal{Y}_\delta, \tag{3.2}$$

$$(\chi_k^1, v)_{\mathcal{Y}_\delta} = \mathcal{M}(\dot{u}_{\delta,k}, v) + \mathcal{A}(u_{\delta,k}, v), \quad \forall v \in \mathcal{Y}_\delta,\ k = 1,\ldots,p+1, \tag{3.3}$$

$$(\chi_{kl}^2, v)_{\mathcal{Y}_\delta} = \mathcal{B}(u_{\delta,k}, u_{\delta,l}, v), \quad \forall v \in \mathcal{Y}_\delta,\ k,l = 1,\ldots,p+1. \tag{3.4}$$

Then, we can express $\hat{e}^p$ as

$$\hat{e}^p = \mu\cdot\chi^0 + \sum_{k=1}^{p+1}\psi_k^p(\mu)\chi_k^1 + \sum_{k,l=1}^{p+1}\psi_k^p(\mu)\psi_l^p(\mu)\chi_{kl}^2.$$

The dual norm of the residual can be expressed as

$$\begin{aligned}\|\hat{e}^p\|_{\mathcal{Y}_\delta} &= \mu^2(\chi^0,\chi^0)_{\mathcal{Y}_\delta} + 2\mu\sum_{m=1}^{p+1}(\chi^0,\chi_m^1)_{\mathcal{Y}_\delta} + 2\mu\sum_{m,n=1}^{p+1}(\chi^0,\chi_{mn}^2)_{\mathcal{Y}_\delta} \\ &\quad + \sum_{k,m=1}^{p+1}\psi_k^p(\mu)\psi_m^p(\mu)(\chi_k^1,\chi_m^1)_{\mathcal{Y}_\delta} + 2\sum_{k,m,n=1}^{p+1}\psi_k^p(\mu)\psi_m^p(\mu)\psi_n^p(\mu)(\chi_k^1,\chi_{mn}^2)_{\mathcal{Y}_\delta} \\ &\quad + \sum_{k,l,m,n=1}^{p+1}\psi_k^p(\mu)\psi_l^p(\mu)\psi_m^p(\mu)\psi_n^p(\mu)(\chi_{kl}^2,\chi_{mn}^2)_{\mathcal{Y}_\delta}.\end{aligned} \tag{3.5}$$

The *offline-online* decomposition is clear from the expression. In the *offline* stage, we first solve (3.2)-(3.4) to obtain the Riesz representations χ^0, $\{\chi_k^1\}_{k=1}^{p+1}$, and $\{\chi_{kl}^2\}_{k,l=1}^{p+1}$. Note that there are $1+(p+1)+(p+1)^2$ representations, each requiring a $\mathcal{Y}_\delta$-solve. Recalling that the matrix associated with the $\mathcal{Y}_\delta$ inner product is given by $\mathbf{Y} = \mathbf{M}_{\Delta t}^Q \otimes \mathbf{A}_h^{\text{space}}$, each $\mathcal{Y}_\delta$-solve requires K inversions of the $\mathbf{A}_h^{\text{space}}$ operator, where K is the number of time steps. It is important to note that the computation of the representations does not require a solution of a coupled space-time system,

as the matrix $\mathbf{Y}$ is block diagonal. In other words, the computational cost is not higher than that for the classical time-marching reduced basis formulation. After computing the representations, we compute the $\mathcal{Y}_\delta$ inner product of all permutations of representations, e.g. $(\chi^0, \chi^0)_{\mathcal{Y}_\delta}$, $(\chi^0, \chi^1_k)_{\mathcal{Y}_\delta}$.

In the *online* stage, we obtain the dual norm of the residual by evaluating (3.5) using the inner products computed in the *offline* stage. The computational cost scales as $(p+1)^4$ and is independent of the cost of the truth discretization. Note that, unlike in the classical reduced-basis formulation based on time-marching, the online residual evaluation cost of our space-time formulation is independent of the number of time steps, K.

3.2.2. *Inf-Sup Constant and Associated Lower Bound*

Here, we present a procedure for computing an inf-sup lower bound, $\beta^p_{\delta,\mathrm{LB}}(\mu)$, that permits an *offline-online* decomposition. The particular procedure presented is specifically designed for the N_μ-p interpolation-based reduced basis approximation introduced in Section 3.1. Let us first define the supremizing operator $S^c : \mathcal{X}_\delta \to \mathcal{Y}_\delta$ associated with the solution $u^c_\delta = u(\mu^c)$ at the centroid of the subdomain $\mathcal{D}_{\mathrm{work}}$, μ^c, by

$$(S^c w, v)_{\mathcal{Y}_\delta} = \partial\mathcal{G}(w, u^c_\delta, v), \quad \forall w \in \mathcal{X}_\delta, \forall v \in \mathcal{Y}_\delta.$$

The inf-sup constant for u^c_δ is given by

$$\beta^c_\delta = \inf_{w\in\mathcal{X}_\delta} \frac{\|S^c w\|_{\mathcal{Y}_\delta}}{\|w\|_{\mathcal{X}_\delta}}.$$

Let us also introduce the following correction factors at the interpolation points,

$$\beta^-_{\delta,k} \equiv \inf_{w\in\mathcal{X}_\delta} \frac{\partial\mathcal{G}(w, u_{\delta,k}, S^c w)}{\|S^c w\|^2_{\mathcal{Y}_\delta}} \quad \text{and} \quad \beta^+_{\delta,k} \equiv \sup_{w\in\mathcal{X}_\delta} \frac{\partial\mathcal{G}(w, u_{\delta,k}, S^c w)}{\|S^c w\|^2_{\mathcal{Y}_\delta}}, \tag{3.6}$$

for $k = 1, \ldots, p+1$. Then, we construct an inf-sup lower bound according to

$$\beta^p_{\delta,\mathrm{LB}}(\mu) = \beta^c_\delta \cdot \left(\sum_{\substack{k=1,\ldots,p+1 \\ \psi^p_k(\mu)>0}} \beta^-_{\delta,k}\psi^p_k(\mu) + \sum_{\substack{k=1,\ldots,p+1 \\ \psi^p_k(\mu)<0}} \beta^+_{\delta,k}\psi^p_k(\mu) \right), \quad \forall \mu \in \mathcal{D}_{\mathrm{work}}. \tag{3.7}$$

We have the following proposition:

Proposition 3.2. *The inf-sup lower bound constructed using the above procedure satisfies* $\beta^p_{\delta,LB}(\mu) \le \beta^p_\delta(\mu)$, $\forall \mu \in \mathcal{D}_{work}$.

Proof. Since $S^c w \in \mathcal{Y}_\delta$, $\forall w \in \mathcal{X}_\delta$, we can bound the inf-sup constant from below

as

$$\begin{aligned}\beta_\delta^p(\mu) &\equiv \inf_{w\in\mathcal{X}_\delta}\sup_{v\in\mathcal{Y}_\delta}\frac{\partial\mathcal{G}(w,\tilde{u}_\delta^p(\mu),v)}{\|w\|_{\mathcal{X}_\delta}\|v\|_{\mathcal{Y}_\delta}} = \inf_{w\in\mathcal{X}_\delta}\sup_{v\in\mathcal{Y}_\delta}\sum_{k=1}^{p+1}\psi_k^p(\mu)\frac{\partial\mathcal{G}(w,u_{\delta,k},v)}{\|w\|_{\mathcal{X}_\delta}\|v\|_{\mathcal{Y}_\delta}} \\ &\geq \inf_{w\in\mathcal{X}_\delta}\sum_{k=1}^{p+1}\psi_k^p(\mu)\frac{\partial\mathcal{G}(w,u_{\delta,k},S^cw)}{\|w\|_{\mathcal{X}_\delta}\|S^cw\|_{\mathcal{Y}_\delta}} = \inf_{w\in\mathcal{X}_\delta}\sum_{k=1}^{p+1}\psi_k^p(\mu)\frac{\|S^cw\|_{\mathcal{Y}_\delta}}{\|w\|_{\mathcal{X}_\delta}}\frac{\partial\mathcal{G}(w,u_{\delta,k},S^cw)}{\|S^cw\|_{\mathcal{Y}_\delta}^2} \\ &= \inf_{w\in\mathcal{X}_\delta}\frac{\|S^cw\|_{\mathcal{Y}_\delta}}{\|w\|_{\mathcal{X}_\delta}}\sum_{k=1}^{p+1}\psi_k^p(\mu)\frac{\partial\mathcal{G}(w,u_{\delta,k},S^cw)}{\|S^cw\|_{\mathcal{Y}_\delta}^2}, \quad \forall\mu\in\mathcal{D}_{\rm work}, \end{aligned} \tag{3.8}$$

Note that we have

$$\frac{\|S^cw\|_{\mathcal{Y}_\delta}}{\|w\|_{\mathcal{X}_\delta}} \geq \inf_{z\in\mathcal{X}_\delta}\frac{\|S^cz\|_{\mathcal{Y}_\delta}}{\|z\|_{\mathcal{X}_\delta}} = \beta_\delta(\mu^c) \equiv \beta_\delta^c > 0, \quad \forall w\in\mathcal{X}_\delta,$$

and the first term of (3.8) is bounded from below by $\beta_\delta^c > 0$. The second term involving summation over $p+1$ terms may be bounded from below by the correction factors defined in (3.6). Namely, if $\psi_k^p(\mu) > 0$, then we may bound the contribution from the k-th term from below by using $\beta_{\delta,k}^-$; if $\psi_k^p(\mu) < 0$, then the contribution may be bounded from below by using $\beta_{\delta,k}^+$. In other words, the final expression of (3.8) is bounded from below by

$$\begin{aligned}\beta_\delta^p(\mu) \geq \left(\inf_{w\in\mathcal{X}_\delta}\frac{\|S^cw\|_{\mathcal{Y}_\delta}}{\|w\|_{\mathcal{X}_\delta}}\right)\Bigg(&\sum_{\substack{k=1,\ldots,p+1\\ \psi_k^p(\mu)>0}}\psi_k^p(\mu)\inf_{w\in\mathcal{X}_\delta}\frac{\partial\mathcal{G}(w,u_{\delta,k},S^cw)}{\|S^cw\|_{\mathcal{Y}_\delta}^2} \\ &+\sum_{\substack{k=1,\ldots,p+1\\ \psi_k^p(\mu)<0}}\psi_k^p(\mu)\sup_{w\in\mathcal{X}_\delta}\frac{\partial\mathcal{G}(w,u_{\delta,k},S^cw)}{\|S^cw\|_{\mathcal{Y}_\delta}^2}\Bigg) \\ = \beta_\delta^c\left(\sum_{\substack{k=1,\ldots,p+1\\ \psi_k^p(\mu)>0}}\beta_{\delta,k}^-\psi_k^p(\mu) + \sum_{\substack{k=1,\ldots,p+1\\ \psi_k^p(\mu)<0}}\beta_{\delta,k}^+\psi_k^p(\mu)\right), &\quad \forall\mu\in\mathcal{D}_{\rm work},\end{aligned}$$

which concludes the proof. □

Remark 3.1. For small intervals, the correction factors are close to unity. To see this, we note that

$$\begin{aligned}\frac{|\partial\mathcal{G}(w,u_{\delta,k},S^cw)|}{\|S^cw\|_{\mathcal{Y}_\delta}^2} &= \frac{|\partial\mathcal{G}(w,u_\delta^c+(u_{\delta,k}-u_\delta^c),S^cw)|}{\|S^cw\|_{\mathcal{Y}_\delta}^2} \\ &\leq \frac{|\partial\mathcal{G}(w,u_\delta^c,S^cw)|}{\|S^cw\|_{\mathcal{Y}_\delta}^2} + \frac{|\mathcal{B}(w,u_{\delta,k}-u_\delta^c,S^cw)|}{\|S^cw\|_{\mathcal{Y}_\delta}^2} \\ &\leq 1+\frac{1}{2}\frac{\rho_\delta^2}{\beta_\delta^c}\|u_{\delta,k}-u_\delta^c\|_{\mathcal{X}}\ .\end{aligned}$$

Thus, as $|\mathcal{D}_{\text{work}}| \to 0$ and $\|u_{\delta,k} - u^c_\delta\|_{\mathcal{X}_\delta} \to 0$, the correction factors converge to 1.

Remark 3.2. The inf-sup lower bound construction procedure presented here produces a tighter lower bound than the natural norm Successive Constraint Method (SCM)[7] that uses the $p+1$ interpolations as the SCM sampling points, i.e.

$$\beta^p_\delta(\mu) \geq \beta^p_{\delta,\text{LB}}(\mu) \geq \beta^p_{\delta,\text{LB,SCM}}(\mu), \quad \forall \mu \in \mathcal{D}_{\text{work}},$$

where $\beta^p_{\delta,\text{LB,SCM}}(\mu)$ is the SCM inf-sup lower bound. A detailed derivation is provided in Appendix B.

Again, the *offline-online* decomposition is clear from the structure of (3.7). In the *offline* stage, for each $\mathcal{D}_{\text{work}}$, we evaluate the inf-sup constant β^c_δ at the centroid and correction factors $\beta^\pm_{\delta,k}$ at each of the $p+1$ interpolation points. In the *online* stage, we identify the parameter subdomain $\mathcal{D}_j$ to which μ belongs, set it as $\mathcal{D}_{\text{work}}$, and evaluate $\beta^p_{\delta,\text{LB}}(\mu)$ using (3.7).

Let us demonstrate that none of the *offline* computations require solutions to a fully-coupled space-time problem, and the computational cost scales linearly with K. The inf-sup constant at the centroid, β^c_δ, can be obtained by finding the largest eigenvalue of a generalized eigenproblem $\mathbf{P}\mathbf{v} = \lambda \mathbf{Q}\mathbf{v}$ with

$$\mathbf{P} \equiv \mathbf{X} = \dot{\mathbf{M}}^S_{\Delta t} \otimes \left(\mathbf{M}^{\text{space}}_h (\mathbf{A}^{\text{space}}_h)^{-1} \mathbf{M}^{\text{space}}_h\right) + \overline{\mathbf{M}}^S_{\Delta t} \otimes \mathbf{A}^{\text{space}}_h + \text{diag}(\boldsymbol{e}_K) \otimes \mathbf{M}^{\text{space}}_h$$
$$\mathbf{Q} \equiv (\mathbf{G}^c)^T \mathbf{Y}^{-1} \mathbf{G}^c$$

and setting $\beta^c_\delta = \lambda^{-1/2}_{\max}$. Here, $\mathbf{G}^c \in \mathbb{R}^{(K\cdot\mathcal{N})\times(K\cdot\mathcal{N})}$ is the Jacobian matrix of the residual operator linearized about $u^c_\delta = u_\delta(\mu^c)$; the $(li)(kj)$ entry of the matrix is given by

$$(\mathbf{G}^c)_{(li)(kj)} = \partial\mathcal{G}(\sigma^k \otimes \phi_j, u^c_\delta, \tau^l \otimes \phi_i).$$

Note that $\mathbf{G}^c$ is block lower bidiagonal due to our choice of the basis functions for the spaces $S_{\Delta t}$ and $Q_{\Delta t}$ in the Petrov-Galerkin formulation. If the eigenproblem is solved using a Lanczos-based method, each Lanczos step requires action of $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{Q}^{-1}$ on a vector in $\mathbb{R}^{K\cdot\mathcal{N}}$. The application of $\mathbf{P}$ requires $\mathcal{O}(K)$ operations due to the tensor-product structure of the matrices that constitutes $\mathbf{X}$; for instance, to compute $(\overline{\mathbf{M}}^S_{\Delta t} \otimes \mathbf{A}^{\text{space}}_h)\mathbf{v}$, we first compute $\mathbf{A}^{\text{space}}_h \mathbf{v}^k$, $k = 1, \ldots, K$, and then take a linear combination of (at most) three $\mathbf{A}^{\text{space}}_h \mathbf{v}^k$'s according to the weights specified in $\overline{\mathbf{M}}^S_{\Delta t}$. The application of $\mathbf{Q}$ requires the application of $\mathbf{G}^c$, $(\mathbf{G}^c)^T$, and $\mathbf{Y}^{-1}$, each of which requires $\mathcal{O}(K)$ operations due to the block bidiagonal or block diagonal structure of the matrices. Finally, the application of $\mathbf{Q}^{-1} = (\mathbf{G}^c)^{-1}\mathbf{Y}(\mathbf{G}^c)^{-T}$ is accomplished by the following three-step procedure (analogous to that of $\mathbf{X}^{-1}$): 1) $(\mathbf{G}^c)^{-T}$, which corresponds to a backward solve of a linearized K-step time marching problem; 2) $\mathbf{Y}$, which requires application of $\mathbf{A}^{\text{space}}_h$ onto K spatial vectors; and 3) $(\mathbf{G}^c)^{-1}$, which corresponds to a forward solve of a linearized K-step time marching problem. Thus, each Lanczos step of the inf-sup eigenproblem requires $\mathcal{O}(K)$ operations.

The calculation of the correction factors require the extreme eigenvalues of a generalized eigenproblem $\mathbf{P}\mathbf{v} = \lambda\mathbf{Q}\mathbf{v}$ with

$$\mathbf{P} \equiv \frac{1}{2}\left((\mathbf{G}^c)^T\mathbf{Y}^{-1}\mathbf{G}^k + (\mathbf{G}^k)^T\mathbf{Y}^{-1}\mathbf{G}^c\right)$$
$$\mathbf{Q} \equiv (\mathbf{G}^c)^T\mathbf{Y}^{-1}\mathbf{G}^c.$$

Here, $\mathbf{G}^k \in \mathbb{R}^{(K\cdot\mathcal{N})\times(K\cdot\mathcal{N})}$ is the Jacobian matrix corresponding to the residual operator linearized about the solution at the interpolation point $u_{\delta,k}$. Application of $\mathbf{P}$ again requires $\mathcal{O}(K)$ operations due to the block bidiagonal and block diagonal structure of $\mathbf{G}^c$ and $\mathbf{Y}$, respectively. The $\mathbf{Q}$ matrix is identical to that used for the inf-sup constant calculation; thus, application of $\mathbf{Q}$ and $\mathbf{Q}^{-1}$ can be carried out in $\mathcal{O}(K)$ operations.

3.2.3. *Sobolev Embedding Constant*

The final piece required for the BRR theory is the L^4-$\mathcal{X}_\delta$ Sobolev embedding constant. Details of the approximation of the embedding constant are provided in Appendix A; here we state the main results.

Due to the nonlinearity, we are not able to analyze the L^4-$\mathcal{X}_\delta$ embedding problem analytically. However, we can analyze closely related linear problems: L^2-$\mathcal{X}$ embedding and L^2-$\mathcal{X}_\delta$ embedding. Using the Fourier decomposition in space and time, we can show that the L^2-$\mathcal{X}$ embedding constant is bounded by

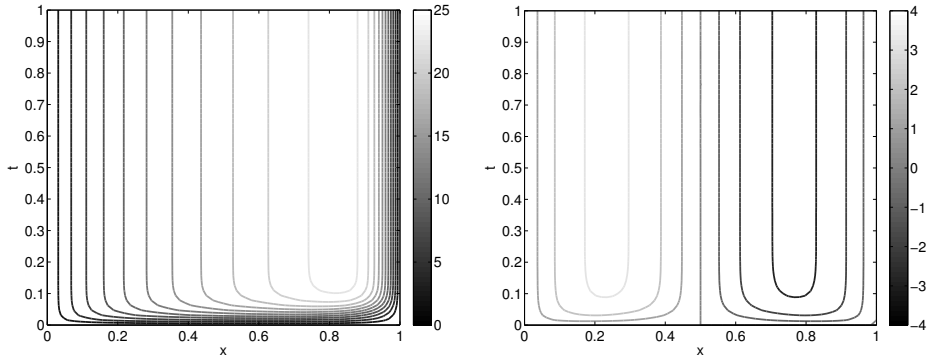
$$\theta \equiv \sup_{w\in\mathcal{X}} \frac{\|w\|_{L^2(I;L^2(\Omega))}}{\|w\|_{\mathcal{X}}} \leq \left(\frac{1}{4T^2} + \pi^2\right)^{-1/2}$$

for $\Omega = (0,1)$ and $I = (0,T]$ with $T > \sqrt{5}/(4\pi)$.

For a uniform temporal discretization, the asymptotic behavior of the L^2-$\mathcal{X}_\delta$ embedding constant in the limit of $\Delta t \to 0$ (for a fixed T) can also be analyzed; the constant approaches a constant for T sufficiently small and grows weakly with the final time for T sufficiently large. Specifically, as $\Delta t \to 0$ ($K \to \infty$) for a fixed T,

$$\theta_\delta \equiv \sup_{w\in\mathcal{X}_\delta} \frac{\|w\|_{L^2(I;L^2(\Omega))}}{\|w\|_{\mathcal{X}_\delta}} \sim \begin{cases} (\frac{1}{4T^2} + \pi^2)^{-1/2}, & T \leq T_{\text{thresh}} \\ C\sqrt{T}, & T > T_{\text{thresh}} \end{cases}$$

for some threshold time T_{thresh} and a constant C independent of T. The T-dependent behavior of the embedding constant θ_δ is due to the mesh-dependent norm $\|\cdot\|_{\mathcal{X}_\delta}$, which includes the mesh-dependent averaged term $\|\bar{\cdot}\|_{L^2(I;V)}$. For $T \leq T_{\text{thresh}}$, the supremizer of θ_δ is the lowest frequency mode in time, which approximates the supremizer of the continuous embedding constant θ; for $T > T_{\text{thresh}}$, the supremizer of θ_δ is the highest frequency mode in time, and hence the behavior of the supremum θ_δ is not predicted by the continuous counterpart θ. For an arbitrary temporal discretization, we are unable to analytically analyze the L^2-$\mathcal{X}_\delta$ embedding constant; however, numerical experiments suggest that, for $T = 1$, the constant is bounded by $\theta_\delta \leq 0.41$ on any quasi-uniform temporal discretization.



(a) Case 1: $g = 1$ (b) Case 2: $g = \frac{1}{2} - x$

Fig. 1. The solution to the Burgers problem Case 1 and Case 2 for $\mathrm{Pe} = 20$.

The L^4-$\mathcal{X}_\delta$ embedding constant, ρ_δ, can be approximated using the fixed-point iteration algorithm of Deparis[3] (see also Manzoni[10] for the analysis of the algorithm). Numerical experiments suggest that the embedding constant depends rather strongly on the temporal grading of the space-time mesh due to the mesh dependence of the $\mathcal{X}_\delta$ norm; thus, we compute the L^4-$\mathcal{X}_\delta$ embedding constant on a particular space-time mesh and use that constant to construct the BRR error bound.

We conclude the section with two remarks. We first emphasize that the mesh dependence originates in the mesh-dependent norm and does not reflect any fundamental ill-posedness. We also note that the origin of the mesh dependence can be traced back to the loss of L-stability for the Crank-Nicolson scheme; a more stable temporal discretization removes the anomaly.[17]

4. Numerical Results

4.1. *Model Problems*

We consider two different forcing functions in this section. First is a constant function, $g_1 = 1$, which results in $\mathcal{F}_1(v;\mu) = \mu \cdot \int_I \int_\Omega v dx dt$ with $\mu = \mathrm{Pe}^2$. The solution over the space-time domain for the $\mathrm{Pe} = 20$ case is shown in Figure 1(a). As the Peclet number increases, the boundary layer at $x = 1$ gets thinner and the initial transition time decreases. The second case uses a spatially linear source function, $g_2 = \frac{1}{2} - x$, which results in $\mathcal{F}_2(v;\mu) = \mu \cdot \int_I \int_\Omega (\frac{1}{2} - x) v dx dt$. The solution for this second case with $\mathrm{Pe} = 20$ is shown in Figure 1(b). This case develops an internal layer at $x = 1/2$, which becomes thinner as the Peclet number increases. These two cases exhibit different stability properties, as we will show shortly.

For purposes of comparison, we provide here a short summary of the time-marching $L^2(\Omega)$ error bound developed by Nguyen *et al.*[11] A parameter that dictates the effectivity of the time-marching $L^2(\Omega)$ formulation is the stability parameter ω^k,

defined as[a]

$$\omega^k \equiv \inf_{v \in V_h} \frac{4b(v, u(\mu), v) + a(v, v)}{\|v\|_{L^2(\Omega)}}, \quad k = 1, \ldots, K.$$

In particular, a negative value of ω^k implies that the $L^2(\Omega)$ error estimate grows exponentially over that period of time. All results shown in this section use the exact value of ω^k instead of a lower bound obtained using the successive constraint method (SCM) as done in Nguyen *et al.*[11]; i.e. we use the most favorable stability constant for the $L^2(\Omega)$ time-marching formulation.

Finally, we provide details of the space-time meshes employed. For all cases considered, the spatial mesh consisting of 128 uniformly spaced elements of size $h = 1/128$. On the other hand, the temporal mesh is varied as a function of the Peclet number to effectively resolve the initial transient. Namely, for $\mathrm{Pe} \leq 6$, we employ a temporal mesh consisting of 128 uniformly spaced elements of size $\Delta t = 1/128$. For $\mathrm{Pe} > 6$, we divide the temporal mesh into two regions: the initial transient region, $t \in (0, 3/\mathrm{Pe}]$, discretized by 64 uniformly spaced elements of size $\Delta t = (3/\mathrm{Pe})/64$; the remaining region, $t \in (3/\mathrm{Pe}, 1]$, discretized by 64 uniformly spaced elements of size $\Delta t = (1 - 3/\mathrm{Pe})/64$. For each of N_μ temporal meshes, $\mathcal{T}^{\mathrm{time}}_{\Delta t, j}$, $j = 1, \ldots, N_\mu$, we use the Peclet number at the centroid of the associated parameter subdomain $\mathcal{D}_j$ as the reference Peclet number over the domain.

4.2. *Stability: Small Parameter Intervals*

We will first demonstrate the improved stability of the space-time *a posteriori* error estimate compared to the $L^2(\Omega)$ time-marching error estimate. For the space-time formulation, we monitor the variation in the inf-sup constant, β^p_δ, and the effectivity, $\Delta^p_\delta / \|e\|_{\mathcal{X}_\delta}$, with the Peclet number. For the $L^2(\Omega)$ time-marching formulation, we monitor several quantities: the minimum (normalized) stability constant, $\min_k \omega^k/\mathrm{Pe}$; the final stability constant, ω^K/Pe; the maximum effectivity, $\max_k \Delta^k_{L^2} / \|e^k\|_{L^2(\Omega)}$; and the final effectivity, $\Delta^K_{L^2} / \|e^K\|_{L^2(\Omega)}$.

For each case, the reduced basis approximation is obtained using the $p = 2$ interpolation over a short interval of $\mathcal{D} = [\mathrm{Pe} - 0.1, \mathrm{Pe} + 0.1]$. Note that, the use of the short interval implies that $\tau^p_\delta \ll 1$, which reduces the BRR-based error bound to

$$\Delta^p_\delta(\mu) \approx \frac{1}{\beta^p_{\delta,\mathrm{LB}}(\mu)} \epsilon^p_\delta.$$

In addition, as the supremizer evaluated at the centroid of the interval is close to the true supremizer over a short interval, $\beta^p_{\delta,\mathrm{LB}}(\mu) \approx \beta^p_\delta(\mu)$, $\forall \mu \in \mathcal{D}$. In other words, we consider the short intervals to ensure a good inf-sup lower bound such that we can focus on stability independent of the quality of the inf-sup *lower bound*; we

[a]In the original paper by Nguyen *et al.*, the variable ρ^k is used for the stability constant. Here, we use ω^k to avoid confusion with the L^4-$\mathcal{X}_\delta$ embedding constant for the space-time formulation.

	space-time		$L^2(\Omega)$ time-marching			
Pe	β_δ^p	$\frac{\Delta_\delta^p}{\|e\|_{\mathcal{X}_\delta}}$	$\min_k \frac{\omega^k}{\mathrm{Pe}}$	$\frac{\omega^K}{\mathrm{Pe}}$	$\max_k \frac{\Delta_{L^2}^k}{\|e^k\|}$	$\frac{\Delta_{L^2}^K}{\|e^K\|}$
1	0.993	1.14	9.87	9.87	3.87	1.30
10	0.665	2.22	0.982	1.32	3.18	2.11
50	0.303	7.00	0.114	0.924	7.73	5.10
100	0.213	9.75	0.0203	0.862	11.7	6.95
200	0.149	12.4	-0.0072	0.820	18.0	9.59

Table 1. Summary the inf-sup constant and effectivity for the space-time formulation and the stability constant and effectivity for the $L^2(\Omega)$ time-marching formulation for Case 1 with $g = 1$.

will later assess the effectiveness of the lower bound. The effectivity reported is the worst case value observed on 40 sampling points over the interval.[b]

Table 1 shows the variation in the stability constant and the effectivity for Case 1 for Pe = 1, 10, 50, 100, and 200. The stability constant for the space-time formulation gradually decreases with Pe; accordingly, the effectivity worsens from 1.14 for Pe = 1 to 12.4 for Pe = 200. Note that the effectivity of $\mathcal{O}(10)$ is more than adequate for the purpose of reduced-order approximation as the error typically rapidly converges (i.e. exponentially) with the number of reduced bases. The $L^2(\Omega)$ time-marching formulation also performs well for this case. This is because, even for the Pe = 200 case, the stability constant ω^k/Pe takes on a negative value over a very short time interval and is asymptotically stable. (See Nguyen *et al.*[11] for the detailed behavior of the stability constant over time.)

Table 2 shows the variation in the stability constant and the effectivity for Case 2 for Pe = 1, 10, 20, 50, and 100. Note that the asymptotic stability constant for the $L^2(\Omega)$ time-marching formulation is negative for $\mathrm{Pe} \gtrsim 18.9$; consequently, the error bound grows exponentially with time even for a moderate value of the Peclet number, rendering the error bound meaningless. The stability constant for the space-time formulation is much better behaved. The effectivity of 41.2 at Pe = 50 is a significant improvement over the 10^{28} for the $L^2(\Omega)$ time-marching formulation, and the error estimate remains meaningful even for the Pe = 100 case.

4.3. N_μ-p Interpolation over a Wide Range of Parameters

Now we demonstrate that our certified reduced basis method provides accurate and certified solutions over a wide range of parameters using a reasonable number of snapshots. Here, we employ a simple (and rather crude) N_μ-p adaptive procedure to construct certified reduced basis approximations over the entire $\mathcal{D}$ with an error bound of $\Delta_{\mathrm{tol}} = 0.01$. Our N_μ-p approximation space is described in terms of a

[b]The 40 sampling points are equally-spaced between $[\mathrm{Pe} - 0.099, \mathrm{Pe} + 0.099]$. We have found that the variation in the effectivity across the sampling points is small (less than 10%) over the small intervals considered.

	space-time		$L^2(\Omega)$ time-marching			
Pe	β_δ^p	$\frac{\Delta_\delta^p}{\|e\|_{\mathcal{X}_\delta}}$	$\min_k \frac{\omega^k}{\mathrm{Pe}}$	$\frac{\omega^K}{\mathrm{Pe}}$	$\max_k \frac{\Delta_{L^2}^k}{\|e^k\|}$	$\frac{\Delta_{L^2}^K}{\|e^K\|}$
1	0.999	1.01	9.84	9.84	2.80	2.80
10	0.877	1.15	0.727	0.727	3.12	3.12
20	0.547	1.84	-0.0675	-0.0675	12.4	12.4
30	0.213	4.99	-0.606	-0.606	3.7×10^4	3.7×10^4
50	0.038	41.2	-1.67	-1.67	6.5×10^{28}	6.5×10^{28}
100	0.0077	237	-4.43	-4.43	–	–

Table 2. Summary the inf-sup constant and effectivity for the space-time formulation and the stability constant and effectivity for the $L^2(\Omega)$ time-marching formulation for Case 2 with $g = \frac{1}{2} - x$.

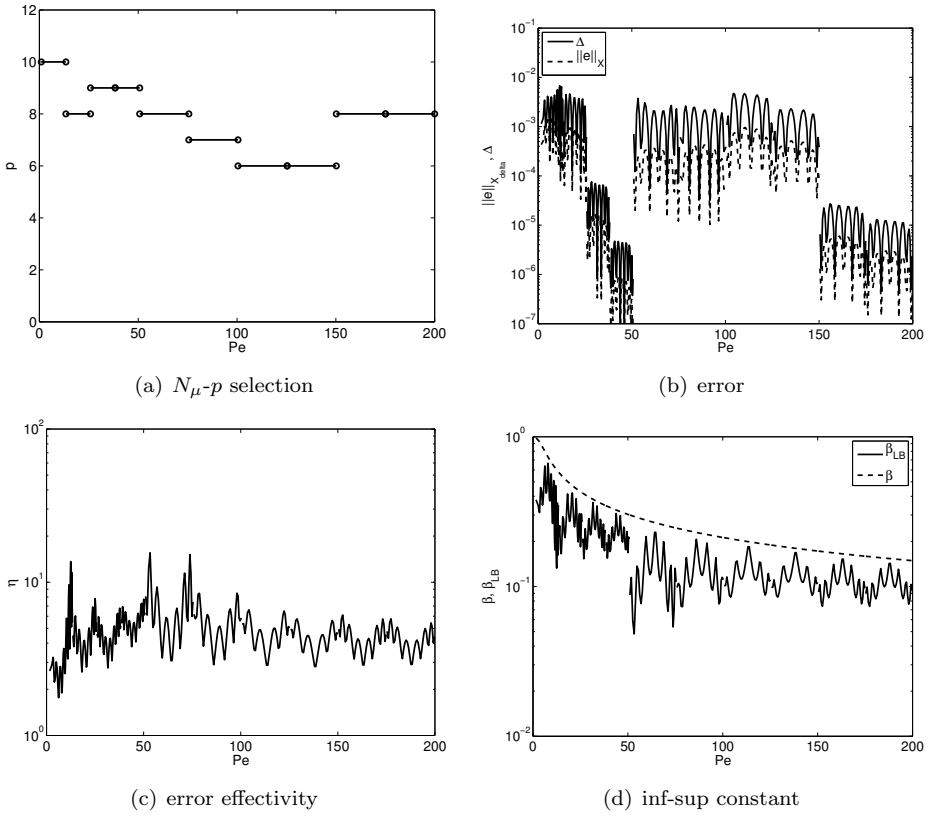
set $\mathrm{Pe}^{\mathrm{set}}$ consisting of $N_\mu + 1$ points that delineate the endpoints of the parameter intervals and an N_μ-tuple $P^{\mathrm{set}} = (p_1, \ldots, p_{N_\mu})$ specifying the polynomial degree over each interval. Starting from a single $p = 1$ interval over the entire $\mathcal{D}$, we recursively apply one of the following two operations to each interval $[\mathrm{Pe}_{\mathrm{L}}, \mathrm{Pe}_{\mathrm{U}}] = [\mathrm{Pe}_j^{\mathrm{set}}, \mathrm{Pe}_{j+1}^{\mathrm{set}}]$ with polynomial degree p_j:

(a) if $\min_\mu \beta_{\delta,\mathrm{LB}}^p(\mu) \leq 0$, subdivide $[\mathrm{Pe}_{\mathrm{L}}, \mathrm{Pe}_{\mathrm{U}}]$ into $[\mathrm{Pe}_{\mathrm{L}}, \mathrm{Pe}_{\mathrm{M}}] \cup [\mathrm{Pe}_{\mathrm{M}}, \mathrm{Pe}_{\mathrm{U}}]$ where $\mathrm{Pe}_{\mathrm{M}} = (\mathrm{Pe}_{\mathrm{L}} + \mathrm{Pe}_{\mathrm{U}})/2$, assign p_j to both intervals, and update $\mathrm{Pe}^{\mathrm{set}}$ and P^{set}.

(b) if $\min_\mu \beta_{\delta,\mathrm{LB}}^p(\mu) > 0$ but $\max_\mu \tau_\delta^p(\mu) \geq 1$ or $\max_\mu \Delta_\delta^p(\mu) \geq \Delta_{\mathrm{tol}}$, then increase p_j to $p_j + 1$.

The operation (a) decreases the width of the parameter interval, which increases the effectiveness of the supremizer S_j^c over $\mathcal{D}_j$ and improves the inf-sup lower bound. The operation (b) aims to decrease the residual (and hence the error) by using a higher-order interpolation, i.e. p-refinement. Thus, in our adaptive procedure, the N_μ and p refinement serves two distinct purposes: improving the stability estimate and improving the approximability of the space, respectively. In particular, we assume that the solution dependence on the parameter is smooth and use (only) p-refinement to improve the approximability; this is in contrast to typical hp adaptation where both h- and p-refinement strategies are used to improve the approximability for potentially irregular functions.

The result of applying the N_μ-p adaptive procedure to Case 1 is summarized in Figure 2. Here, we show variations over the parameter domain $\mathcal{D} = [1, 200]$ of key quantities: a) the approximation polynomial degree; b) the error and error bound; c) the error effectivity; and d) the inf-sup constant and associated lower bound. First, Figure 2(a) shows that the entire parameter domain is covered using just 10 intervals consisting of 89 total interpolation points; this is despite the use of the crude adaptation process whose inefficiency is reflected in excessively accurate estimates in some of the intervals, as shown in Figure 2(b). Second, we note that



(a) N_μ-p selection (b) error

(c) error effectivity (d) inf-sup constant

Fig. 2. The error, effectivity, and inf-sup constant behaviors on the final N_μ-p adapted interpolation for Case 1.

the maximum error bound of 10^{-2} is clearly satisfied over the entire parameter range. Third, Figure 2(c) shows that the effectivity is of order 5. Finally, we observe in Figure 2(d) that the inf-sup lower bound procedure provides relatively sharp lower bounds thanks to the adaptive N_μ-p interpolation strategy that considers the behavior of the stability estimate.

Table 3 shows the p-convergence behavior of our certified basis formulation over the final interval, $\mathcal{D}_{10} = [175.13, 200.00]$.[c] Each variable is sampled at 40 equispaced sampling points over $\mathcal{D}_{10}$ and the worst case values are reported. The table confirms that the error (and the normalized residual) converges rapidly with p. The rapid convergence suggests that the error effectivity of $\mathcal{O}(10)$ is more than adequate, as improving the error by a factor of 10 only requires 1 or 2 additional points.

[c]Using the N_μ-p adaptive procedure, this $p = 8$, $\mathcal{D}_{10} = [175, 13, 200.00]$ interval is created by subdividing a $p = 8$, $\mathcal{D}_9 = [150.25, 200.00]$ interval in the final step. This results in the use of the $p = 8$ interpolant over the interval $\mathcal{D}_{10}$ in the final N_μ-p adapted configuration despite the error meeting the specified tolerance for $p = 5$.

p	$\max_\mu \tau_\delta^p(\mu)$	$\max_\mu \Delta_\delta^p(\mu)$	$\max_\mu \|e(\mu)\|_{\mathcal{X}_\delta}$	$\max_\mu \frac{\Delta_\delta^p(\mu)}{\|e(\mu)\|_{\mathcal{X}_\delta}}$	$\min_\mu \frac{\beta_{\delta,\mathrm{LB}}^p(\mu)}{\beta_\delta^p(\mu)}$
1	1.82×10^4	-	1.14×10^1	-	0.61
2	3.57×10^2	-	6.67×10^{-1}	-	0.62
3	3.03×10^1	-	9.37×10^{-2}	-	0.61
4	2.06×10^0	-	1.11×10^{-2}	-	0.61
5	2.53×10^{-1}	6.63×10^{-3}	1.48×10^{-3}	5.14	0.56
6	3.24×10^{-2}	7.71×10^{-4}	1.86×10^{-4}	5.06	0.52
7	4.39×10^{-3}	1.02×10^{-4}	2.38×10^{-5}	5.38	0.49
8	6.18×10^{-4}	1.30×10^{-5}	3.00×10^{-6}	5.77	0.47

Table 3. The p-convergence behavior over the final interval of Case 1, Pe $\in [175.13, 200.00]$.

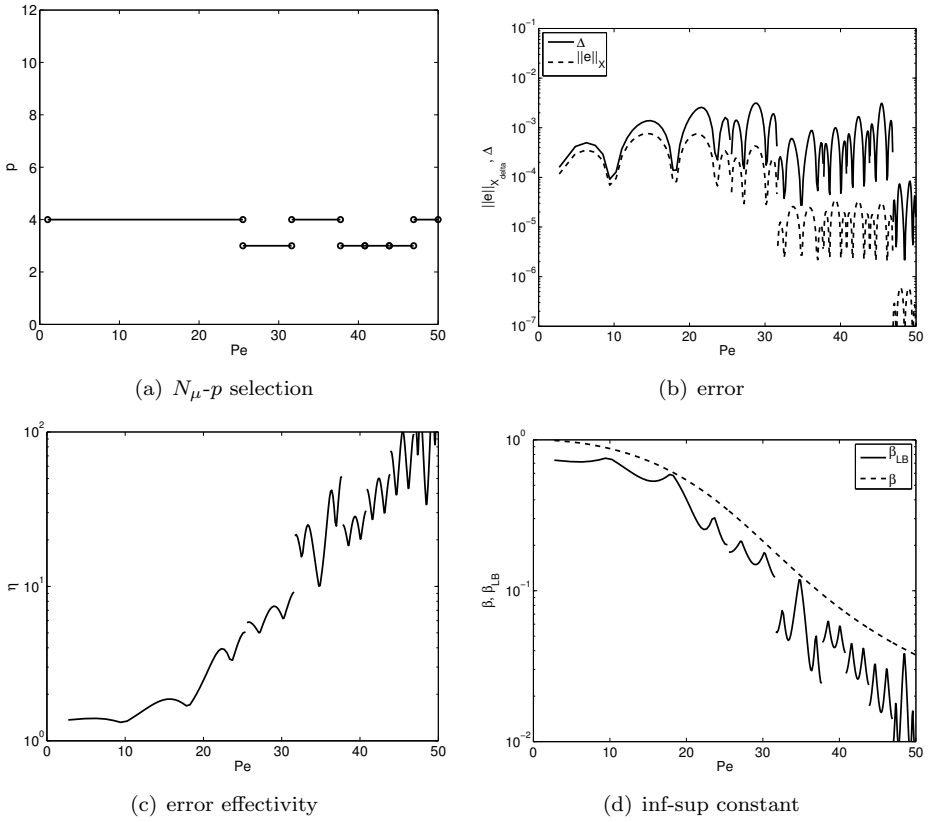
The higher p not only provides higher accuracy but also concomitantly enables construction of the BRR-based error bounds by decreasing τ_δ^p. Note also that the inf-sup effectivity decreases with p in general as a larger number of "inf" operations are required to construct $\beta_{\delta,\mathrm{LB}}^p$ using the procedure introduced in Section 3.2.2.

Figure 3 summarizes the N_μ-p interpolation strategy, the error behavior, and the stability constant variation for Case 2 over $\mathcal{D} = [1, 50]$. We recall from Section 4.2 that this problem is less stable than Case 1; the classical formulation produces exponentially growing error bounds. First, Figure 3(a) shows that the N_μ-p adaptive procedure utilizes 7 intervals consisting of 31 total interpolation points. Second, Figure 3(b) verifies that the maximum error bound incurred over $\mathcal{D}$ is less than 10^{-2}; we note that the error (and error bound) is smaller for higher Peclet number because the proximity condition of the BRR theory ($\tau_\delta^p(\mu) < 1$) forces a smaller residual than actually necessary to meet the error tolerance. Third, Figure 3(c) shows that, due to the unstable nature of the problem, the effectivity worsens as the Peclet number increases; nevertheless, unlike in the classical time-marching based formulation, our error bounds remain meaningful over the entire parameter range. Finally, we note that the size of the interval in the high Peclet number regime is dictated by the necessity to maintain a positive inf-sup lower bound; for instance, for the $p = 4$ interpolation, we were unable to maintain a positive value of $\beta_{\delta,\mathrm{LB}}^p$ over a single interval of $[46, 50]$, necessitating the split into two smaller intervals.

Table 4 shows the p-convergence behavior of the reduced basis formulation over $\mathcal{D}_8 = [46.94, 50]$. Similar to the previous case, the normalized residual, the error bound, and the error converge exponentially with p. The worst effectivity over the 40 sampling points is of $\mathcal{O}(100)$.

Appendix A. Sobolev Embedding Constants

In this appendix, we study the behavior of the L^4-$\mathcal{X}_\delta$ embedding constant required for the Brezzi-Rappaz-Raviart theory. Unfortunately, due to the nonlinearity, we are not able to analyze the L^4-$\mathcal{X}_\delta$ problem analytically. To gain some insight into the behavior of the embedding constant using analytical techniques, let us consider



(a) N_μ-p selection (b) error

(c) error effectivity (d) inf-sup constant

Fig. 3. The error, effectivity, and inf-sup constant behaviors on the final N_μ-p adapted interpolation for Case 2.

p	$\max_\mu \tau_\delta^p(\mu)$	$\max_\mu \Delta_\delta^p(\mu)$	$\max_\mu \|e(\mu)\|_{\mathcal{X}_\delta}$	$\max_\mu \frac{\Delta_\delta^p(\mu)}{\|e(\mu)\|_{\mathcal{X}_\delta}}$	$\min_\mu \frac{\beta_{\delta,\mathrm{LB}}^p(\mu)}{\beta_\delta^p(\mu)}$
1	5.25×10^3	-	7.43×10^{-2}	-	0.21
2	2.04×10^1	-	1.03×10^{-3}	-	0.55
3	1.80×10^0	-	2.79×10^{-5}	-	0.33
4	3.86×10^{-2}	8.36×10^{-5}	6.05×10^{-7}	91.30	0.35
5	3.15×10^{-2}	1.13×10^{-5}	1.54×10^{-8}	137.85	0.24

Table 4. The p-convergence behavior over the last interval of Case 2, $\mathrm{Pe} \in [46.94, 50.00]$.

two closely related linear problems, L^2-$\mathcal{X}$ embedding and L^2-$\mathcal{X}_\delta$ embedding, in Appendix A.1 and A.2. Then, we will numerical investigate the behavior of the

L^4-$\mathcal{X}_\delta$ embedding constant in Appendix A.3.[d]

A.1. L^2-$\mathcal{X}$ *Embedding*

Let us first consider L^2-$\mathcal{X}$ embedding. The embedding constant is defined by

$$\theta \equiv \sup_{w\in\mathcal{X}} \frac{\|w\|_{L^2(I;L^2(\Omega))}}{\|w\|_{\mathcal{X}}} \ ,$$

which is related to the minimum eigenvalue of an eigenproblem: find $(w,\lambda)\in\mathcal{X}\times\mathbb{R}$ such that $\|w\|^2_{L^2(I;L^2(\Omega))}=1$ and

$$(w,v)_{\mathcal{X}} = \lambda(w,v)_{L^2(I;L^2(\Omega))}, \quad \forall v\in\mathcal{X} \ ;$$

the embedding constant is given by $\theta=\lambda_{\min}^{-1/2}$. The application of the Fourier decomposition in the spatial domain[e] results in an eigenproblem in time: find eigenpairs $(w^{k_x},\lambda^{k_x})\in H^1_{(0)}(I)\times\mathbb{R}$ such that

$$\frac{1}{k_x^2\pi^2}\int_I \dot{v}^{k_x}(t)\dot{w}^{k_x}(t)dt + k_x^2\pi^2\int_I v^{k_x}(t)w^{k_x}(t)dt + v^{k_x}(T)w^{k_x}(T) = \lambda^{k_x}\int_I v^{k_x}(t)w^{k_x}(t)dt, \quad \forall v^{k_x}\in H^1_{(0)}(I),$$

where $v^{k_x}\in H^1_{(0)}(I)$ is the temporally-varying Fourier coefficient associated with the k_x-mode and $H^1_{(0)}(I)\equiv\{v\in H^1(I): v(t=0)=0\}$. Note that the homogeneous Dirichlet condition is enforced at $t=0$ and a Robin condition, $\dot{w}/(k_x^2\pi^2)+w=0$, is enforced at $t=T$. The eigenmodes of the continuous problem are given by

$$v^{k_x,k_t}(t)=\sin(m_{k_x,k_t}t), \quad k_t=1,2,\ldots,$$

where the wave number m_{k_x,k_t} satisfies

$$\tan(m_{k_x,k_t}T) = -\frac{m_{k_x,k_t}}{k_x^2\pi^2} \ , \tag{A.1}$$

and the associated eigenvalues are given by

$$\lambda^{k_x,k_t} = k_x^2\pi^2 + \frac{m^2_{k_x,k_t}}{k_x^2\pi^2} \ . \tag{A.2}$$

Without loss of generality, we order the wave number such that, for each k_x, $m_{k_x,k_t=1} < m_{k_x,k_t=2} < \cdots$.

We now deduce the minimum eigenvalue (and hence the embedding constant). A close inspection of the root condition (A.1) shows that, for any given k_x, the minimum wave number, $m_{k_x,k_t=1}$, lies in the interval $(\pi/(2T),\pi/T)$. It follows that,

[d]Analysis in this appendix is "formal"; for brevity, some of the assumptions or arguments required related to completeness or compactness may be omitted.

[e]We could directly analyze the spatial discretization with appropriate modification of the k_x Fourier symbol per the usual von Neumann analysis. Here we consider a continuous-in-space case for simplicity.

for $T > \sqrt{5}/(4\pi)$, the minimum eigenvalue is given for $k_x = 1$, and the eigenvalue lies in $\lambda_{\min} = \lambda^{k_x=1,k_t=1} \in (\pi^2 + 1/(4T^2), \pi^2 + 1/T^2)$. (Appropriate bounding constant may be deduced from the expression for the continuous problem even for $T \leq \sqrt{5}/(4\pi)$.) Thus, for $T > \sqrt{5}/(4\pi)$, the L^2-$\mathcal{X}$ embedding constant lies in

$$\left(\pi^2 + \frac{1}{T^2}\right)^{-1/2} < \theta < \left(\pi^2 + \frac{1}{4T^2}\right)^{-1/2} .$$

Note that these bound for the L^2-$\mathcal{X}$ embedding constant is not significantly different from the standard L^2-H^1_0 embedding constant, $\theta_{L^2\text{-}H^1_0} = \left(\pi^2/T^2 + \pi^2\right)^{-1/2}$. For instance, for $T = 1$, the embedding constant lies in $\theta \in (0.3033, 0.3173)$; the direct computation of the root condition (A.1) and the eigenvalue (A.2) produces the wave number of $m_{k_x=1,k_t=1} = 2.8596$ and the embedding constant of $\theta = 0.3057$.

We finally observe that the eigenvalues of the space-time eigenproblem (A.2) is a sum of two terms, one of which increases with the Fourier coefficient k_x and the other which decreases with k_x. This is precisely the subtlety introduced by the space-time norms. In the case of L^2-$\mathcal{X}$ embedding, this subtlety turns out not to be an issue, bur rather just a complication; however, this subtlety is an issue in proving the existence and boundedness of the L^4-$\mathcal{X}$ embedding constant in higher dimensions.

A.2. L^2-$\mathcal{X}_\delta$ Embedding

Now let us consider L^2-$\mathcal{X}_\delta$ embedding. The embedding constant is defined by

$$\theta_\delta \equiv \sup_{w \in \mathcal{X}_\delta} \frac{\|w\|_{L^2(I;L^2(\Omega))}}{\|w\|_{\mathcal{X}_\delta}},$$

where we recall that $\|w\|^2_{\mathcal{X}_\delta} = \|\dot{w}\|^2_{L^2(I;V')} + \|\bar{w}\|^2_{L^2(I;V)} + \|w(T)\|^2_H$. Similar to the L^2-$\mathcal{X}$ embedding problem, the solution is given by the eigenproblem: find $(w, \lambda) \in \mathcal{X}_\delta \times \mathbb{R}$ such that $\|w\|^2_{L^2(I;L^2(\Omega))} = 1$ and

$$(w, v)_{\mathcal{X}_\delta} = \lambda (w, v)_{L^2(I;L^2(\Omega))}, \quad \forall v \in \mathcal{X}_\delta \ ;$$

the embedding constant is given by $\theta_\delta = \lambda^{-1/2}_{\min}$. However, as the $\mathcal{X}_\delta$ norm is dependent on the temporal mesh by construction, we must consider temporally discrete spaces for our analysis. Let $V_{\Delta t} \subset H^1_0(I)$ be the piecewise linear temporal approximation space. Then, the Fourier decomposition in the spatial domain results in an eigenproblem: find eigenpairs $(w^{k_x}_\delta, \lambda^{k_x}) \in V_{\Delta t} \times \mathbb{R}$ such that

$$\frac{1}{k_x^2\pi^2} \int_I \dot{v}^{k_x}(t)\dot{w}^{k_x}(t)dt + k_x^2\pi^2 \int_I \bar{v}^{k_x}(t)\bar{w}^{k_x}(t)dt + v^{k_x}(T)w^{k_x}(T) = \lambda^{k_x} \int_I v^{k_x}(t)w^{k_x}(t)dt, \quad \forall v^{k_x} \in V_{\Delta t},$$

where $\bar{v}^{k_x}$ over the I^k is given by $(\Delta t^k)^{-1} \int_{I^k} v^{k_x} dx$.

We first analyze the embedding constant for $V_{\Delta t}$ with a constant time step, $\Delta t = \Delta t^1 = \cdots = \Delta t^K$. The k-th entry of the k_t-th eigenmode $\mathbf{v}^{k_x,k_t} \in \mathbb{R}^K$ is given by

$$(\mathbf{v}^{k_x,k_t})_k = \sin(m_{k_x,k_t} t) \ ,$$

where the wave number m_{k_x,k_t} is the k_t-th root of

$$\sin(m_{k_x,k_t} T) + \frac{\cos(m_{k_x,k_t} T)\sin(m_{k_x,k_t}\Delta t)}{2+\cos(m_{k_x,k_t}\Delta t)} \left(\frac{3}{k_x^2\pi^2\Delta t} - \frac{k_x^2\pi^2\Delta t}{4} \right) = 0 \qquad \text{(A.3)}$$

and the associated eigenvalue is given by

$$\lambda^{k_x,k_t} = \frac{\frac{2}{k_x^2\pi^2\Delta t}(1-\cos(m_{k_x,k_t}\Delta t)) + \frac{k_x^2\pi^2}{2}\Delta t(1+\cos(m_{k_x,k_t}\Delta t))}{\frac{1}{3}\Delta t(2+\cos(m_{k_x,k_t}\Delta t))} \ ; \qquad \text{(A.4)}$$

we have found the wave numbers associated with the continuous L^2-$\mathcal{X}$ embedding problem serves as good initializers for the root finding problem (A.3). Numerically, we have observed that the eigenvalue is minimized for $k_t = 1$ for T less than some threshold T_{thresh} and for $k_t = K$ for $T > T_{\text{thresh}}$. We now separately analyze these two branches of the solution.

The behavior of the $k_t = 1$ branch is similar to that of the continuous L^2-$\mathcal{X}$ embedding problem. In particular in the limit of $K \to \infty$ (and $\Delta t \to 0$), the discrete root condition (A.3) becomes the continuous root condition (A.1); similarly, the discrete eigenvalue (A.4) approaches the continuous eigenvalue (A.2).

The behavior of the $k_t = K$ branch is dissimilar to that of the continuous problem. In the limit of $K \to \infty$, this branch is approximated by $m_{k_x,k_t}\Delta t = \pi - n_{k_x,k_t}\Delta t$ for $n_{k_x,k_t}\Delta t \to 0$. For $K \to \infty$, the discrete root condition (A.3) simplifies to

$$\tan(n_{k_x,k_t} T) = \frac{3 n_{k_x,k_t}}{k_x^2\pi^2}$$

and the discrete eigenvalue expression (A.4) also simplifies to

$$\lambda^{k_x,k_t} = 3\left(\frac{4}{k_x^2\pi^2\Delta t^2} + \frac{k_x^2\pi^2\Delta t^2 n_{k_x,k_t}^2}{4} \right) \ .$$

We identify, for a given n_{k_x,k_t}, the minimizing wave number k_x^* based on a continuous relaxation as $k_x^* = 2/(\pi\Delta t\sqrt{n_{k_x,k_t}})$. The substitution of the minimizing wave number k_x^* in the expression for the eigenvalue yields $\lambda^{k_x^*,k_t} = 6 n_{k_x^*,k_t}$. For $T > 3/\pi^2$, the smallest positive root of (A.3) lies in the interval $n_{\min} \in [\frac{\pi}{T}, \frac{3}{2}\frac{\pi}{T}]$. Thus, it follows that, as $K \to \infty$, the minimum eigenvalue asymptotically behaves as $\lambda_{\min} \sim 1/T$. In other words, the L^2-$\mathcal{X}_\delta$ embedding constant asymptotically behaves as $\theta_\delta \sim \sqrt{T}$ for $T > T_{\text{thresh}}$. Thus, for a sufficiently large T, the embedding constant scales weakly with the final time T.

Unfortunately, for $V_{\Delta t}$ with non-constant time stepping, we cannot deduce the embedding constant analytically. Here, we numerically demonstrate that the L^2-$\mathcal{X}_\delta$

	mesh grading factor, q							
K	-2	-1	0	1	2	3	4	5
2	0.3242	0.3336	0.3473	0.3689	0.3817	0.3989	0.4046	0.4076
4	0.3302	0.3237	0.3110	0.3461	0.3875	0.4036	0.4073	0.4080
8	0.3197	0.3099	0.3069	0.3464	0.3882	0.4036	0.4075	0.4081
16	0.3089	0.3067	0.3060	0.3457	0.3882	0.4034	0.4074	0.4081
32	0.3065	0.3060	0.3058	0.3458	0.3882	0.4034	0.4073	0.4081
64	0.3059	0.3058	0.3058	0.3458	0.3882	0.4034	0.4073	0.4078

Table 5. The variation in the L^2-$\mathcal{X}_\delta$ embedding constant with the number of time intervals, K, and the mesh grading factor, q, for $T = 1$.

embedding constant is indeed bounded for all quasi-uniform meshes. In particular, we compute the embedding constant on temporal meshes characterized by the number of elements, K, and a logarithmic mesh grading factor, q, where $q = 0$ corresponds to a uniform mesh, $q > 0$ implies that elements are clustered toward $t = 0$. For q sufficiently large, the first temporal element is of order $\Delta t^1 \approx 10^{-q} T$. Without loss of generality, we pick $T = 1$.

The result of the calculation is summarized in Table 5. The table confirms that, on a uniform temporal mesh ($q = 0$), the embedding constant converges to the semi-analytical value of 0.3057 as K increases. The embedding constant increases with the mesh grading factor, q, due to the presence of the mesh dependent term $\|\cdot\|_{L^2(I;V)}$ in $\|\cdot\|_{\mathcal{X}_\delta}$. In fact, the embedding constant $\sup_{w\in\mathcal{X}_\delta} \|w\|_{L^2(I;L^2(\Omega))}/\|w\|_{\mathcal{X}}$ associated with the norm $\|\cdot\|_{\mathcal{X}}$ — which has $\|\cdot\|_{L^2(I;V)}$ in place of $\|\cdot\|_{L^2(I;V)}$ — is in fact independent of mesh grading and is bounded from above for any mesh by the L^2-$\mathcal{X}$ embedding constant θ; for instance, 0.3057 for $T = 1$. In any event, Table 5 suggests that the L^2-$\mathcal{X}_\delta$ embedding constant asymptote to ≈ 0.41 as $q \to \infty$ and the constant is bounded for all quasi-uniform meshes.

A.3. L^4-$\mathcal{X}_\delta$ Embedding

Recall that the L^4-$\mathcal{X}_\delta$ embedding constant is defined as

$$\rho_\delta \equiv \sup_{w\in\mathcal{X}_\delta} \frac{\|w\|_{L^4(I;L^4(\Omega))}}{\|w\|_{\mathcal{X}_\delta}}.$$

To solve the maximization problem, we employ the fixed-point iteration algorithm of Deparis[3] in the space-time setting.[f] We first define an operator $z : \mathcal{X} \setminus 0 \to L^2(I; L^2(\Omega))$,

$$z(w) = \frac{1}{\|w\|^2_{L^4(I;L^4)}} w^2 \ .$$

[f]We have found Deparis' algorithm to be more robust than the Newton continuation algorithm of Veroy and Patera.[16]

	mesh grading factor, q						
K	-2	-1	0	1	2	3	4
2	0.4147	0.4306	0.4638	0.5018	0.5561	0.6064	0.7032
4	0.4333	0.4407	0.4562	0.5618	0.6816	0.8437	1.0345
8	0.4372	0.4363	0.4555	0.6247	0.8072	1.0223	1.2773
16	0.4339	0.4351	0.5542	0.7387	0.9578	1.2115	1.5263
32	0.4321	0.4343	0.4770	0.8756	1.1303	1.4323	1.7746
64	0.4170	0.4166	0.4171	1.0276	1.3341	1.6840	1.8387

Table 6. The variation in the L^4-$\mathcal{X}_\delta$ embedding constant with the number of time intervals, K, and the mesh grading factor, q, for $T = 1$.

We then introduce an eigenproblem: for a fixed u, find $(w, \lambda) \in \mathcal{X} \times \mathbb{R}$ such that $\|w\|_{\mathcal{X}} = 1$ and

$$\int_I \int_\Omega z(u) w v dx dt = \lambda (w, v)_{\mathcal{X}} \ , \quad \forall v \in \mathcal{X} \ ; \tag{A.5}$$

we denote the largest eigenvalue and the associated eigenfunction by $\lambda_{\max}(z(u))$ and $w_{\max}(z(u))$, respectively. Note that the L^4-$\mathcal{X}_\delta$ supremizer, u^*, is the fixed point $u^* = w_{\max}(z(u^*))$ and the embedding constant is $\rho_\delta = \sqrt{\lambda_{\max}(z(u^*))}$. Deparis' fixed point algorithm is given as follows: initialize $u^0 = 1$; for $l \geq 1$, set

$$u^l = w_{\max}(z(u^{l-1})) \quad \text{and} \quad \lambda^l = \lambda_{\max}(z(u^{l-1})) \ .$$

As $l \to \infty$, the fixed-point algorithm locates at least a local supremizer.

Remark Appendix A.1. The construction of the Krylov space for an efficient solution of the eigenproblem (A.5) by the Lanczos algorithm requires the application of X^{-1}. This space-time solve is efficiently carried out using the the block-tridiagonal decomposition and time-marching solves noted in Remark 2.1.

The numerical values of the embedding constant on different meshes is shown in Table 6. Similar to the L^2-$\mathcal{X}_\delta$ embedding constant, the L^4-$\mathcal{X}_\delta$ embedding constant increases with the number of temporal time steps, K, and the mesh grading factor, q. Based on the table alone, the boundedness of the embedding constant for any quasi-uniform mesh is inconclusive. For this reason, we compute the L^4-$\mathcal{X}_\delta$ embedding constant for each space-time finite element mesh and employ the constant in the construction of BRR bounds. (It is worth noting that, the embedding constant $\sup_{w \in \mathcal{X}_\delta} \|w\|_{L^4(I;L^4(\Omega))} / \|w\|_{\mathcal{X}}$ associated with the continuous $\mathcal{X}$-norm appears to converge to ≈ 0.42 independent of the mesh grading factor.)

Appendix B. Comparison of Inf-Sup Lower Bound Construction Procedures

This appendix details the relationship between the inf-sup lower bound constructed using the procedure developed in Section 3.2.2 and the natural-norm Successive

Constraint Method (SCM) method.[7] For convenience, we refer to our method based on the explicit calculation of the lower and upper bounds of the correction factors as "LU" and that based on the Successive Constraint Method as "SCM." Both LU and SCM procedures are based on the decomposition

$$\beta_\delta^p(\mu) \equiv \inf_{w\in\mathcal{X}_\delta}\sup_{v\in\mathcal{Y}_\delta}\frac{\partial\mathcal{G}(w,\tilde{u}_\delta^p,v)}{\|w\|_{\mathcal{X}_\delta}\|v\|_{\mathcal{Y}_\delta}} \geq \inf_{w\in\mathcal{X}_\delta}\frac{\partial\mathcal{G}(w,\tilde{u}_\delta^p,S^c w)}{\|w\|_{\mathcal{X}_\delta}\|S^c w\|_{\mathcal{Y}_\delta}}$$
$$\geq \underbrace{\left(\inf_{w\in\mathcal{X}_\delta}\frac{\|S^c w\|_{\mathcal{Y}_\delta}}{\|w\|_{\mathcal{X}_\delta}}\right)}_{\beta_\delta^c}\cdot\underbrace{\left(\inf_{w\in\mathcal{X}_\delta}\frac{\partial\mathcal{G}(w,\tilde{u}_\delta^p,S^c w)}{\|S^c w\|_{\mathcal{Y}_\delta}^2}\right)}_{\hat{\beta}_\delta^c(\mu)},$$

where we have identified the inf-sup constant evaluated at the centroid by β_δ^c and the *correction factor* by $\hat{\beta}_\delta^c(\mu)$. Note that the correction factor may be expressed as

$$\hat{\beta}_\delta^c(\mu) = \inf_{w\in\mathcal{X}_\delta}\frac{\partial\mathcal{G}(w,\tilde{u}_\delta^p,S^c w)}{\|S^c w\|_{\mathcal{Y}}^2} = \sum_{k=1}^{p+1}\inf_{w\in\mathcal{X}_\delta}\psi_k^p(\mu)\frac{\partial\mathcal{G}(w,u_{\delta,k},S^c w)}{\|S^c w\|_{\mathcal{Y}_\delta}^2}$$
$$= \sum_{k=1}^{p+1}\inf_{w\in\mathcal{X}_\delta}\psi_k^p(\mu)\frac{(S^k w,u_{\delta,k},S^c w)_{\mathcal{Y}_\delta}}{\|S^c w\|_{\mathcal{Y}_\delta}^2}.$$

Our LU method and SCM differ in the way they construct bounds of $\hat{\beta}_\delta^c(\mu)$.

Let us recast our LU formulation as a linear programming problem, the language in which the SCM is described. We compute a lower bound of the correction factor, $\hat{\beta}_{\delta,\mathrm{LB,LU}}^c(\mu) \leq \hat{\beta}_\delta^c(\mu)$, $\forall\mu\in\mathcal{D}_{\mathrm{work}}$, by first constructing a box in $\mathbb{R}^{p+1}$ that encapsulates the lower and upper bounds of contribution of each term of the correction factor, i.e.

$$B_{\mathrm{LU}} = \prod_{k=1}^{p+1}\left[\inf_{w\in\mathcal{X}_\delta}\frac{(S^k w,u_{\delta,k},S^c w)_{\mathcal{Y}}}{\|S^c w\|_{\mathcal{Y}_\delta}^2},\ \sup_{w\in\mathcal{X}_\delta}\frac{(S^k w,u_{\delta,k},S^c w)_{\mathcal{Y}}}{\|S^c w\|_{\mathcal{Y}_\delta}^2}\right].$$

Then, we solve a (rather simple) linear programming problem

$$\hat{\beta}_{\delta,\mathrm{LB,LU}}^c(\mu) = \inf_{y\in B^{\mathrm{LU}}}\sum_{k=1}^{p+1}\psi_k^p(\mu)y_k,$$

the solution to which is given by choosing either extrema for each coordinate of the bounding box B_{LU} based on the sign of $\psi_k^p(\mu)$, as explicitly stated in Section 3.2.2.

Let us now consider a special case of SCM where the SCM sampling points are the interpolation points, μ^k, $k=1,\ldots,p+1$, of the N_μ-p interpolation scheme. The SCM bounding box is given by

$$B_{\mathrm{SCM}} = \prod_{k=1}^{p+1}\left[-\frac{\gamma_{\delta,k}}{\beta_\delta^c},\frac{\gamma_{\delta,k}}{\beta_\delta^c}\right].$$

where

$$\gamma_{\delta,k} \equiv \sup_{w\in\mathcal{X}_\delta}\frac{\|S^k w\|_{\mathcal{Y}_\delta}}{\|w\|_{\mathcal{X}_\delta}}.$$

Since the kernel of $B_{\rm LU}$ is bounded by

$$\left|\frac{(S^k w, u_k, S^c w)_{\mathcal{Y}_\delta}}{\|S^c w\|^2_{\mathcal{Y}_\delta}}\right| \le \frac{\|S^k w\|_{\mathcal{Y}_\delta}}{\|w\|_{\mathcal{X}_\delta}} \frac{\|w\|_{\mathcal{X}_\delta}}{\|S^c w\|_{\mathcal{Y}_\delta}} \le \sup_{w\in\mathcal{X}_\delta} \frac{\|S^k w\|_{\mathcal{Y}_\delta}}{\|w\|_{\mathcal{X}_\delta}} \left(\inf_{w\in cY_\delta} \frac{\|S^c w\|_{\mathcal{Y}_\delta}}{\|w\|_{\mathcal{X}_\delta}}\right)^{-1} = \frac{\gamma_{\delta,k}}{\beta^c_\delta},$$

for $k = 1, \ldots, p+1$, we have

$$B_{\rm LU} \subset B_{\rm SCM}.$$

Furthermore, as the SCM sampling points correspond to the interpolation points, the SCM linear programming constraints

$$\sum_{k=1}^{p+1} \psi^p_k(\mu_l) y_k \ge \hat{\beta}^c_\delta(\mu_l), \quad l = 1, \ldots, p+1$$

simplify to (using $\psi^p_k(\mu_l) = \delta_{kl}$)

$$y_k \ge \hat{\beta}^c_\delta(\mu_k), \quad k = 1, \ldots, p+1,$$

where

$$\hat{\beta}^c_\delta(\mu_k) = \inf_{w\in\mathcal{X}_\delta} \frac{(S^k w, u_{\delta,k}, S^c w)_{\mathcal{Y}_\delta}}{\|S^c w\|^2_{\mathcal{Y}_\delta}}.$$

We recognize that the this constraint is in fact identical to the lower bound box constraint of $B_{\rm LU}$. Thus, the space over which the SCM lower bound is computed,

$$D^{\rm LB}_{\rm SCM} = \{y \in B^{\rm SCM} : y_k \ge \hat{\beta}^c_\delta(\mu_k), \quad k = 1, \ldots, p+1\},$$

satisfies

$$B_{\rm LU} \subset D^{\rm LB}_{\rm SCM}.$$

More specifically, $D^{\rm LB}_{\rm SCM}$ has the same lower bounds as $B_{\rm LU}$ but has looser upper bounds than $B_{\rm LU}$. Consequently, we have

$$\inf_{y\in D^{\rm LB}_{\rm SCM}} \sum_{k=1}^{p+1} \psi^p_k(\mu) y_k = \hat{\beta}^c_{\delta,{\rm LB,SCM}}(\mu) \le \hat{\beta}^c_{\delta,{\rm LB,LU}}(\mu) = \inf_{y\in B^{\rm LU}} \sum_{k=1}^{p+1} \psi^p_k(\mu) y_k \le \hat{\beta}^c_\delta.$$

Thus, if the SCM sampling points are the same as the interpolation points of the N_μ-p interpolation scheme, then our LU formulation gives a tighter inf-sup lower bound than SCM.

Acknowledgments

This work was supported by OSD/AFOSR/MURI Grant FA9550-09-1-0613, ONR Grant N00014-11-1-0713, and the Deutsche Forschungsgemeinschaft (DFG) under Ur-63/9 and GrK1100.

References

1. D. Amsallem and C. Farhat. Interpolation method for adapting reduced-order models and application to aeroelasticity. *AIAA J*, 46(7):1803–1813, 2008.
2. F. Brezzi, J. Rappaz, and P. A. Raviart. Finite dimensional approximation of nonlinear problems. part I: Branches of nonsingular solutions. *Numer. Math.*, 36:1–25, 1980.
3. S. Deparis. Reduced basis error bound computation of parameter-dependent Navier-Stokes equations by the natural norm approach. *SIAM J. Numer. Anal.*, 46(4):2039–2067, 2008.
4. J. L. Eftang, A. T. Patera, and E. M. Rønquist. An "*hp*" certified reduced basis method for parametrized elliptic partial differential equations. *SIAM J. Sci. Comput.*, 32(6):3170–3200, 2010.
5. M. A. Grepl and A. T. Patera. *A posteriori* error bounds for reduced-basis approximations of parametrized parabolic partial differential equations. *Math. Model. Numer. Anal.*, 39(1):157–181, 2005.
6. B. Haasdonk and M. Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *Math. Model. Numer. Anal.*, 42(2):277–302, 2008.
7. D. B. P. Huynh, D. J. Knezevic, Y. Chen, J. S. Hesthaven, and A. T. Patera. A natural-norm successive constraint method for inf-sup lower bounds. *Comput. Methods Appl. Mech. Engrg.*, 199:1963–1975, 2010.
8. D. J. Knezevic, N.-C. Nguyen, and A. T. Patera. Reduced basis approximation and *a posteriori* error estimation for the parametrized unsteady Boussinesq equations. *Math. Mod. Meth. Appl. S.*, 21(7):1415–1442, 2011.
9. Y. Maday. Private communication, Dec. 2012.
10. A. Manzoni. *Reduced models for optimal control, shape optimization and inverse problems in haemodynamics*. PhD thesis, École Polytechnique Fédérale de Lausanne, 2012.
11. N.-C. Nguyen, G. Rozza, and A. T. Patera. Reduced basis approximation and a posteriori error estimation for the time-dependent viscous Burgers' equation. *Calcolo*, 46:157–185, 2009.
12. A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer, New York, 1997.
13. C. Schwab and R. Stevenson. Space-time adaptive wavelet methods for parabolic evolution problems. *Math. Comp.*, 78(267):1293–1318, 2009.
14. K. Urban and A. T. Patera. An improved error bound for reduced basis approximation of linear parabolic problems. *Math. Comp.*, page submitted, 2012.
15. K. Urban and A. T. Patera. A new error bound for reduced basis approximation of parabolic partial differential equations. *C. R. Acad. Sci. Paris, Ser. I*, 350:203–207, 2012.
16. 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*, 47:773–788, 2005.
17. M. Yano. A space-time Petrov-Galerkin certified reduced basis method: Application to the Boussinesq equations. *SIAM J. Sci. Comput.*, submitted.