REDUCED BASIS ERROR BOUND COMPUTATION OF PARAMETER-DEPENDENT NAVIER–STOKES EQUATIONS BY THE NATURAL NORM APPROACH

SIMONE DEPARIS*

Abstract.
This work focuses on the *a posteriori* error estimation for the reduced basis method applied to partial differential equations with quadratic nonlinearity and affine parameter dependence. We rely on natural norms — *local* parameter-dependent norms — to provide a sharp and computable lower bound of the inf-sup constant. We prove a formulation of the Brezzi–Rappaz–Raviart existence and uniqueness theorem in the presence of two distinct norms. This allows us to relax the existence condition and to sharpen the field variable error bound. We also provide a robust algorithm to compute the Sobolev embedding constants involved in the error bound and in the inf-sup lower bound computation. We apply our method to a steady natural convection problem in a closed cavity, with Grashof number varying from 10 to 10^7.

Key words. Reduced basis methods, a posteriori error estimation, Brezzi–Rappaz–Raviart theory, steady incompressible Navier–Stokes equations, natural convection.

AMS subject classifications. 35Q30, 46E35,65N30, 76D03, 76D05.

1. Introduction. In this work we are interested in the numerical approximation of parameter ($\boldsymbol{\mu}$) dependent quadratic nonlinear partial differential equations — in particular the natural convection problem for Navier–Stokes equations — and the prediction of an "output of interest" which is a functional of the field variable $\boldsymbol{u}^{\mathrm{e}}(\boldsymbol{\mu})$,

$$s^{\mathrm{e}}(\boldsymbol{\mu}) = \ell(\boldsymbol{u}^{\mathrm{e}}(\boldsymbol{\mu})) \in \mathbb{R}, \tag{1.1}$$

where ℓ is a continuous linear form. The solution $\boldsymbol{u}^{\mathrm{e}}(\boldsymbol{\mu})$ may not be unique, nevertheless we consider only one solution branch; hence we presume local uniqueness.

The outputs are related to energies or forces, stresses or strains, flow-rates or pressure drops, temperatures or fluxes and are functions of an "input" parameter P-vector $\boldsymbol{\mu} \in \mathcal{D} \subset \mathbb{R}^P$, which is related to geometry, physical properties, boundary conditions, or loads.

We directly adopt the description and notation of Sen et al. [32], where the natural norms are introduced and applied to the error bound computation of coercive and non-coercive linear elliptic partial differential equations.

The field variable $\boldsymbol{u}^{\mathrm{e}}(\boldsymbol{\mu}) \in X^{\mathrm{e}}$ — say velocity, pressure, or temperature — satisfies in weak form the $\boldsymbol{\mu}$-parametrized nonlinear partial differential equation

$$a(\boldsymbol{u}^{\mathrm{e}}(\boldsymbol{\mu}), \boldsymbol{v}; \boldsymbol{\mu}) = f(\boldsymbol{v}), \quad \forall\, \boldsymbol{v} \in X^{\mathrm{e}}. \tag{1.2}$$

Here X^{e} is the appropriate function space with norm $\|\cdot\|_X$, $a(\cdot,\cdot;\boldsymbol{\mu})$ is continuous, quadratic in the first variable, and linear in the second one, and f is a continuous linear form.

Unfortunately, to achieve the desired accuracy, the evaluation $\boldsymbol{\mu} \rightarrow s^{\mathrm{e}}(\boldsymbol{\mu})$ by discrete projection methods like finite element or spectral methods is simply too costly in the many-query and real-time contexts often of interest in engineering. Low-order models — we consider here reduced basis approximations — are increasingly popular

*Mechanical Engineering Department, Massachusetts Institute of Technology,Rm 3-266, 77 Massachusetts Avenue, Cambridge, MA 02139 USA simone.deparis@epfl.ch

in the engineering analysis, parameter estimation, design optimization, and control contexts.

In the reduced basis approach [1, 18, 6, 22, 20, 13, 23, 11, 12, 25], we approximate $s^{\mathrm{e}}(\boldsymbol{\mu})$, $\boldsymbol{u}^{\mathrm{e}}(\boldsymbol{\mu})$ with $s_N(\boldsymbol{\mu}), \boldsymbol{u}_N(\boldsymbol{\mu})$: given $\boldsymbol{\mu} \in \mathcal{D}$,

$$s_N(\boldsymbol{\mu}) = \ell(\boldsymbol{u}_N(\boldsymbol{\mu})),$$

where $\boldsymbol{u}_N(\boldsymbol{\mu}) \in W_N$ satisfies[1]

$$a(\boldsymbol{u}_N(\boldsymbol{\mu}), \boldsymbol{v}; \boldsymbol{\mu}) = f(\boldsymbol{v}), \quad \forall\, \boldsymbol{v} \in W_N. \tag{1.3}$$

Here W_N is a problem-specific space of dimension $N \leq N_{\max}$ that focuses on the (typically very smooth) parametric manifold of interest — $\{\boldsymbol{u}^{\mathrm{e}}(\boldsymbol{\mu}) \,|\, \boldsymbol{\mu} \in \mathcal{D}\}$ — and thus the reduced basis method enjoys very rapid convergence as N increases [6, 15]; $N_{\max}$ is the maximal size of our reduced basis.

Our own effort is dedicated to the development of *a posteriori* error estimators for reduced basis approximations in the nonlinear case [13, 23, 35, 17, 11, 12, 24]. Since, in general, we can not find the exact (our superscript "e" above) solution, we replace $s^{\mathrm{e}}(\boldsymbol{\mu}), \boldsymbol{u}^{\mathrm{e}}(\boldsymbol{\mu})$ with a Galerkin finite element approximation[2] $s^{\mathcal{N}}(\boldsymbol{\mu})$, $\boldsymbol{u}^{\mathcal{N}}(\boldsymbol{\mu})$. $\mathcal{N}$ represents the dimension of our finite element space.

For some fixed sufficiently large $\mathcal{N} = \mathcal{N}_{\mathrm{t}}$, we presume that the finite element and the exact solution are very close. We shall call $\boldsymbol{u}^{\mathcal{N}_{\mathrm{t}}}(\boldsymbol{\mu})$ a "truth" solution, upon which we provide inexpensive — complexity *independent* of $\mathcal{N}_{\mathrm{t}}$ — and sharp error bounds $\Delta_N(\boldsymbol{\mu})$ and $\Delta_N^s(\boldsymbol{\mu})$ such that

$$\begin{aligned} &\|\boldsymbol{u}^{\mathcal{N}_{\mathrm{t}}}(\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})\|_X \leq \Delta_N(\boldsymbol{\mu}) \quad \text{and} \\ &|s^{\mathcal{N}_{\mathrm{t}}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu})| \leq \Delta_N^s(\boldsymbol{\mu}), \quad \forall\, \boldsymbol{\mu} \in \mathcal{D}. \end{aligned}$$

In the absence of such rigorous error bounds we can not efficiently determine if N is too small — and our reduced basis approximation unacceptably inaccurate — or if N is too large — and our reduced basis approximation unnecessarily expensive —, or we can not establish the very *existence* of a "truth" solution $\boldsymbol{u}^{\mathcal{N}_{\mathrm{t}}}(\boldsymbol{\mu})$ [8, 34, 33]. We can not determine in "real-time" if critical design conditions and constraints are satisfied. And, in fact, we can not even construct an efficient and well-conditioned reduced basis approximation space W_N [17, 16].

The dramatic *dimension reduction* — from $\mathcal{N}_t$ to N —, in conjunction with *offline/online computational procedures* [13, 23, 2, 9], yields very large savings in the many-query and real-time contexts: the online complexity depends only on the size of the reduced basis space, N, which is typically orders of magnitude smaller than the dimension of the finite element space, $\mathcal{N}_{\mathrm{t}}$. Hence we may consider a highly accurate truth approximation, at the detriment of the offline complexity only.

Relative to earlier work [35, 17, 34, 33], we adapt the *natural norms* technique proposed in [32] to the quadratic case in order to improve the construction and evaluation of a lower bound of the inf-sup constant (which we shall subsequently call "inf-sup lower bound"). We also extend the Brezzi–Rappaz–Raviart theory for the case of mixed norms (in our case, X- and natural norms); this enables to better control the effectivity of our error bound and to significantly reduce the constraint on

[1]For simplicity, we consider a purely primal approach; we shall subsequently describe also a primal-dual formulation.

[2]Other projection methods, like spectral or spectral elements methods, may be used as well, cf. [11, 12].

the dual norm of the residual for the existence and uniqueness of a (finite element) solution in the neighborhood of our reduced basis approximation.

The effectivity relative to the natural norm of the field variable error is proven to be very accurate: locally of order one when the dual norm of the residual is small. However, when considering only a primal formulation — as in the example that we provide —, we have no control of the effectivity on the output. To partially address this issue, we also propose a primal-dual formulation, whose apparent advantages would require further investigation, in particular in conjunction with deflation.

Our improvements enable a feasible and fast inf-sup lower bound construction based on a relatively small number of local norms. The inf-sup lower bound is of order one in the *local* natural norm and hence is intrinsically much simpler to compute than in existing works ([34, 17, 33]). The lower bound components are also simplified: they rely on the computation of Sobolev embedding constants and eigenvalues with respect to the natural norms. We develop robust tools to compute them based on multiple initial guesses modes and, regarding the Sobolev embedding constants, the restriction to a small discrete subspace to provide the initial guess and on a fixed point algorithm.

As a result, we are now able to cover much larger parameter sets than in earlier works. For example, in [17, 33] the authors considered natural convection with Grashof number (ratio buoyancy over viscosity) from 10 to 10^4. Instead, we have applied the reduced basis method for Grashof number up to 10^7 without increasing — actually by simplifying — the complexity of the error bound computation. To enable the finite element resolution of our problem at these values of the Grashof number and to speed up the offline work, we implemented a parallel code and a domain decomposition preconditioner based on a sequential one proposed in [5].

We first introduce the basic assumptions that we are going to use along the paper and define our natural norms for nonlinear partial differential equations (Section 2). We then prove an existence and unicity result based on the Brezzi–Rappaz–Raviart theory and provide an output error bound based on a pure primal approach (Section 3). In Section 4 we develop our inf-sup lower bound construction that enables a feasible offline/online computational strategy (Section 5). We then apply our approach to a steady natural convection problem (Section 6), provide the numerical results (Section 7), and present our conclusions (Section 8). The Appendix is devoted to the output error bound by a primal-dual formulation.

2. Problem statement.

2.1. Reduced basis formulation. For any $\boldsymbol{\mu} \equiv (\mu_1 \cdots \mu_P)$ in the closed input domain $\mathcal{D} \subset \mathbb{R}^P$, $a(\cdot,\cdot;\boldsymbol{\mu})\colon X^{\mathrm{e}} \times X^{\mathrm{e}} \to \mathbb{R}$ is a parameter-dependent form, which is quadratic with respect to its first argument and linear with respect to the second one. The forms ℓ and $f\colon X^{\mathrm{e}} \to \mathbb{R}$ are parameter-independent and linear (extension to parameter-dependent linear forms is straightforward; we provide also an extension to quadratic outputs).

Our "truth" or "reference" finite element approximation to the exact output and field variable, $s(\boldsymbol{\mu}) \equiv s^{\mathcal{N}_{\mathrm{t}}}(\boldsymbol{\mu})$ and $\boldsymbol{u}(\boldsymbol{\mu}) \equiv \boldsymbol{u}^{\mathcal{N}_{\mathrm{t}}}(\boldsymbol{\mu}) \in X^{\mathcal{N}_{\mathrm{t}}} \equiv X$, for a given $\boldsymbol{\mu} \in \mathcal{D}$,

$$
\begin{aligned}
s(\boldsymbol{\mu}) &= \ell(\boldsymbol{u}(\boldsymbol{\mu})) \quad \text{and} \\
a(\boldsymbol{u}(\boldsymbol{\mu}), \boldsymbol{v}; \boldsymbol{\mu}) &= f(\boldsymbol{v}), \quad \forall\, \boldsymbol{v} \in X.
\end{aligned} \tag{2.1}
$$

We assume that $\mathcal{N}_{\mathrm{t}}$ is chosen sufficiently large so that $s(\boldsymbol{\mu})$ and $\boldsymbol{u}(\boldsymbol{\mu})$ are essentially indistinguishable from $s^{\mathrm{e}}(\boldsymbol{\mu})$ and $\boldsymbol{u}^{\mathrm{e}}(\boldsymbol{\mu})$, respectively. We shall build our reduced basis approximation upon this "truth" approximation; and we shall evaluate the error in our reduced basis approximation with respect to this "truth" approximation. The online complexity (and stability) of our reduced basis approach is *independent* of $\mathcal{N}_\mathrm{t}$ (cf. [17, 33, 32]); hence, we may choose $\mathcal{N}_\mathrm{t}$ to be "arbitrarily" large at *no detriment to (online) performance*.

We shall suppose that our form a is "affine" in the parameter: for some fixed integers Q_0 and Q_1 — typically larger than P, sometimes by a considerable factor — we require

$$a(\boldsymbol{u},\boldsymbol{v};\boldsymbol{\mu}) = \sum_{q=1}^{Q_0} \Theta_0^q(\boldsymbol{\mu}) a_0^q(\boldsymbol{u},\boldsymbol{v}) + \frac{1}{2}\sum_{q=1}^{Q_1} \Theta_1^q(\boldsymbol{\mu}) a_1^q(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}), \quad \forall \boldsymbol{u},\boldsymbol{v} \in X,\ \forall \boldsymbol{\mu} \in \mathcal{D}, \tag{2.2}$$

where $\Theta_0^q, \Theta_1^q\colon\ \mathcal{D} \to \mathbb{R}$, $a_0^q\colon X \times X \to \mathbb{R}$, $1 \le q \le Q_0$, and $a_1^q\colon X \times X \times X \to \mathbb{R}$, $1 \le q \le Q_1$, are parameter-dependent functions and parameter-independent continuous bilinear and trilinear forms, respectively. (In particular, this assumption presume a quadratic nonlinearity in our partial differential equation.) Without loss of generality, we assume that, for $1 \le q \le Q_1$ and $\forall \boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in X$, $a_1^q(\boldsymbol{u},\boldsymbol{w},\boldsymbol{v}) = a_1^q(\boldsymbol{w},\boldsymbol{u},\boldsymbol{v})$. We shall further assume that $\Theta_i^q \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q_i$, $i = 0, 1$.

We denote the inner product and norm associated with our Hilbert space $X(\equiv X^{\mathcal{N}_\mathrm{t}}$, finite dimensional) as $(\boldsymbol{w},\boldsymbol{v})_X$ and $\|\boldsymbol{v}\|_X = \sqrt{(\boldsymbol{v},\boldsymbol{v})_X}$, respectively. We further define the dual norm for any bounded linear functional f as

$$\|f\|_{X'} \equiv \sup_{\boldsymbol{v}\in X} \frac{f(\boldsymbol{v})}{\|\boldsymbol{v}\|_X}.$$

For example, when considering two dimensional Navier-Stokes equations, the exact space X^e shall satisfy $H_0^1(\Omega) \times H_0^1(\Omega) \times L_0^2(\Omega) \subset X^\mathrm{e} \subset H^1(\Omega) \times H^1(\Omega) \times L^2(\Omega)$: here $\Omega \subset \mathbb{R}^2$ is a spatial domain with suitably regular boundary $\partial\Omega$; $L^2(\Omega)$ the Hilbert space of integrable functions; $H^1(\Omega)$ is the usual Hilbert space of derivative-square-integrable functions; $H_0^1(\Omega)$ indicates the functions with null trace on $\partial\Omega$ and $L_0^2(\Omega)$ the zero mean ones. The typical choice for our inner product $(\cdot,\cdot)_X$ is the standard $H^1(\Omega) \times H^1(\Omega) \times L^2(\Omega)$ inner product.

If in this example the Reynolds number is the parameter, then the affine decomposition (2.2) becomes, for $\boldsymbol{\mu} = (\mathrm{Re})$, $\boldsymbol{u} = (u_1, u_2, p)$, $\boldsymbol{v} = (v_1, v_2, q)$, $Q_0 = 2$, and $Q_1 = 1$,

$$\begin{aligned}
a_0^1(\boldsymbol{u},\boldsymbol{v}) &= -\int_\Omega p\partial_i v_i - \int_\Omega q\partial_i u_i + \lambda\int_\Omega p + \gamma\int_\Omega q && \Theta_0^1(\boldsymbol{\mu}) = 1\\
a_0^2(\boldsymbol{u},\boldsymbol{v}) &= \int_\Omega \partial_i u_j \partial_i v_j && \Theta_0^2(\boldsymbol{\mu}) = \mathrm{Re}^{-1}\\
a_1^1(\boldsymbol{u},\boldsymbol{z},\boldsymbol{v}) &= \int_\Omega u_j\partial_i z_i v_j + \int_\Omega z_j \partial_i u_i v_j && \Theta_1^1(\boldsymbol{\mu}) = 1
\end{aligned}$$

(with summation over repeated indices and partial derivatives with respect to x_1, x_2).

Recalling (2.2) and the symmetry of the a_1^q's in the first two variables, the Fréchet derivative of $a(\cdot,\cdot;\boldsymbol{\mu})$ with respect to the first variable at a point $\boldsymbol{w} \in X$ can be expressed as

$$da(\boldsymbol{w};\boldsymbol{\mu})(\boldsymbol{u},\boldsymbol{v}) = \sum_{q=1}^{Q_0} \Theta_0^q(\boldsymbol{\mu}) a_0^q(\boldsymbol{u},\boldsymbol{v}) + \sum_{q=1}^{Q_1} \Theta_1^q(\boldsymbol{\mu}) a_1^q(\boldsymbol{u},\boldsymbol{w},\boldsymbol{v}), \quad \forall\, \boldsymbol{u},\boldsymbol{v} \in X,\ \forall\, \boldsymbol{\mu} \in \mathcal{D}.$$

For any $\boldsymbol{\mu}$ in $\mathcal{D}$ and any solution $\boldsymbol{u}(\boldsymbol{\mu})$ of (2.1) in the region of our interest, we assume that $da(\boldsymbol{u}(\boldsymbol{\mu});\boldsymbol{\mu})$ is "stable" and continuous in the sense that there exist $\beta_0 > 0$ and $\gamma_0 \in \mathbb{R}$ such that $\forall \boldsymbol{\mu} \in \mathcal{D}$,

$$0 < \beta_0 < \beta(\boldsymbol{\mu}) \equiv \inf_{\boldsymbol{w}\in X} \sup_{\boldsymbol{v}\in X} \frac{da(\boldsymbol{u}(\boldsymbol{\mu});\boldsymbol{\mu})(\boldsymbol{w},\boldsymbol{v})}{\|\boldsymbol{w}\|_X \|\boldsymbol{v}\|_X}, \tag{2.3}$$

$$\infty > \gamma_0 > \gamma(\boldsymbol{\mu}) \equiv \sup_{\boldsymbol{w}\in X} \sup_{\boldsymbol{v}\in X} \frac{da(\boldsymbol{u}(\boldsymbol{\mu});\boldsymbol{\mu})(\boldsymbol{w},\boldsymbol{v})}{\|\boldsymbol{w}\|_X \|\boldsymbol{v}\|_X}.$$

It then follows that in the neighborhood of $\boldsymbol{u}(\boldsymbol{\mu})$ the solution to (2.1) is unique. We further assume that ℓ and f are in X' — bounded linear functionals.

2.2. Natural norms. Let $\boldsymbol{u}_N(\boldsymbol{\mu})$ be the reduced basis approximation to $\boldsymbol{u}(\boldsymbol{\mu})$ given by (1.3). We next introduce [35, 17, 7, 14, 32] the parametrized linear operator $T_N^{\boldsymbol{\mu}}\colon X \to X$, such that

$$(T_N^{\boldsymbol{\mu}}\boldsymbol{w},\boldsymbol{v})_X = da(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})(\boldsymbol{w},\boldsymbol{v}), \forall\, \boldsymbol{v},\boldsymbol{w} \in X. \tag{2.4}$$

Our method — in particular our inf-sup lower bound construction — requires a discrete set of K parameter values, $\mathcal{V}^K \equiv \{\overline{\boldsymbol{\mu}}_1,\ldots,\overline{\boldsymbol{\mu}}_K\} \subset \mathcal{D}$ — upon which to construct local norms —, a fixed integer $\overline{N} \le N_{\max}$,[3] and an indicator function $\mathcal{I}^K\colon\ \mathcal{D} \to \mathcal{V}^K$ which associates to any $\boldsymbol{\mu}$ in $\mathcal{D}$ a member of $\mathcal{V}^K$.[4] Typically, $\mathcal{I}^K$ defines the "closest" — in a sense that we have to define — element of $\mathcal{D}$.

We assume that, $\forall \overline{\boldsymbol{\mu}} \in \mathcal{V}^K$,

$$0 < \beta_{\overline{N}}(\overline{\boldsymbol{\mu}}) \equiv \inf_{\boldsymbol{w}\in X} \sup_{\boldsymbol{v}\in X} \frac{da(\boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}});\overline{\boldsymbol{\mu}})(\boldsymbol{w},\boldsymbol{v})}{\|\boldsymbol{w}\|_X \|\boldsymbol{v}\|_X} \equiv \inf_{\boldsymbol{w}\in X} \frac{\|T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}\|_X}{\|\boldsymbol{w}\|_X}, \tag{2.5}$$

$$\infty > \gamma_{\overline{N}}(\overline{\boldsymbol{\mu}}) \equiv \sup_{\boldsymbol{w}\in X} \sup_{\boldsymbol{v}\in X} \frac{da(\boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}});\overline{\boldsymbol{\mu}})(\boldsymbol{w},\boldsymbol{v})}{\|\boldsymbol{w}\|_X \|\boldsymbol{v}\|_X} \equiv \sup_{\boldsymbol{w}\in X} \frac{\|T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}\|_X}{\|\boldsymbol{w}\|_X}. \tag{2.6}$$

Note that this is true if $\boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}})$ is close enough to $\boldsymbol{u}(\overline{\boldsymbol{\mu}})$ (even though this is not required here); from the Cauchy-Schwarz inequality, $\boldsymbol{v} = T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}$ is the inner supremizer in (2.5) and (2.6).

We then introduce, for any given $\overline{\boldsymbol{\mu}} \in \mathcal{V}^K$ and $\overline{N}$, our "natural norm"

$$|||\boldsymbol{v}|||_{\overline{\boldsymbol{\mu}}} \equiv \|T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{v}\|_X, \quad \forall\, \boldsymbol{v} \in X. \tag{2.7}$$

(To simplify the notation we have dropped the index $_{\overline{N}}$ from the norm symbol.) This norm is the extension of the natural norm introduced in [32] for the linear case. Note that, thanks to our assumptions on $\beta_{\overline{N}}(\overline{\boldsymbol{\mu}})$ and $\gamma_{\overline{N}}(\overline{\boldsymbol{\mu}})$, (2.7) does indeed define a norm, which is equivalent to $\|\cdot\|_X$.

2.3. Trilinear forms continuity constants. In the development of our error bound, we assume that there exists a positive $\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ such that

$$\left|da(\boldsymbol{z}^2;\boldsymbol{\mu})(\boldsymbol{v},\boldsymbol{w}) - da(\boldsymbol{z}^1;\boldsymbol{\mu})(\boldsymbol{v},\boldsymbol{w})\right| \le \rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\, |||\boldsymbol{z}^2 - \boldsymbol{z}^1|||_{\overline{\boldsymbol{\mu}}}\, |||\boldsymbol{v}|||_{\overline{\boldsymbol{\mu}}}\, \|\boldsymbol{w}\|_X, \tag{2.8}$$

for all $\boldsymbol{z}^1$, $\boldsymbol{z}^2$, $\boldsymbol{v}$, and $\boldsymbol{w}$ in X.

[3] In practice, $\overline{N}$ may depend on $\overline{\boldsymbol{\mu}}$ and the results of this work still hold.

[4] The process by which we select "good" $\mathcal{V}^K$ and $\mathcal{I}^K$ is described in Section 5.

Similarly, in the construction of an inf-sup lower bound, we assume that there exists a positive $\rho_{X,\overline{\boldsymbol{\mu}}}$ such that

$$\left|\sum_{q=1}^{Q_1}\Theta_1^q(\overline{\boldsymbol{\mu}})a_1^q(\boldsymbol{v},\boldsymbol{z},\boldsymbol{w})\right| \le \rho_{X,\overline{\boldsymbol{\mu}}}\,\|\boldsymbol{z}\|_X\,|||\boldsymbol{v}|||_{\overline{\boldsymbol{\mu}}}\,\|\boldsymbol{w}\|_X, \tag{2.9}$$

$$\left|\sum_{q=1}^{Q_1}\left(\Theta_1^q(\boldsymbol{\mu})-\Theta_1^q(\overline{\boldsymbol{\mu}})\right)a_1^q(\boldsymbol{v},\boldsymbol{z},\boldsymbol{w})\right| \le \max_{q=1,\dots,Q_1}\frac{|\Theta_1^q(\boldsymbol{\mu})-\Theta_1^q(\overline{\boldsymbol{\mu}})|}{|\Theta_1^q(\overline{\boldsymbol{\mu}})|}\rho_{X,\overline{\boldsymbol{\mu}}}\,\|\boldsymbol{z}\|_X\,|||\boldsymbol{v}|||_{\overline{\boldsymbol{\mu}}}\,\|\boldsymbol{w}\|_X, \tag{2.10}$$

for all $\boldsymbol{z}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ in X.

To a given problem, we have to provide these constants; in Sections 6.3 and 6.5 we show how to compute these constants, which are related to the Sobolev embedding theorem, for the model problem at hand. The algorithm proposed in section 6.5 can be easily extended to other examples.

3. Existence and uniqueness. Let $\boldsymbol{\mu}\in\mathcal{D}$, $\overline{\boldsymbol{\mu}}\in\mathcal{V}^K$, $N\le N_{\max}$, and $\overline{N}\le N_{\max}$; we define (implicitly dropping the subscripts $_{\overline{N}}$ and $_N$)

$$\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \equiv \inf_{\boldsymbol{w}\in X}\sup_{\boldsymbol{v}\in X}\frac{(T_N^{\boldsymbol{\mu}}\boldsymbol{w},\boldsymbol{v})_X}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}\|\boldsymbol{v}\|_X} = \inf_{\boldsymbol{w}\in X}\frac{\|T_N^{\boldsymbol{\mu}}\boldsymbol{w}\|_X}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}} = \inf_{\boldsymbol{w}\in X}\frac{\|T_N^{\boldsymbol{\mu}}\boldsymbol{w}\|_X}{\|T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}\|_X}, \tag{3.1}$$

$$\gamma_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \equiv \sup_{\boldsymbol{w}\in X}\sup_{\boldsymbol{v}\in X}\frac{(T_N^{\boldsymbol{\mu}}\boldsymbol{w},\boldsymbol{v})_X}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}\|\boldsymbol{v}\|_X} = \sup_{\boldsymbol{w}\in X}\frac{\|T_N^{\boldsymbol{\mu}}\boldsymbol{w}\|_X}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}} = \sup_{\boldsymbol{w}\in X}\frac{\|T_N^{\boldsymbol{\mu}}\boldsymbol{w}\|_X}{\|T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}\|_X}. \tag{3.2}$$

We assume that $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})>0$ (this is the case when $\beta(\boldsymbol{\mu})>0$, cf. (2.3), and $\boldsymbol{u}_N(\boldsymbol{\mu})$ is "nearby" to $\boldsymbol{u}(\boldsymbol{\mu})$) and also that we explicitly know either $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ or a rigorous positive lower bound to it. The results in this section extend straightforward in the latter case by replacing $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ by its lower bound.

3.1. Uniqueness. We are now ready to prove a uniqueness result in the neighborhood — with respect to the natural norms — of a given reduced basis approximation. Let $B_{\overline{\boldsymbol{\mu}}}(\boldsymbol{u},\epsilon)$ be an open ball around $\boldsymbol{u}\in X$, of radius $\epsilon>0$, and metric induced by $|||\cdot|||_{\overline{\boldsymbol{\mu}}}$.

LEMMA 3.1. *Assume that $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})>0$. There can be at most one solution $\boldsymbol{u}(\boldsymbol{\mu})$ to* (2.1) *such that*

$$\boldsymbol{u}(\boldsymbol{\mu})\in B_{\overline{\boldsymbol{\mu}}}\left(\boldsymbol{u}_N(\boldsymbol{\mu}),\frac{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}\right).$$

Proof. The proofs of this lemma and the following one are based on the original proofs in [4] and is extended to the case of two distinct norms. For these proofs we need to introduce the operators $A(\cdot;\boldsymbol{\mu}):X\mapsto X'$ and $DA(\boldsymbol{z},\boldsymbol{\mu}):X\mapsto X'$, for $\boldsymbol{z}$ in X, associated to a (and f) and da, and the operator $H(\cdot;\boldsymbol{\mu}):X\to X$,

$$\langle A(\boldsymbol{u};\boldsymbol{\mu}),\boldsymbol{v}\rangle = a(\boldsymbol{u},\boldsymbol{v};\boldsymbol{\mu})-f(\boldsymbol{v})\ \forall\boldsymbol{u},\boldsymbol{v}\in X$$

$$\langle DA(\boldsymbol{z},\boldsymbol{\mu})\boldsymbol{u},\boldsymbol{v}\rangle = da(\boldsymbol{z};\boldsymbol{\mu})(\boldsymbol{u},\boldsymbol{v})\ \forall\boldsymbol{u},\boldsymbol{v}\in X. \tag{3.3}$$

$$H(\boldsymbol{w};\boldsymbol{\mu}) = \boldsymbol{w}-DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})^{-1}A(\boldsymbol{w};\boldsymbol{\mu}); \tag{3.4}$$

$DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})$ is invertible thanks to our assumption $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})>0$. A fixed point of H must be a solution to (2.1) and vice-versa; we are going to show that there exists at most one.

We consider $H(\boldsymbol{z}^2,\boldsymbol{\mu})-H(\boldsymbol{z}^1,\boldsymbol{\mu})$, which from the definition (3.4) can be expressed as

$$H(\boldsymbol{z}^2,\boldsymbol{\mu})-H(\boldsymbol{z}^1,\boldsymbol{\mu}) = (\boldsymbol{z}^2-\boldsymbol{z}^1) - DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})^{-1}\left(A(\boldsymbol{z}^2;\boldsymbol{\mu})-A(\boldsymbol{z}^1;\boldsymbol{\mu})\right). \quad (3.5)$$

We observe that since our nonlinearity is quadratic,

$$A(\boldsymbol{z}^2;\boldsymbol{\mu})-A(\boldsymbol{z}^1;\boldsymbol{\mu}) = DA(\frac{\boldsymbol{z}^1+\boldsymbol{z}^2}{2};\boldsymbol{\mu})(\boldsymbol{z}^2-\boldsymbol{z}^1). \quad (3.6)$$

Applying (3.5) and (3.6) gives

$$\begin{aligned} DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})\left(H(\boldsymbol{z}^2;\boldsymbol{\mu})-H(\boldsymbol{z}^1;\boldsymbol{\mu})\right) &= DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})\left(\boldsymbol{z}^2-\boldsymbol{z}^1\right)-\left(A(\boldsymbol{z}^2;\boldsymbol{\mu})-A(\boldsymbol{z}^1;\boldsymbol{\mu})\right) \\ &= \left[DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})-DA\left(\frac{\boldsymbol{z}^1+\boldsymbol{z}^2}{2};\boldsymbol{\mu}\right)\right](\boldsymbol{z}^2-\boldsymbol{z}^1). \end{aligned}$$

We then have, recalling (2.8),

$$\left\langle DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})(H(\boldsymbol{z}^2;\boldsymbol{\mu})-H(\boldsymbol{z}^1;\boldsymbol{\mu})),\boldsymbol{v}\right\rangle \le \rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})|||\boldsymbol{u}_N(\boldsymbol{\mu})-\frac{1}{2}(\boldsymbol{z}^2+\boldsymbol{z}^1)|||_{\overline{\boldsymbol{\mu}}}|||\boldsymbol{z}^2-\boldsymbol{z}^1|||_{\overline{\boldsymbol{\mu}}}\|\boldsymbol{v}\|_X.$$

Furthermore, from the definition of $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$, $T_N^{\boldsymbol{\mu}}$, and $DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})$ ((3.1), (2.4) and (3.3) respectively) we have

$$\begin{aligned} &\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})|||H(\boldsymbol{z}^2;\boldsymbol{\mu})-H(\boldsymbol{z}^1;\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}}\|T_N^{\boldsymbol{\mu}}(H(\boldsymbol{z}^2;\boldsymbol{\mu})-H(\boldsymbol{z}^1;\boldsymbol{\mu}))\|_X \\ &\quad\le \left\langle DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})(H(\boldsymbol{z}^2;\boldsymbol{\mu})-H(\boldsymbol{z}^1;\boldsymbol{\mu})), T_N^{\boldsymbol{\mu}}(H(\boldsymbol{z}^2;\boldsymbol{\mu})-H(\boldsymbol{z}^1;\boldsymbol{\mu}))\right\rangle. \quad (3.7) \end{aligned}$$

Let $\alpha>0$; if $\boldsymbol{z}^1$ and $\boldsymbol{z}^2$ are in $B_{\overline{\boldsymbol{\mu}}}(\boldsymbol{u}_N(\boldsymbol{\mu}),\alpha)$, then $|||\boldsymbol{u}_N(\boldsymbol{\mu})-\frac{1}{2}(\boldsymbol{z}^2+\boldsymbol{z}^1)|||_{\overline{\boldsymbol{\mu}}}<\alpha$ and the last two inequalities give

$$|||H(\boldsymbol{z}^2;\boldsymbol{\mu})-H(\boldsymbol{z}^1;\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}} < \frac{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\alpha}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}|||\boldsymbol{z}^2-\boldsymbol{z}^1|||_{\overline{\boldsymbol{\mu}}},$$

which implies that $H(\cdot;\boldsymbol{\mu})$ is a contraction mapping for

$$\alpha < \frac{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}.$$

It follows that there can be only (at most) one fixed point of $H(\cdot;\mu)$ inside $B_{\overline{\boldsymbol{\mu}}}(\boldsymbol{u}_N(\boldsymbol{\mu}),\frac{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})})$ and hence at most one solution $\boldsymbol{u}(\boldsymbol{\mu})$. ☐

3.2. Existence. We define the following quantities:

$$\epsilon_N(\boldsymbol{\mu}) \equiv \|A(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})\|_{X'}, \quad (3.8)$$

$$\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \equiv \frac{2\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\epsilon_N(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})^2}, \quad (3.9)$$

$$\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \equiv \frac{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}\left[1-\sqrt{1-\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}\right]. \quad (3.10)$$

Lemma 3.2. *If $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})>0$ and $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\le 1$, there exists a unique solution $\boldsymbol{u}(\boldsymbol{\mu})$ to* (2.1) *such that*

$$\boldsymbol{u}(\boldsymbol{\mu}) \in B_{\overline{\boldsymbol{\mu}}}\Big(\boldsymbol{u}_N(\boldsymbol{\mu}),\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\Big).$$

Proof. Let $\alpha > 0$ and $\boldsymbol{z}$ in $B_{\overline{\boldsymbol{\mu}}}(\boldsymbol{u}_N(\boldsymbol{\mu}), \alpha)$; consider

$$\begin{aligned} &H(\boldsymbol{z};\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu}) = \boldsymbol{z} - \boldsymbol{u}_N(\boldsymbol{\mu}) \\ &\quad - DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})^{-1}\left(A(\boldsymbol{z};\boldsymbol{\mu}) - A(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})\right) - DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})^{-1} A(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu}). \end{aligned}$$

Using arguments similar to those invoked in the previous lemma (among others that $\boldsymbol{u}(\boldsymbol{\mu})$ is a fixed point of $H(\cdot;\boldsymbol{\mu})$), we have that

$$\begin{aligned} &\langle DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})(H(\boldsymbol{z};\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})), \boldsymbol{v}\rangle \\ &= \langle DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})(\boldsymbol{z} - \boldsymbol{u}_N(\boldsymbol{\mu})), \boldsymbol{v}\rangle - \langle A(\boldsymbol{z};\boldsymbol{\mu}) - A(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu}), \boldsymbol{v}\rangle - \langle A(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu}), \boldsymbol{v}\rangle \\ &= \left\langle \left[DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu}) - DA\left(\frac{\boldsymbol{z} + \boldsymbol{u}_N(\boldsymbol{\mu})}{2};\boldsymbol{\mu}\right)\right](\boldsymbol{z} - \boldsymbol{u}_N(\boldsymbol{\mu})), \boldsymbol{v}\right\rangle - \langle A(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu}), \boldsymbol{v}\rangle \\ &\qquad \le \frac{1}{2}\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) |||\boldsymbol{z} - \boldsymbol{u}_N(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}}^2 \|\boldsymbol{v}\|_X + \epsilon_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\|\boldsymbol{v}\|_X, \end{aligned} \tag{3.11}$$

for all $\boldsymbol{v}$ in X. Similarly to (3.7),

$$\begin{aligned} &\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) |||H(\boldsymbol{z};\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}} \|T_N^{\boldsymbol{\mu}}(H(\boldsymbol{z};\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu}))\|_X. \\ &\qquad \le \langle DA(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})(H(\boldsymbol{z};\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})), T_N^{\boldsymbol{\mu}}(H(\boldsymbol{z};\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu}))\rangle \end{aligned} \tag{3.12}$$

Since $\boldsymbol{z}$ is in the open ball $B_{\overline{\boldsymbol{\mu}}}(\boldsymbol{u}_N(\boldsymbol{\mu}), \alpha)$, choosing $\boldsymbol{v} = T_N^{\boldsymbol{\mu}}(H(\boldsymbol{z};\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu}))$, (3.11), and (3.12) give

$$|||H(\boldsymbol{z};\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}} \le \frac{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{2\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})} |||\boldsymbol{z} - \boldsymbol{u}_N(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}}^2 + \frac{\epsilon_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})} < \frac{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\alpha^2}{2\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})} + \frac{\epsilon_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})} \tag{3.13}$$

We now look for values of α (in particular, the smallest will give us a sharper error bound), such that

$$\frac{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\alpha^2}{2\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})} + \frac{\epsilon_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})} \le \alpha.$$

Provided that $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \le 1$, we have that the smallest value is given by $\alpha = \Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$. Then for this choice and from (3.13) we have

$$|||H(\boldsymbol{z};\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}} < \alpha,$$

i.e., H maps $\in B_{\overline{\boldsymbol{\mu}}}(\boldsymbol{u}_N(\boldsymbol{\mu}), \Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}))$ into itself. The proof follows from $\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \le \frac{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}$, the fact that H is a contraction (cf. proof of the previous lemma), and the Contraction Mapping Theorem. ◻

3.3. A posteriori error estimation.

Theorem 3.3. *Let $\boldsymbol{\mu} \in \mathcal{D}$ and assume that for a $\overline{\boldsymbol{\mu}} \in \mathcal{V}^K$, $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) > 0$ and $\tau_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \le 1$. Then there exists a unique solution $\boldsymbol{u}(\boldsymbol{\mu})$ such that*

$$|||\boldsymbol{u}(\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}} < \frac{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})};$$

the error with respect to the natural norm is

$$|||\boldsymbol{u}(\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}} \le \Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \tag{3.14}$$

with effectivity

$$\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \leq \left[\frac{2\gamma_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})} + \tau_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\right] |||\boldsymbol{u}(\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}}. \tag{3.15}$$

This theorem provides an error bound with respect to the natural norms that, compared to previous results (cf. [17, 33]), firstly provides better effectivity in the field variable, since we expect $\gamma_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ and $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ to be of order one; and secondly indirectly sharpens the existence's condition since in (3.9) again $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ is of order one.

Note: we expect that $\gamma_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ and $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ are of order one, since $\overline{\boldsymbol{\mu}} = \mathcal{I}^K(\boldsymbol{\mu})$ is "close" to $\boldsymbol{\mu}$ and hence $T_N^{\boldsymbol{\mu}}$ and $T_N^{\overline{\boldsymbol{\mu}}}$ have similar spectra. In (3.9) the term that is the worse controlled is $\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$; this is related to a Sobolev embedding constant (cf. [33] and Section 6). However, the equivalent condition in [33] reads

$$\tau_N(\boldsymbol{\mu}) = \frac{2\rho\,\epsilon_N(\boldsymbol{\mu})}{\beta_N(\boldsymbol{\mu})^2},$$

where ρ, $\beta_N(\boldsymbol{\mu})$ respectively, is the Sobolev embedding constant equivalent to (2.8), (3.1) respectively, with respect to the X-norm only. Since $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \geq \frac{\beta_N(\boldsymbol{\mu})}{\gamma(\overline{\boldsymbol{\mu}})}$ and $\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \leq \frac{\rho}{\beta(\overline{\boldsymbol{\mu}})^2}$, we have that

$$\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \leq \tau_N(\boldsymbol{\mu})\frac{\gamma(\overline{\boldsymbol{\mu}})^2}{\beta(\overline{\boldsymbol{\mu}})^2},$$

i.e., if $\gamma_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ and $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ are of order one, the condition $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \leq 1$ is less stringent than $\tau_N(\boldsymbol{\mu}) < 1$, which is required in [33].

In our simulations, for large Grashof number, we actually get that $\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ is one order smaller than $\frac{\rho}{\beta_N(\boldsymbol{\mu})^2}$, hence $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ is also one order smaller than $\tau_N(\boldsymbol{\mu})$.

Proof. We just need to prove (3.15). Let $\boldsymbol{e}(\boldsymbol{\mu}) = \boldsymbol{u}(\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})$ and $\hat{\boldsymbol{e}}(\boldsymbol{\mu})$ such that

$$\begin{aligned}(\hat{\boldsymbol{e}}(\boldsymbol{\mu}), \boldsymbol{v})_X &= -\langle A(\boldsymbol{u}_N(\boldsymbol{\mu}); \boldsymbol{\mu}), \boldsymbol{v}\rangle \\ &= f(\boldsymbol{v}) - a(\boldsymbol{u}_N(\boldsymbol{\mu}), \boldsymbol{v}; \boldsymbol{\mu}) = a(\boldsymbol{u}(\boldsymbol{\mu}), \boldsymbol{v}; \boldsymbol{\mu}) - a(\boldsymbol{u}_N(\boldsymbol{\mu}), \boldsymbol{v}; \boldsymbol{\mu}).\end{aligned} \tag{3.16}$$

Note that, from (3.8), $\|\hat{\boldsymbol{e}}(\boldsymbol{\mu})\|_X = \epsilon_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ and that, since our non-linearity is quadratic,

$$\begin{aligned}&a(\boldsymbol{u}(\boldsymbol{\mu}), \boldsymbol{v}; \boldsymbol{\mu}) - a(\boldsymbol{u}_N(\boldsymbol{\mu}), \boldsymbol{v}; \boldsymbol{\mu}) = \\ &\quad = da(\boldsymbol{u}_N(\boldsymbol{\mu}); \boldsymbol{\mu})(\boldsymbol{e}(\boldsymbol{\mu}), \boldsymbol{v}) + \frac{1}{2}\Big(da(\boldsymbol{u}(\boldsymbol{\mu}); \boldsymbol{\mu}) - da(\boldsymbol{u}_N(\boldsymbol{\mu}); \boldsymbol{\mu})\Big)(\boldsymbol{e}(\boldsymbol{\mu}), \boldsymbol{v})\end{aligned}$$

for all $\boldsymbol{v}$ in X. This, (3.16), (3.2), and (2.8) yield to

$$\begin{aligned}\|\hat{\boldsymbol{e}}(\boldsymbol{\mu})\|_X^2 &= da(\boldsymbol{u}_N(\boldsymbol{\mu}); \boldsymbol{\mu})(\boldsymbol{e}(\boldsymbol{\mu}), \hat{\boldsymbol{e}}(\boldsymbol{\mu})) + \frac{1}{2}\Big(da(\boldsymbol{u}(\boldsymbol{\mu}); \boldsymbol{\mu}) - da(\boldsymbol{u}_N(\boldsymbol{\mu}); \boldsymbol{\mu})\Big)(\boldsymbol{e}(\boldsymbol{\mu}), \hat{\boldsymbol{e}}(\boldsymbol{\mu})) \\ &\leq \gamma_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})|||\boldsymbol{e}(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}}\|\hat{\boldsymbol{e}}(\boldsymbol{\mu})\|_X + \frac{1}{2}\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\,|||\boldsymbol{e}(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}}^2\|\hat{\boldsymbol{e}}(\boldsymbol{\mu})\|_X\end{aligned}$$

and therefore

$$\epsilon_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \leq \gamma_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})|||\boldsymbol{e}(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}} + \frac{1}{2}\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})|||\boldsymbol{e}(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}}^2. \tag{3.17}$$

Since $0 \le \tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \le 1$ and $1 - \sqrt{1 - \tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})} \le \tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$, from (3.9)-(3.10) we get

$$\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \le 2\frac{\epsilon_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})} \quad \text{and} \tag{3.18}$$

$$\frac{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \le \tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}). \tag{3.19}$$

Then, from (3.17) and (3.18)

$$\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \le \frac{2\gamma_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}|||\boldsymbol{e}(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}} + \frac{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}|||\boldsymbol{e}(\boldsymbol{\mu})|||^2_{\overline{\boldsymbol{\mu}}};$$

we split the square of $|||\boldsymbol{e}(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}}$ and get

$$\begin{aligned}\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) &\le \frac{2\gamma_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}|||\boldsymbol{e}(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}} + \frac{\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}|||\boldsymbol{e}(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}}\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\\ &\le \frac{2\gamma_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})}|||\boldsymbol{e}(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}} + \tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})|||\boldsymbol{e}(\boldsymbol{\mu})|||_{\overline{\boldsymbol{\mu}}},\end{aligned}$$

thanks to (3.14) in the first inequality and (3.19) in the second one.
$\square$

3.4. Output error bound. Often the precision required in the prediction of the output is obtained by the primal approximation alone. This is the case in the model problems that we propose; in fact, the requirements to ensure existence and unicity already implies a very small error in the output.

For a given $\boldsymbol{\mu} \in \mathcal{D}$, we are interested in bounding the error on the output $s(\boldsymbol{\mu}) = \ell(\boldsymbol{u}(\boldsymbol{\mu}))$. For $\overline{\boldsymbol{\mu}} \in \mathcal{V}^K$, the *natural* dual norm of the output linear functional is defined as

$$|||\ell|||_{\overline{\boldsymbol{\mu}}'} = \sup_{\boldsymbol{v}\in Y}\frac{|\ell(\boldsymbol{v})|}{|||\boldsymbol{v}|||_{\overline{\boldsymbol{\mu}}}}.$$

(Note that in the case of an affine output, $|||\ell|||_{\overline{\boldsymbol{\mu}}'} \le \sum_{q=1}^{Q_\ell}|\theta_\ell^q(\boldsymbol{\mu})|\ |||\ell^q|||_{\overline{\boldsymbol{\mu}}'}$.) If $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) > 0$ and $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \le 1$, we can then state the output error bound as

$$|s(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu})| \le |||\ell|||_{\overline{\boldsymbol{\mu}}'}\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \equiv \Delta^s_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}).$$

However, we do not have any effectivity result about these bounds. In our example, the effectivities are at high values of the Grashof number reasonable, but not as accurate at low values. In the Appendix we propose a dual approach that partially addresses this issue; further improvements shall require deflation, cf. [32].

4. Lower bound of the inf-sup constant. We are going to build a lower bound for the inf-sup parameter $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$. The ingredients, similarly to [32], are the selection of a candidate supremizer, a Taylor expansion with remainder, and a lower bound for the terms in the Taylor expansion.

It is clear that, for $\boldsymbol{\mu} = \overline{\boldsymbol{\mu}}$ in (3.1) and (3.2), $\beta_{\overline{\boldsymbol{\mu}}}(\overline{\boldsymbol{\mu}}) = \gamma_{\overline{\boldsymbol{\mu}}}(\overline{\boldsymbol{\mu}}) = 1$; the natural norm can thus be viewed as a generalization of the usual energy norm (for symmetric, coercive operators) to the nonlinear case. It can be demonstrated that

$$\beta(\overline{\boldsymbol{\mu}})\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \le \beta(\boldsymbol{\mu}) \le \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\gamma(\overline{\boldsymbol{\mu}}).$$

In what follows we shall require an "intermediate" inf-sup parameter — an approximation to $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ — which we denote as $\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$: for given $\boldsymbol{\mu} \in \mathcal{D}$ and $\overline{\boldsymbol{\mu}} \in \mathcal{V}^K$,

$$\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \equiv \inf_{\boldsymbol{w} \in X} \frac{(T_N^{\boldsymbol{\mu}} \boldsymbol{w}, T_N^{\overline{\boldsymbol{\mu}}} \boldsymbol{w})_X}{|||\boldsymbol{w}|||^2_{\overline{\boldsymbol{\mu}}}}.$$

It follows directly from the Cauchy–Schwarz inequality — or equivalently, we may observe that $T_N^{\overline{\boldsymbol{\mu}}} \boldsymbol{w}$ is a candidate supremizer $\boldsymbol{v}$ in (3.1) — that

$$\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \leq \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}), \quad \forall\, \boldsymbol{\mu} \in \mathcal{D}.$$

(Note that $\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ is not necessarily positive.)

We can also show that $\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ is a "good" lower bound for $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$:

Lemma 4.1. *Let*

$$\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}} = \sup_{\boldsymbol{w} \in X} \frac{\left(T_N^{\boldsymbol{\mu}} \boldsymbol{w} - T_N^{\overline{\boldsymbol{\mu}}} \boldsymbol{w}, T_N^{\boldsymbol{\mu}} \boldsymbol{w} + T_N^{\overline{\boldsymbol{\mu}}} \boldsymbol{w}\right)_X}{\|T_N^{\overline{\boldsymbol{\mu}}} \boldsymbol{w}\|_X^2}$$

and assume that

$$\sup_{\boldsymbol{w} \in X} \frac{\|T_N^{\boldsymbol{\mu}} \boldsymbol{w} - T_N^{\overline{\boldsymbol{\mu}}} \boldsymbol{w}\|_X}{\|T_N^{\overline{\boldsymbol{\mu}}} \boldsymbol{w}\|_X} = o(\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}})$$

as $\boldsymbol{\mu} \to \overline{\boldsymbol{\mu}}$. Then

$$1 - \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) = -\frac{1}{2} \delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}} + o(|\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}}|^2) \tag{4.1}$$

and there exists $c(\overline{\boldsymbol{\mu}}, \boldsymbol{\mu})$ in $[0,1]$ such that

$$c(\overline{\boldsymbol{\mu}}, \boldsymbol{\mu}) \left[1 - \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\right] + 2 \left[\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) - \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\right] = o(|\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}}|^2). \tag{4.2}$$

In general, we can assume that $\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}}$ is of order $o(|\boldsymbol{\mu} - \overline{\boldsymbol{\mu}}|)$, hence, as for the linear case [32], $1 - \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) = o(|\boldsymbol{\mu} - \overline{\boldsymbol{\mu}}|)$ as $\boldsymbol{\mu} \to \overline{\boldsymbol{\mu}}$ and

$$0 \leq \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) - \overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \leq \frac{1}{2} \left[1 - \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\right] + o(|\overline{\boldsymbol{\mu}} - \boldsymbol{\mu}|^2).$$

Proof. From the definition of $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ and $\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$, $1 - \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})^2 = -\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}}$ and $\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})^2 \leq \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})^2 \leq \overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})^2 + \frac{1}{2}\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}} + o(|\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}}|^2)$, hence

$$\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})^2 - \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})^2 = -\frac{c(\overline{\boldsymbol{\mu}}, \boldsymbol{\mu})}{2} \delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}} + o(|\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}}|^2),$$

where $c(\overline{\boldsymbol{\mu}}, \boldsymbol{\mu}) \in [0,1]$. From the Taylor series of the square root follows (4.1) and

$$\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) - \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) = \frac{c(\overline{\boldsymbol{\mu}}, \boldsymbol{\mu})}{4} \frac{\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}}}{\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})} + o(|\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}}|^2).$$

Since $\frac{1}{1-x} = 1 + x + o(x)$, and from (4.1), $\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) - \beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) = \frac{c(\overline{\boldsymbol{\mu}}, \boldsymbol{\mu})}{4} \delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}} + o(|\delta^{\boldsymbol{\mu}}_{\overline{\boldsymbol{\mu}}}|^2)$, which proves (4.2). ∎

4.1. Local lower bound. For a fixed $\overline{\boldsymbol{\mu}}$ in $\mathcal{V}^K$, we expand $\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ as follows

$$\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) = \inf_{\boldsymbol{w}\in X} \frac{\left(T_N^{\boldsymbol{\mu}}\boldsymbol{w}, T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}\right)_X}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2} \tag{4.3}$$

$$= \inf_{\boldsymbol{w}\in X} \Big\{ 1 + \frac{1}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2} \sum_{q=1}^{Q_0} \left(\Theta_0^q(\boldsymbol{\mu}) - \Theta_0^q(\overline{\boldsymbol{\mu}})\right) a_0^q(\boldsymbol{w}, T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}) \tag{4.3a}$$

$$+ \frac{1}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2} \sum_{q=1}^{Q_1} \left(\Theta_1^q(\boldsymbol{\mu}) - \Theta_1^q(\overline{\boldsymbol{\mu}})\right) a_1^q(\boldsymbol{w}, \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}), T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}) \tag{4.3b}$$

$$+ \frac{1}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2} \sum_{q=1}^{Q_1} \Theta_1^q(\overline{\boldsymbol{\mu}})\, a_1^q(\boldsymbol{w},\, \boldsymbol{u}_N(\boldsymbol{\mu}) - \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}})\,, T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}) \Big\} \tag{4.3c}$$

$$+ \frac{1}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2} \sum_{q=1}^{Q_1} \left(\Theta_1^q(\boldsymbol{\mu}) - \Theta_1^q(\overline{\boldsymbol{\mu}})\right) a_1^q(\boldsymbol{w},\, \boldsymbol{u}_N(\mu) - \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}})\,, T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}) \Big\}. \tag{4.3d}$$

We approximate the partial derivative of $\frac{\partial \boldsymbol{u}_{\overline{N}}}{\partial \mu_p}$ at $\overline{\boldsymbol{\mu}}$ with respect to the p^{th} parameter in our reduced space $W_{\overline{N}}$:

$$da(\boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}); \overline{\boldsymbol{\mu}})(\frac{\partial \boldsymbol{u}_{\overline{N}}}{\partial \mu_p}(\overline{\boldsymbol{\mu}}), \boldsymbol{v}) = - \left[\sum_{q=1}^{Q_0} \frac{\partial \Theta_0^q(\overline{\boldsymbol{\mu}})}{\partial \mu_p} a_0^q(\boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}), \boldsymbol{v}) + \frac{1}{2} \sum_{q=1}^{Q_1} \frac{\partial \Theta_1^q(\overline{\boldsymbol{\mu}})}{\partial \mu_p} a_1^q(\boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}), \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}), \boldsymbol{v}) \right], \tag{4.4}$$

for all $\boldsymbol{v}$ in $W_{\overline{N}}$. (If f affinely depends on the parameter, we have to add $\sum_{q=1}^{Q_f} \frac{\partial \Theta^f(\overline{\boldsymbol{\mu}})}{\partial \mu_p} f^q(\boldsymbol{v})$ to the right hand side of (4.4).) We then rewrite

$$a_1^q(\boldsymbol{w}, \boldsymbol{u}_N(\boldsymbol{\mu}) - \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}), T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}) = \sum_{p=1}^{P} \kappa_p \left(\mu_p - \overline{\mu}_p\right) a_1^q(\boldsymbol{w}, \frac{\partial \boldsymbol{u}_{\overline{N}}}{\partial \mu_p}(\overline{\mu}), T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}) \;+\; a_1^q(\boldsymbol{w}, \boldsymbol{\alpha}(\boldsymbol{\mu}, \overline{\boldsymbol{\mu}}, \boldsymbol{\kappa}), T_{\overline{N}}^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}),$$

where $\boldsymbol{\kappa} \in \mathbb{R}^P$ and $\boldsymbol{\alpha}(\boldsymbol{\mu}, \overline{\boldsymbol{\mu}}, \boldsymbol{\kappa}) \in W_{N_{\max}}$ is defined as

$$\boldsymbol{\alpha}(\boldsymbol{\mu}, \overline{\boldsymbol{\mu}}, \boldsymbol{\kappa}) = \boldsymbol{u}_N(\boldsymbol{\mu}) - \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}) - \sum_{p=1}^{P} \kappa_p \left(\mu_p - \overline{\mu}_p\right) \frac{\partial \boldsymbol{u}_{\overline{N}}}{\partial \mu_p}(\overline{\boldsymbol{\mu}}).$$

As for the linear case [32], we perform a Taylor expansion with respect to $\boldsymbol{\mu}$ at $\overline{\boldsymbol{\mu}}$. The first order terms in (4.3a), (4.3b), and (4.3c) are collected: for $p = 1, ..., P$,

$$a_p^{\overline{\boldsymbol{\mu}}}(\cdot,\cdot) \equiv \sum_{q=1}^{Q_0} \frac{\partial \Theta_0^q}{\partial \mu_p}(\overline{\boldsymbol{\mu}}) a_0^q(\cdot,\cdot) + \sum_{q=1}^{Q_1} \frac{\partial \Theta_1^q}{\partial \mu_p}(\overline{\boldsymbol{\mu}}) a_1^q(\cdot, \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}), \cdot) + \sum_{q=1}^{Q_1} \Theta_1^q(\overline{\boldsymbol{\mu}}) a_1^q(\cdot, \frac{\partial \boldsymbol{u}_{\overline{N}}}{\partial \mu_p}(\overline{\boldsymbol{\mu}}), \cdot). \tag{4.5}$$

We denote the extreme eigenvalues with respect to $|||\cdot|||_{\overline{\boldsymbol{\mu}}}$ of $a_p^{\overline{\boldsymbol{\mu}}}$, a_0^q and $a_1^q(\cdot, \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}), \cdot)$ as

$$\lambda_{p,\inf}^{\overline{\boldsymbol{\mu}}} = \inf_{\boldsymbol{w}\in X} \frac{a_p^{\overline{\boldsymbol{\mu}}}(\boldsymbol{w}, T_N^{\overline{\boldsymbol{\mu}}}\boldsymbol{w})}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2}, \quad \lambda_{p,\sup}^{\overline{\boldsymbol{\mu}}} = \sup_{\boldsymbol{w}\in X} \frac{a_p^{\overline{\boldsymbol{\mu}}}(\boldsymbol{w}, T_N^{\overline{\boldsymbol{\mu}}}\boldsymbol{w})}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2}, \quad p = 1, ..., P,$$

$$\gamma_{0,q,\inf}^{\overline{\boldsymbol{\mu}}} = \inf_{\boldsymbol{w}\in X} \frac{a_0^q(\boldsymbol{w}, T_N^{\overline{\boldsymbol{\mu}}}\boldsymbol{w})}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2}, \quad \gamma_{0,q,\sup}^{\overline{\boldsymbol{\mu}}} = \sup_{\boldsymbol{w}\in X} \frac{a_0^q(\boldsymbol{w}, T_N^{\overline{\boldsymbol{\mu}}}\boldsymbol{w})}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2}, \quad q = 1, ..., Q_0,$$

$$\gamma_{1,q,\inf}^{\overline{\boldsymbol{\mu}}} = \inf_{\boldsymbol{w}\in X} \frac{a_1^q(\boldsymbol{w}, \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}), T_N^{\overline{\boldsymbol{\mu}}}\boldsymbol{w})}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2}, \quad \gamma_{1,q,\sup}^{\overline{\boldsymbol{\mu}}} = \sup_{\boldsymbol{w}\in X} \frac{a_1^q(\boldsymbol{w}, \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}), T_N^{\overline{\boldsymbol{\mu}}}\boldsymbol{w})}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2}, \quad q = 1, ..., Q_1,$$

where $\inf|\sup$ refers to two different quantities.

From our assumption (2.9) we have that

$$\frac{1}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2} \left| \sum_{q=1}^{Q_1} \Theta_1^q(\overline{\boldsymbol{\mu}}) a_1^q(\boldsymbol{w}, \boldsymbol{\alpha}(\boldsymbol{\mu}, \overline{\boldsymbol{\mu}}, \boldsymbol{\kappa}), T_N^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}), \right| \leq \rho_{X,\overline{\boldsymbol{\mu}}} \|\boldsymbol{\alpha}(\boldsymbol{\mu}, \overline{\boldsymbol{\mu}}, \boldsymbol{\kappa})\|_X. \tag{4.6}$$

Similarly, from (2.10),

$$\frac{1}{|||\boldsymbol{w}|||_{\overline{\boldsymbol{\mu}}}^2} \left| \sum_{q=1}^{Q_1} \left(\Theta_1^q(\boldsymbol{\mu}) - \Theta_1^q(\overline{\boldsymbol{\mu}})\right) a_1^q(\boldsymbol{w}, \boldsymbol{u}_N(\boldsymbol{\mu}) - \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}}), T_N^{\overline{\boldsymbol{\mu}}}\boldsymbol{w}), \right|$$
$$\leq \max_{q=1,\ldots,Q_1} \frac{|\Theta_1^q(\boldsymbol{\mu}) - \Theta_1^q(\overline{\boldsymbol{\mu}})|}{|\Theta_1^q(\overline{\boldsymbol{\mu}})|} \rho_{X,\overline{\boldsymbol{\mu}}} \|\boldsymbol{u}_N(\boldsymbol{\mu}) - \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}})\|_X.$$

As a result, our bound to (4.3) is defined, for any $\boldsymbol{\kappa}$ in $\mathbb{R}^P$, as

$$\begin{aligned}
\overline{\beta}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \geq & \overline{B}_{\overline{\boldsymbol{\mu}}}^{LB}(\boldsymbol{\mu}, \boldsymbol{\kappa}) \equiv 1 + \sum_{p=1}^{P} \min_{\lambda_p = \lambda_{p,\,\inf|\sup}^{\overline{\boldsymbol{\mu}}}} \kappa_p(\mu_p - \overline{\mu}_p)\lambda_p \\
& + \sum_{q=1}^{Q_0} \min_{\gamma_q = \gamma_{0,q,\,\inf|\sup}^{\overline{\boldsymbol{\mu}}}} \left(\Theta_0^q(\boldsymbol{\mu}) - \Theta_0^q(\overline{\boldsymbol{\mu}}) - \sum_{p=1}^{P} \kappa_p(\mu_p - \overline{\mu}_p) \frac{\partial \Theta_0^q}{\partial \mu_p}(\overline{\boldsymbol{\mu}}) \right) \gamma_q \\
& + \sum_{q=1}^{Q_1} \min_{\gamma_q = \gamma_{1,q,\,\inf|\sup}^{\overline{\boldsymbol{\mu}}}} \left(\Theta_1^q(\boldsymbol{\mu}) - \Theta_1^q(\overline{\boldsymbol{\mu}}) - \sum_{p=1}^{P} \kappa_p(\mu_p - \overline{\mu}_p) \frac{\partial \Theta_1^q}{\partial \mu_p}(\overline{\boldsymbol{\mu}}) \right) \gamma_q \\
& - \rho_{X,\overline{\boldsymbol{\mu}}} \left(\|\boldsymbol{\alpha}(\boldsymbol{\mu}, \overline{\boldsymbol{\mu}}, \boldsymbol{\kappa})\|_X + \max_{q=1,\ldots,Q_1} \frac{|\Theta_1^q(\boldsymbol{\mu}) - \Theta_1^q(\overline{\boldsymbol{\mu}})|}{|\Theta_1^q(\overline{\boldsymbol{\mu}})|} \|\boldsymbol{u}_N(\boldsymbol{\mu}) - \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}})\|_X \right),
\end{aligned} \tag{4.7}$$

where we denote the minimum of a function g over two values as

$$\min_{\lambda = \lambda_{\inf|\sup}} g(\lambda) \equiv \min\{g(\lambda_{\inf}), g(\lambda_{\sup})\}.$$

Since (4.7) is valid for any $\boldsymbol{\kappa} \in \mathbb{R}^P$, we define our inf-sup lower bound as

$$\overline{B}_{\overline{\boldsymbol{\mu}}}^{LB}(\boldsymbol{\mu}) \equiv \max_{\boldsymbol{\kappa}\in\mathbb{R}^P} \overline{B}_{\overline{\boldsymbol{\mu}}}^{LB}(\boldsymbol{\mu}, \boldsymbol{\kappa}), \tag{4.8}$$

4.2. $\boldsymbol{\kappa}$ optimization. Ideally, for a given $\boldsymbol{\mu}$, we have to find the value of $\boldsymbol{\kappa}$ that maximizes $\overline{B}^{LB}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu},\boldsymbol{\kappa})$. The optimization of $\boldsymbol{\kappa}$ relies on quadratic programming. We propose a simpler heuristic method: the bound involving the eigenvalues $\lambda^{\overline{\boldsymbol{\mu}}}_{p,\,\inf\,|\,\sup}$ is sharper than (4.6), therefore we compute $\boldsymbol{\kappa}$ such that $\|\boldsymbol{\alpha}(\boldsymbol{\mu},\overline{\boldsymbol{\mu}},\boldsymbol{\kappa})\|_X$ is the smallest. Moreover, in our algorithms we will require an X-orthonormal basis for $W_{N_{\max}}$, hence this minimization is achieved by the least-square solution $\boldsymbol{\kappa}^*$ that satisfies

$$\begin{bmatrix} \vdots \\ (\mu_p-\overline{\mu}_p)\partial_p U_{\overline{N},j}(\overline{\boldsymbol{\mu}}) \\ \vdots \end{bmatrix}_{j,p}^{T} \begin{bmatrix} \vdots \\ (\mu_p-\overline{\mu}_p)\partial_p U_{\overline{N},j}(\overline{\boldsymbol{\mu}}) \\ \vdots \end{bmatrix}_{j,p} \boldsymbol{\kappa}^* \\ = \begin{bmatrix} \vdots \\ (\mu_p-\overline{\mu}_p)\partial_p U_{\overline{N},j}(\overline{\boldsymbol{\mu}}) \\ \vdots \end{bmatrix}_{j,p}^{T} \begin{bmatrix} \vdots \\ U_{N,j}(\boldsymbol{\mu}) - U_{\overline{N},j}(\overline{\boldsymbol{\mu}}) \\ \vdots \end{bmatrix}_{j}, \quad (4.9)$$

where $U_{N,j}(\boldsymbol{\mu})$, $U_{\overline{N},j}(\overline{\boldsymbol{\mu}})$, and $\partial_p U_{\overline{N},j}(\overline{\boldsymbol{\mu}})$ are the j^{th} components of $\boldsymbol{u}_N(\boldsymbol{\mu})$, $\boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}})$, and $\frac{\partial \boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}})}{\partial \mu_p}$ with respect to the basis of $W_{N_{\max}}$.

5. Computational strategy.

5.1. Offline algorithm. Our offline algorithm includes the computation of the reduced basis ingredients — optimal selection of the basis, computation of matrices and vector in the reduced space (c.f., e.g., [32]) — and of the a posteriori error estimation ingredients — selection of $\mathcal{V}^K=\{\overline{\boldsymbol{\mu}}_1,...,\overline{\boldsymbol{\mu}}_K\}$, solution of the eigenvalue problems, and computation of the Sobolev embedding constants. Because of the presence of the reduced basis approximations $\boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}})$ in the definition of the natural norms, these two stages are dependent from each other. (In the linear case, cf. [32], these steps are ordered: first to compute the a posteriori error estimation ingredients, then to select the reduced basis.)

In the first step, we provide surrogates $\overline{B}^{LB}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ and $\rho_{\overline{\boldsymbol{\mu}}}(\overline{\boldsymbol{\mu}})$ in order to compute an error bound extrapolation. We expect that $\overline{B}^{LB}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ is of order one, so we set its surrogate to, say, 0.2; to compute a surrogate $\rho_{\overline{\boldsymbol{\mu}}}(\overline{\boldsymbol{\mu}})$, we replace $\boldsymbol{u}_{\overline{N}}(\overline{\boldsymbol{\mu}})$ by $\boldsymbol{u}(\overline{\boldsymbol{\mu}})$ in (2.4) (hence in (2.7)) — in other words we do not need a reduced basis. As a result we have an efficient — online fast — tool to approximately compute the error for a given $\boldsymbol{\mu}$ in $\mathcal{D}$.

In the second step we compute the eigenvalues and the Sobolev embedding constants that we need in our inf-sup lower bound construction.

We hence propose the following algorithm

1. Manually set K, select a representative set $\mathcal{V}^K=\{\overline{\boldsymbol{\mu}}_1,...,\overline{\boldsymbol{\mu}}_K\}$, and compute surrogates for each $\rho_{\overline{\boldsymbol{\mu}}}$;
2. start an optimal search algorithm: for a large random[5] set of $\boldsymbol{\mu}$'s
    (a) solve the reduced basis problem with the existing basis;
    (b) select the "nearest" $\overline{\boldsymbol{\mu}}$ in $\mathcal{V}^K$ and compute $\epsilon_N(\mu)$, $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$, and $\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ by replacing $\rho_{\overline{\boldsymbol{\mu}}}$ and $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ by their surrogates;
    (c) select the optimal $\boldsymbol{\mu}$ as follows

[5] This can be replaced by a deterministic search over a large set of sample points.

i. if for some $\boldsymbol{\mu}$'s, (2a) did not converge, select the "furthest" from the previously selected ones;
ii. if for some $\boldsymbol{\mu}$'s, $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ is bigger than 1, select the one with the largest $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$;
iii. otherwise select the one with the largest $\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$.

3. solve the finite element problem (2.1) for the selected $\boldsymbol{\mu}$;
4. increase N and enrich the reduced basis space[6] W_N with $\boldsymbol{u}(\boldsymbol{\mu})$;[7]
5. if the maximum number of basis wished or the tolerance requested for $\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ are achieved, go to 6, otherwise go to 2;
6. set $\overline{N}$ to the reached number of basis functions and compute all the ingredients for the computation of $\overline{B}^{LB}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$;
7. start a random process: for a large set of $\boldsymbol{\mu}$'s check that the reduced basis problem can be solved, that $\overline{B}^{LB}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ is positive for at least one $\overline{\boldsymbol{\mu}}$ and that $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \leq 1$ for this $\overline{\boldsymbol{\mu}}$.
   If this is ok, then all the online components are ready and we set $N_{\max} = N$, otherwise either add some $\overline{\boldsymbol{\mu}}$'s to $\mathcal{V}^K$ and go to 6; or go to 2 and replace $\beta_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ by $\overline{B}^{LB}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ instead of by 0.2.

Steps 2a. and 2b. are online fast. Moreover, in our parallel offline code, each processor performs the optimal search on a different random set. As a result, we can explore our parameter space very fast.

To ensure that we are on a selected solution branch, in the finite element resolution we perform a homotopy procedure in $\boldsymbol{\mu}$ starting from an already computed solution at a nearby parameter; in the reduced basis we simply start from a solution at nearby parameter used to generate the W_N. (In some applications, it may be necessary to perform homotopy also in the reduced basis context.)

Note that we can admit small errors in the resolution of (2.1); indeed, we need approximations to $\boldsymbol{u}(\boldsymbol{\mu})$ at selected parameter points to build our reduced basis. However, the error bound is given with respect to the exact truth solution $\boldsymbol{u}(\boldsymbol{\mu})$.

The reader is referred to [33] and [32] for the exact offline procedure to compute the reduced basis and a posteriori error estimation components.

5.2. Online algorithm. The online procedure, for given $\boldsymbol{\mu} \in \mathcal{D}$ and $N \leq N_{\max}$, reads

I. solve the reduced basis problem (1.3)[8] and compute $s_N(\boldsymbol{\mu}) = \ell(\boldsymbol{u}(\boldsymbol{\mu}))$;
II. compute $\epsilon_N(\boldsymbol{\mu})$;
III. order $\mathcal{V}^K$ in increasing distance[9] from $\boldsymbol{\mu}$. Compute $\overline{B}^{LB}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ and $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ until we find $\overline{\boldsymbol{\mu}}$ in $\mathcal{V}$ such that $\overline{B}^{LB}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) > 0$ and $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \leq 1$. We set $\mathcal{I}^K(\boldsymbol{\mu}) = \overline{\boldsymbol{\mu}}$;
IV. compute $\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ and $\Delta^s_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$.

If in step III. we do not find any $\overline{\boldsymbol{\mu}}$, our method fails: we can not provide an error bound nor even existence; if possible, N must be increased, otherwise our basis is not rich enough.

[6]This step involves Gramm-Schmidt orthogonalization and the computation of the matrix- and scalar-products necessary for the reduced basis matrices and the online computation of the Y-dual norm of the residual (cf. [23, 32]).

[7]In our model problems, $\boldsymbol{u}(\boldsymbol{\mu})$ is split into its physical components and a stabilizing supremizer is computed; the reduced space is hence enriched by four functions (cf. Section 6.2).

[8]We solve this by a Newton algorithm with initial guess a solution at a nearby parameter used to generate the reduced basis.

[9]The distance function in $\mathcal{D}$ is arbitrary.

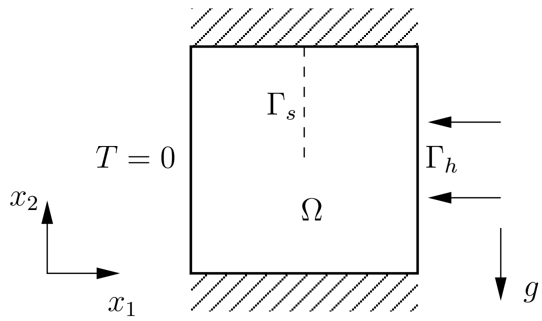

Fig. 1. *Closed cavity. On the left the temperature is constant, on the right the heat flux is constant, and the top and bottom are insulated.*

The computational effort is independent of $\mathcal{N}$: steps I. and II. have dominant complexities $N^3 + Q_1 N^2$ times the number of Newton iterations required and $Q_1^2 N^4 + Q_0 Q_1 N^3$, respectively. The complexity of steps III. and IV. is dominated by that of step II.

6. Application: steady natural convection. We are interested in solving a natural convection problem in a two-dimensional square cavity under vertical gravity (Figure 1), with constant temperature on one side, heated on the opposite side, insulated at the top and bottom, and Prandtl number $\mathrm{Pr} = 1$ (dissipation to conduction ratio). We consider one parameter $\boldsymbol{\mu} = (\frac{1}{\sqrt{\mathrm{Gr}}})$, where Gr is the Grashof number, which is defined as $(\beta \Delta T g L^3)/\nu^2$ (g is the gravity constant, β is the thermal expansion coefficient, ΔT is the temperature difference, L is the length scale, and ν is the kinematic viscosity).

We impose homogeneous Dirichlet boundary conditions for the velocity (u_1, u_2), while for the temperature θ we impose homogeneous Dirichlet boundary conditions on the left side, homogeneous Neumann boundary conditions on the top and the bottom, and constant unitary Neumann boundary conditions on the heated side Γ_h ($\partial_n \theta = 1$ on Γ_h). We scale the pressure such that $\int_\Omega p = 0$.

Our domain $\Omega \subset \mathbb{R}^2$ is a closed square cavity with unitary sides. We want to solve (with summation over repeated indices and partial derivatives with respect to x_1 and x_2)

$$\begin{aligned} u_j \partial_j u_1 &= -\partial_1 p + \frac{1}{\sqrt{\mathrm{Gr}}} \partial^2_{jj} u_1, \\ u_j \partial_j u_2 &= -\partial_2 p + \theta + \frac{1}{\sqrt{\mathrm{Gr}}} \partial^2_{jj} u_2, \\ \partial_i u_i &= 0, \\ u_j \partial_j \theta &= \frac{1}{\sqrt{\mathrm{Gr}}\,\mathrm{Pr}} \partial^2_{jj} \theta \end{aligned}$$

in Ω.

Our outputs of interest are the inverse of the Nusselt number and the flux through the middle half section Γ_s parallel to Γ_h,

$$s_{(1)}(\boldsymbol{\mu}) = \frac{1}{|\Gamma_h|} \int_{\Gamma_h} \theta(\boldsymbol{\mu}), \qquad s_{(2)}(\boldsymbol{\mu}) = \frac{1}{|\Gamma_s|} \int_{\Gamma_s} u_1(\boldsymbol{\mu}).$$

6.1. Affine decomposition. Our reference finite element space X is a subset of $H_0^1(\Omega)\times H_0^1(\Omega)\times H^1(\Omega)\times L^2(\Omega)\times\mathbb{R}$. We denote elements in X as $\boldsymbol{u}=(u_1,u_2,\theta,p,\lambda)$, $\boldsymbol{v}=(v_1,v_2,\chi,q,\gamma)$, and $\boldsymbol{z}=(z_1,z_2,\zeta,\cdots)$.

Recalling (2.2), a decomposes into $Q_0=2$ bilinear and $Q_1=1$ trilinear components,

$$\begin{aligned}
a_0^1(\boldsymbol{u},\boldsymbol{v}) &= -\int_\Omega p\partial_i v_i - \int_\Omega q\partial_i u_i + \lambda\int_\Omega p + \gamma\int_\Omega q - \int_\Omega \theta v_2 && \Theta_0^1(\boldsymbol{\mu})=1\\
a_0^2(\boldsymbol{u},\boldsymbol{v}) &= \int_\Omega \partial_i u_j\partial_i v_j + \int_\Omega \partial_i\theta\partial_i\chi && \Theta_0^2(\boldsymbol{\mu})=\mu_1\\
a_1^1(\boldsymbol{u},\boldsymbol{z},\boldsymbol{v}) &= \int_\Omega u_j\partial_i z_i v_j + \int_\Omega z_j\partial_i u_i v_j + \int_\Omega z_i\partial_i\theta\,\chi + \int_\Omega u_i\partial_i\zeta\,\chi && \Theta_1^1(\boldsymbol{\mu})=1.
\end{aligned}$$

The right hand side, which linearly depends on μ_1, and the (linear) outputs take the form

$$\begin{aligned}
f^1(\boldsymbol{v}) &= \int_{\Gamma_h}\chi && \Theta_f^1(\boldsymbol{\mu})=\mu_1\\
\ell_{(1),0}^1(\boldsymbol{v}) &= \int_{\Gamma_h}\chi && \Theta_{\ell,(1)}^1(\boldsymbol{\mu})=1\\
\ell_{(2),0}^1(\boldsymbol{v}) &= \int_{\Gamma_s}v_1 && \Theta_{\ell,(2)}^1(\boldsymbol{\mu})=1.
\end{aligned}$$

6.2. Reduced basis and supremizer. To achieve faster convergence (cf. [17, 33]), in step 4 of the algorithm in Section 5.1, the solution $\boldsymbol{u}(\boldsymbol{\mu})=(u_1,u_2,\chi,p,\lambda)$ is split into its physical components. To ensure solvability of the reduced system we preliminary add the constant pressure $(0,0,0,1,0)$ and Lagrange multiplier $(0,0,0,0,1)$ to our reduced space. To ensure stability, to each new added pressure mode, we add a supremizer (cf. [26, 27]) $\sigma(\boldsymbol{u}(\boldsymbol{\mu}))\in X$, defined by a null temperature, pressure, and Lagrange multiplier and

$$(\boldsymbol{\sigma}(\boldsymbol{u}(\boldsymbol{\mu})),\boldsymbol{v})_X = \int_\Omega p\partial_i v_i \quad \forall\boldsymbol{v}\in X.$$

(Note: there are alternative definition of the supremizer, cf. [26, 27].)

Therefore step 4 is replaced by

4bis. Add the three basis functions $(u_1,u_2,0,0,0)$, $(0,0,\chi,0,0)$, $(0,0,0,p,0)$ as well as the supremizer $\boldsymbol{\sigma}(\boldsymbol{u}(\boldsymbol{\mu}))$ to our reduced basis space;

6.3. Sobolev embedding constant. We define a semi-norm $|\cdot|_4$ as the restriction of the L^4-norm to the velocity and temperature fields,

$$|\boldsymbol{v}|_4^4 = \int_\Omega \left(v_1^2+v_2^2+\chi^2\right)^2.$$

Proposition For all $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{z}$ in X,

$$\left|a_1^1(\boldsymbol{u},\boldsymbol{z},\boldsymbol{v})\right| \le 2|\boldsymbol{u}|_4\,\|\boldsymbol{z}\|_X\,|\boldsymbol{v}|_4 + |\boldsymbol{u}|_4\,|\boldsymbol{z}|_4\,\|\boldsymbol{v}\|_X.$$

Proof. We denote $u_3 = \theta$, $v_3 = \chi$, and $z_3 = \zeta$; we then have

$$\begin{aligned}
\int_\Omega \sum_{i=1}^{2}\sum_{j=1}^{3} u_i \partial_i z_j v_j \le \int_\Omega \left(\sum_{j=1}^{3} v_j^2\right)^{\frac{1}{2}} \left(\sum_{j=1}^{3}\left(\sum_{i=1}^{2} u_i \partial_i z_j\right)^2\right)^{\frac{1}{2}} \\
\le \left[\int_\Omega \left(\sum_{j=1}^{3} v_j^2\right)^2\right]^{\frac{1}{4}} \left[\int_\Omega \left(\sum_{j=1}^{3}\left(\sum_{i=1}^{2} u_i \partial_i z_j\right)^2\right)^{\frac{2}{3}}\right]^{\frac{3}{4}} \\
\le |\boldsymbol{v}|_4 \left[\int_\Omega \left(\sum_{i=1}^{2} u_i^2\right)^{\frac{2}{3}} \left(\sum_{j=1}^{3}\sum_{i=1}^{2} (\partial_i z_j)^2\right)^{\frac{2}{3}}\right]^{\frac{3}{4}} \\
\le |\boldsymbol{v}|_4 \left[\int_\Omega \left(\sum_{i=1}^{2} u_i^2\right)^2\right]^{\frac{1}{4}} \left[\int_\Omega \sum_{j=1}^{3}\sum_{i=1}^{2} (\partial_i z_j)^2\right]^{\frac{1}{2}} \\
\le |\boldsymbol{v}|_4 \, \|(u_1,u_2)\|_{L^4(\Omega)^2} \, |(z_1,z_2,z_3)|_{H^1(\Omega)^3}. \qquad (6.1)
\end{aligned}$$

From

$$\sum_{i=1}^{2}\sum_{j=1}^{3} \partial_i (u_j z_i) = \sum_{i=1}^{2}\sum_{j=1}^{3} u_j \partial_i z_i + \sum_{i=1}^{2}\sum_{j=1}^{3} z_i \partial_i u_j,$$

integration by parts, and $z_1 = z_2 = 0$ on $\partial\Omega$ we have

$$\int_\Omega \sum_{i=1}^{2}\sum_{j=1}^{3} z_i \partial_i u_j v_j = -\int_\Omega \sum_{i=1}^{2}\sum_{j=1}^{3} u_j z_i \partial_i v_j - \int_\Omega \sum_{i=1}^{2}\sum_{j=1}^{3} u_j \partial_i z_i v_j.$$

We bound the first term on the right hand side by applying (6.1) and the second one by

$$\begin{aligned}
\int_\Omega \sum_{i=1}^{2}\sum_{j=1}^{3} u_j \partial_i z_i v_j \le |\boldsymbol{v}|_4 \left[\int_\Omega \left(\sum_{j=1}^{3} u_j^2 \sum_{i=1}^{2} (\partial_i z_i)^2\right)^{\frac{2}{3}}\right]^{\frac{3}{4}} \\
\le |\boldsymbol{v}|_4 \left[\int_\Omega \left(\sum_{j=1}^{3} u_j^2\right)^2\right]^{\frac{1}{4}} \left[\int_\Omega \sum_{i=1}^{2} (\partial_i z_i)^2\right]^{\frac{1}{2}} = |\boldsymbol{v}|_4 \, |\boldsymbol{u}|_4 \, |(z_1,z_2)|_{H^1_{\mathrm{div}}(\Omega)^2}.
\end{aligned}$$

$\|(u_1,u_2)\|_{L^4(\Omega)^2} \le |\boldsymbol{u}|_4$, $|(z_1,z_2,z_3)|_{H^1(\Omega)^3} < \|\boldsymbol{z}\|_X$, $|(z_1,z_2)|_{H^1_{\mathrm{div}}(\Omega)^2} < \|\boldsymbol{z}\|_X$, and the definition of a_1^1 complete the proof. ☐

From the Sobolev Embedding Theorem we know that there exists a constant ρ_X such that

$$|\boldsymbol{v}|_4 \le \rho_X \|\boldsymbol{v}\|_X \text{ for all } \boldsymbol{v} \in X. \qquad (6.2)$$

and from the equivalence of the norms $\|\cdot\|_X$ and $|||\cdot|||_{\overline{\boldsymbol{\mu}}}$, there exists a constant $\rho_{\overline{\boldsymbol{\mu}}} \le \frac{\rho_X}{\beta(\overline{\boldsymbol{\mu}})}$ such that

$$|\boldsymbol{v}|_4 \le \rho_{\overline{\boldsymbol{\mu}}} |||\boldsymbol{v}|||_{\overline{\boldsymbol{\mu}}} \text{ for all } \boldsymbol{v} \in X. \qquad (6.3)$$

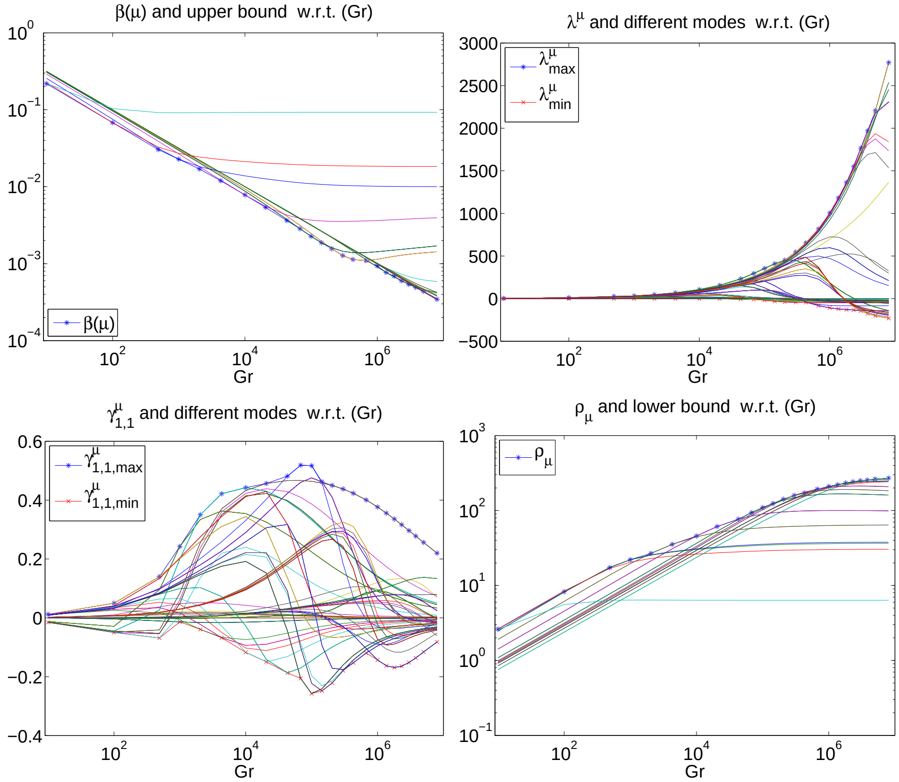

FIG. 2. *Upper left: $\beta(\overline{\boldsymbol{\mu}})$ and upper bounds given by the Rayleigh quotients of the minimizing modes. U.r. and L.l.: $\lambda_1^{\overline{\boldsymbol{\mu}}}$, $\gamma_{1,1}^{\overline{\boldsymbol{\mu}}}$ respectively, and Rayleigh quotients of several maximizing modes (those relative to $\lambda^{\overline{\boldsymbol{\mu}}}$, $\gamma_{0,1}^{\overline{\boldsymbol{\mu}}}$, $\gamma_{0,2}^{\overline{\boldsymbol{\mu}}}$, and $\gamma_{1,1}^{\overline{\boldsymbol{\mu}}}$). L.r.: $\rho_{\overline{\boldsymbol{\mu}}}$ and lower bounds given by the quotients of our L^4 semi-norm and natural norm of the maximizing modes. The x-axis represents the Grashof number relative to $\overline{\boldsymbol{\mu}}$.*

With

$$\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \equiv \frac{\rho_{\overline{\boldsymbol{\mu}}}\rho_X}{\beta(\overline{\boldsymbol{\mu}})} + 2\rho_{\overline{\boldsymbol{\mu}}}^2 \quad \text{and} \quad \rho_{X,\overline{\boldsymbol{\mu}}} \equiv 3\rho_{\overline{\boldsymbol{\mu}}}\rho_X$$

the assumptions (2.8), (2.9), and (2.10) are satisfied.

6.4. Computation of the eigenvalues. The computation of the inf-sup constants $\beta(\overline{\boldsymbol{\mu}})$ and of the extreme eigenvalues $\lambda_{p,\inf|\sup}^{\overline{\boldsymbol{\mu}}}$, $p = 1, ..., P$, $\gamma_{i,q,\inf|\sup}^{\overline{\boldsymbol{\mu}}}$, $q = 1, ..., Q_i$, $i = 0, 1$, heavily depends on the initial guesses. We start by computing $\beta(\overline{\boldsymbol{\mu}})$ for all $\overline{\boldsymbol{\mu}}$ in $\mathcal{V}^K$ from low to high Grashof numbers by inverse iterations. We store all the modes and at each $\overline{\boldsymbol{\mu}}$ we select the initial guess as the stored mode with the smallest Rayleigh quotient.

Then we compute the remaining eigenvalues, also by selecting among the stored modes and by storing the new modes found. In this case, we use a locally optimized block preconditioned gradient method [10]. Figure 2 shows $\beta(\overline{\boldsymbol{\mu}})$, $\lambda_{1,\inf|\sup}^{\overline{\boldsymbol{\mu}}}$, and $\gamma_{1,1,\inf|\sup}^{\overline{\boldsymbol{\mu}}}$ and the comparison with the stored modes; some maximizing modes are almost identical for different values of $\overline{\boldsymbol{\mu}}$.

6.5. Computation of the Sobolev embedding constants. We have developed an algorithm, [19], for the computation of the Sobolev embedding constants ρ_X and $\rho_{\overline{\boldsymbol{\mu}}}$ in (6.2) and (6.3).

Let $\boldsymbol{u}^* \in X$, $\|\boldsymbol{u}^*\| = 1$, be the (possibly non unique) supremizer of

$$\rho = \sup_{\boldsymbol{v} \in X} \frac{|\boldsymbol{v}|_4}{\|\boldsymbol{v}\|}, \tag{6.4}$$

where $\|\cdot\|$ denotes the X or the natural norms, and let $z : X \to L^2(\Omega)$,

$$z(\boldsymbol{v}) = \frac{v_1^2 + v_2^2 + \chi^2}{|\boldsymbol{v}|_4^2}.$$

Note that $\|z(\boldsymbol{v})\|_{L^2(\Omega)} = 1$ for all $\boldsymbol{v}$ in X.

For a non-negative function $z \in L^2(\Omega)$, we introduce the eigenproblem: find $\boldsymbol{u} = (u_1, u_2, \theta, \cdots) \in X$ and $\lambda \in \mathbb{R}$ such that $\|\boldsymbol{u}\| = 1$ and

$$\int_\Omega z\,(u_1 v_1 + u_2 v_2 + \theta\chi) = \lambda\,(\boldsymbol{u}, \boldsymbol{v}),\ \forall \boldsymbol{v} = (v_1, v_2, \chi, \cdots) \in X,$$

where $(\cdot,\cdot)$ is the X or natural scalar products. We denote by $\lambda_{\max}(z)$ and $\boldsymbol{u}_{\max}(z)$ the largest eigenvalue and associated eigenfunction. Note: we choose the eigenfunctions always in the same half-space.

We observe that $\lambda_{\max}(z(\boldsymbol{u}^*)) = \rho^2$ and $\boldsymbol{u}_{\max}(z(\boldsymbol{u}^*)) = \boldsymbol{u}^*$, which suggests the following fixed point algorithm:

Let $\boldsymbol{u}^0 \in X$ be our initial guess; for $k \geq 0$, define

$$\boldsymbol{u}^{k+1} = \boldsymbol{u}_{\max}\left(z(\boldsymbol{u}^k)\right), \lambda^{k+1} = \lambda_{\max}\left(z(\boldsymbol{u}^k)\right).$$

A fixed point of this algorithm is not necessarily the supremizer of (6.4), but it is at least a local supremizer. In fact

$$\lambda^{k+1} - \lambda^k = |\boldsymbol{u}^{k-1}|_4^2 \int_\Omega (z(\boldsymbol{u}^{k-1}) - z(\boldsymbol{u}^{k-2}))z(\boldsymbol{u}^{k-1}) + o(\|z(\boldsymbol{u}^{k-1}) - z(\boldsymbol{u}^{k-2})\|_{L^2(\Omega)}^2)$$

and, since $\|z(\boldsymbol{v})\|_{L^2(\Omega)} = 1$,

$$\int_\Omega (z(\boldsymbol{u}^{k-1}) - z(\boldsymbol{u}^{k-2}))z(\boldsymbol{u}^{k-1}) \geq 0;$$

if $z(\boldsymbol{u}^{k-1})$ and $z(\boldsymbol{u}^{k-2})$ are not parallel (i.e, not equal), this integral is strictly positive.

The most critical point of this algorithm is hence the initial guess. To provide a robust initial guess, we first look for the maximizing mode in the low dimensional subspace generated by the eigenmodes of $T_N^{\overline{\boldsymbol{\mu}}}$ and the already computed supremizers of (6.4). We apply our algorithm in this space and we select our initial guesses among the basis elements (i.e., among the eigenmodes of $T_N^{\overline{\boldsymbol{\mu}}}$ and the supremizers of (6.4)). This process depends on the order that we choose in $\mathcal{V}^K$; to relax this dependency, when the computations are completed for all $\overline{\boldsymbol{\mu}}$ in $\mathcal{V}^K$, we store the modes and restart the computation of all the embedding constants. The last picture in Figure 2 shows the embedding constants and the comparison with the quotient (6.4) of the stored modes. In Figure 3 note how for high Grashof number, $\rho_{\overline{\boldsymbol{\mu}}}$ deflates significantly from its upper bound $\rho_X/\beta(\boldsymbol{\mu})$.

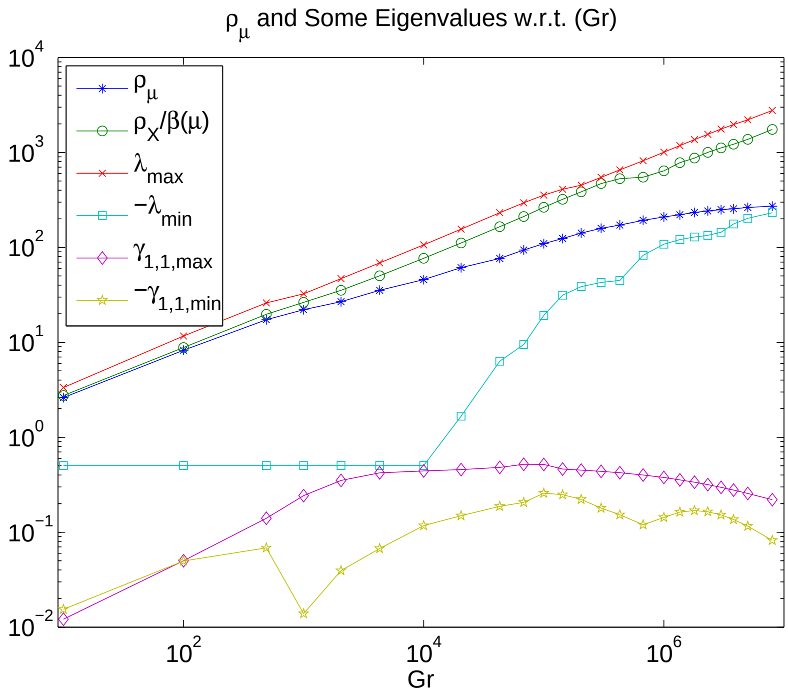

FIG. 3. *Plot of* $\rho_{\overline{\boldsymbol{\mu}}}$, $\rho_X/\beta(\overline{\boldsymbol{\mu}})$, $\lambda^{\overline{\boldsymbol{\mu}}}_{1,\,\inf|\,\sup}$, *and* $\gamma^{\overline{\boldsymbol{\mu}}}_{1,1,\,\inf|\,\sup}$. *Note that* $\rho_{\overline{\boldsymbol{\mu}}} \le \rho_X/\beta(\overline{\boldsymbol{\mu}})$ *and that they are similar for low Grashof number, but that at higher Grashof number the difference is almost of one order of magnitude. This encourages to use the mixed formulation of the Brezzi–Rappaz–Raviart theory.*

N	$\max \tau_N$	$\max \Delta_N$	$\max \Delta^s_{(1),N}$	$\max \Delta^s_{(2),N}$	RBcpu	FEcpu
$2{+}4{\cdot}10$	∞	$4.05{\cdot}10^{-4}$	$5.60{\cdot}10^{-3}$	$4.57{\cdot}10^{-4}$	15+19 ms	
$2{+}4{\cdot}12$	∞	$3.60{\cdot}10^{-4}$	$4.81{\cdot}10^{-3}$	$3.93{\cdot}10^{-4}$	18+35 ms	
$2{+}4{\cdot}14$	2.26	$2.94{\cdot}10^{-5}$	$6.92{\cdot}10^{-4}$	$9.24{\cdot}10^{-5}$	22+66 ms	
$2{+}4{\cdot}16$	0.128	$6.32{\cdot}10^{-6}$	$3.08{\cdot}10^{-4}$	$2.98{\cdot}10^{-5}$	25+108 ms	
$2{+}4{\cdot}18$	0.113	$3.16{\cdot}10^{-6}$	$1.154{\cdot}10^{-4}$	$1.49{\cdot}10^{-5}$	27+157 ms	
$2{+}4{\cdot}19$	0.122	$5.12{\cdot}10^{-7}$	$3.34{\cdot}10^{-5}$	$8.20{\cdot}10^{-6}$	29+188 ms	16×19s

TABLE 1

Convergence rate and CPU times's comparison over 1000 random samples (different for different N). $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) = \infty$ means that for some parameters, our inf-sup lower bound is negative. The maximal error bounds are taken over the samples whom we can provide existence and unicity for. The RBcpu counts the mean CPU time for solving the reduced basis plus the time to compute the error bound (dual norm of the residual and inf-sup lower bounds); the FEcpu counts the mean CPU time for computing the finite element solution with a dual norm of the residual smaller than 10^{-10} *and the reduce basis solution as initial guess.*

7. Numerical results.

7.1. Finite element solver. We implemented a parallel finite element solver that uses the Trilinos[10] library, in particular Epetra as MPI interface, Amesos [31, 30] as dense direct solver, Aztec00 as linear iterative solver with domain decomposition

[10] `http://software.sandia.gov/trilinos/`

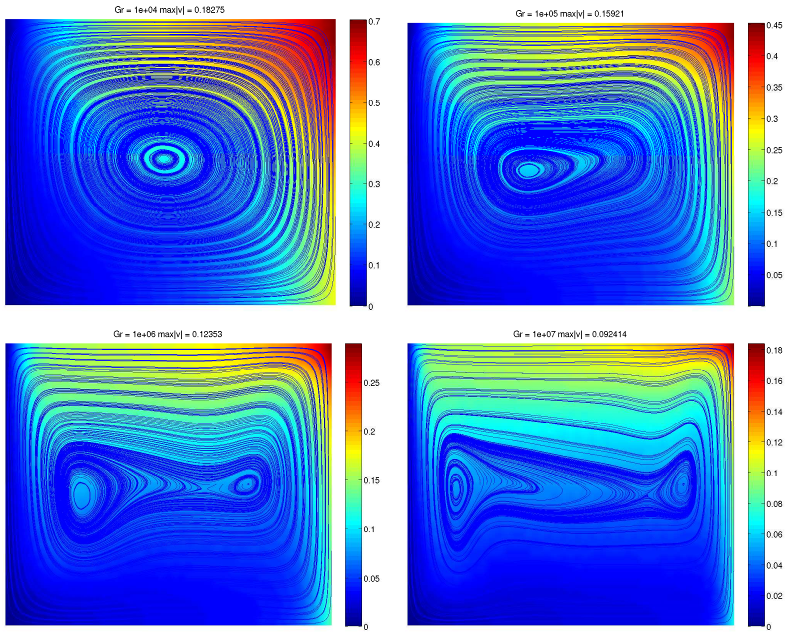

FIG. 4. *Streamlines at Grashof number* 10^4, 10^5, 10^6 *and* 10^7.

preconditioners from IFPACK [29], and Anasazi as eigenvalue solver, in particular LOBPCG [10].

We exploited the PSC[11] resources thanks to the grant number ASC040010P.

We used Taylor–Hood P1/P2 finite elements for a total of 38000 degrees of freedom. We modified a preconditioner proposed by Elman et al. [5]: at a point $\boldsymbol{u} \in X$, we define P^0 as $da(\boldsymbol{u}, \boldsymbol{\mu})$ where the divergence operator is replaced by the pressure mass matrix on the pressure block diagonal. We then construct a one-level Schwarz preconditioner P [29, 28] to P^0 and perform a local LU factorization. We solve the Jacobian system with restarted PGMRES(500) with preconditioner P. The resulting operator P, which depends on $\boldsymbol{u}$ and $\boldsymbol{\mu}$, is fast, distributed, and is an effective preconditioner for our problem up to $\mathrm{Gr} = 10^7$.

We are interested in one solution branch, therefore we use homotopy with respect to Grashof number when solving the finite element problem. In contrast, in our example, the reduced basis problem does not need homotopy: our initial guess is a known solution for a nearby parameter (a solution that has generated our reduced space).

In Figures 4 and 5 we present the streamlines and the profiles for increasing Grashof number. We note the development of two vortices and a boundary layer against the wall. We also show the behavior of the Reynolds number (given by $\sqrt{\mathrm{Gr}} \cdot \max |(u_1, u_1)|$) with respect to the Grashof number (Figure 6), and our outputs (Figure 7, computed by the reduced basis method).

The offline computation, i.e., the computation of the reduced basis and the inf-sup lower bound ingredients, required about 36 wall-time hours on a 16 CPUs cluster.

[11] `http://www.psc.edu`

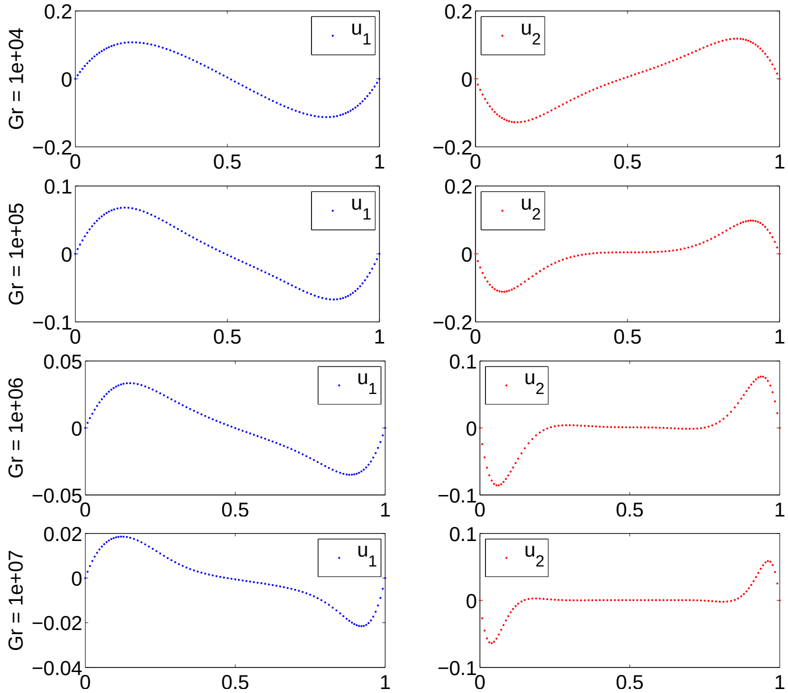

FIG. 5. *Profiles at Grashof number* 10^4, 10^5, 10^6 *and* 10^7. *On the left the vertical central section, on right the horizontal one.*

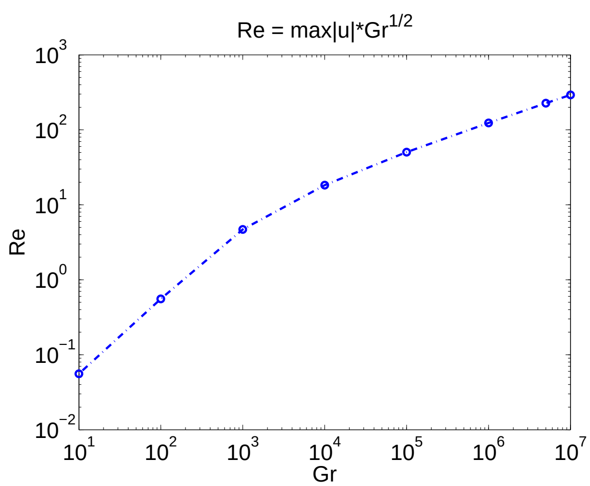

FIG. 6. *Approximative relation between Grashof and Reynolds numbers.*

7.2. Reduced basis resolution and error bounds. Table 1 compares the CPU-time needed by an online reduced basis resolution and a finite element resolution. In the latter, we optimize the initial guess by taking the reduced basis solution, such that the X-dual norm of the residual is already below 10^{-6} and two Newton iterations are usually enough to reduce it to 10^{-10}. Although this optimization is very generous with the finite element code, the wall-time is about two orders of magnitude smaller

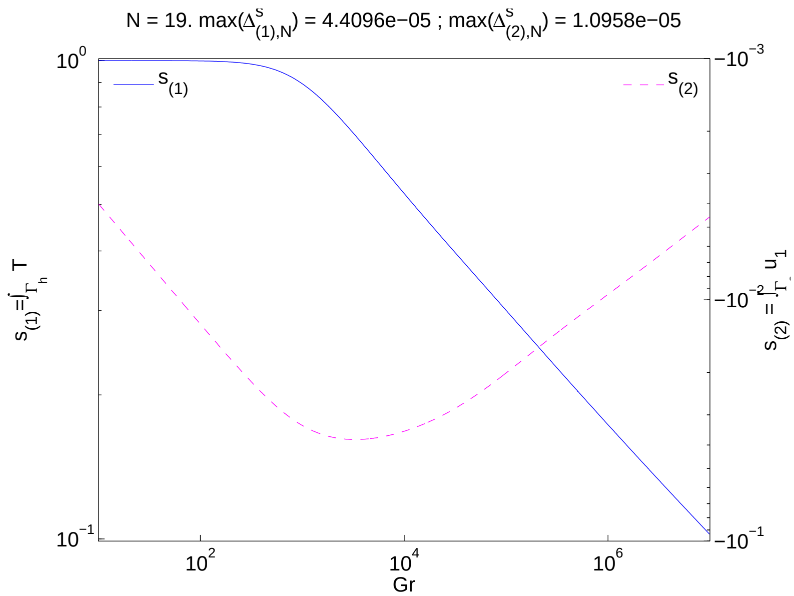

FIG. 7. *Outputs* $s_{(1)}(\boldsymbol{\mu}) = \frac{1}{|\Gamma_h|}\int_{\Gamma_h}\theta(\boldsymbol{\mu})$ *and* $s_{(2)}(\boldsymbol{\mu}) = \frac{1}{|\Gamma_s|}\int_{\Gamma_s} u_1(\boldsymbol{\mu})$ *with* $N = N_{\max} = 2{+}4{\cdot}19$.

when using the reduced basis method. Moreover, the online reduced basis problem can be solved on a notebook (one CPU and about 1GB of memory), while the finite element problem has been solved on a cluster using 16 CPUs.

Once $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) \leq 1$, our greedy algorithm seeks for a reduction in the error, not in $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ itself; our random set changes at every new search, which explains why in the last two lines of the table $\tau_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ increases.

Figure 8 shows that our inf-sup lower bound is everywhere between 0.1 and 1. We also recognize the elements of $\mathcal{V}^K = \{\overline{\boldsymbol{\mu}}_1, \ldots, \overline{\boldsymbol{\mu}}_{24}\}$ where $\overline{B}^{LB}_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) = 1$.

The error bounds for $N = 2{+}4{\cdot}19$ are small and, also important, the effectivities are small thanks to our natural norms approach (cf. (3.15) and Figure 9). Note how the error bounds are closely related to the parameters that generate our reduced basis, while the effectivities are closely related to the inf-sup lower bound (Figure 8). Since our inf-sup lower bound is better "nearby a $\overline{\boldsymbol{\mu}}$" (cf. (4.1) and (4.2)) the effectivity is better "nearby a $\overline{\boldsymbol{\mu}}$".

In Figure 7 we show our outputs $s_{(1)}$ and $s_{(2)}$ with respect to Grashof number, computed with our reduced basis on 1000 random points; the error bound is always smaller than $5{\cdot}10^{-5}$. The effectivities for the first output (Figure 10) are less sharp than for the field variable; anyway, when the error bound is small, a "bad" effectivity is acceptable, and when the error bound is the largest the effectivity is good.

8. Conclusions. We have modified the Brezzi–Rappaz–Raviart theory and defined the natural norm in the quadratic case; the latter enables a feasible and locally sharp inf-sup lower bound construction. These three components provide a fast and feasible error bounds computation, which is essential in both the online and offline stages of the reduced basis method. Our approach has proven good effectivities in the field variable error bounds and empirical good effectivities in the output error bounds.

One essential point is that the error bounds — with respect to the *truth* finite element solution — are exact, hence the computation of the inf-sup lower bound

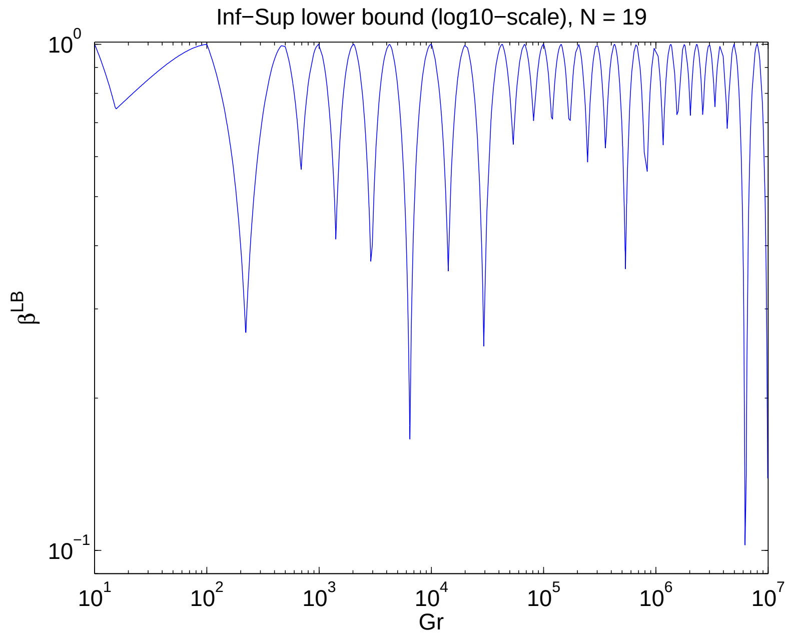

FIG. 8. $\overline{B}^{LB}_{\boldsymbol{\mu}}(\boldsymbol{\mu})$, *with* $N = N_{\max} = 2{+}4{\cdot}19$ *and* $\boldsymbol{\kappa}$ *satisfying* (4.9)

components have to be exact, too. We have developed an algorithm to remedy a robust Sobolev embedding constants computation.

We are hence able firstly, to prepare all the inf-sup lower bound components offline with a clear methodology and in a reasonable CPU wall-time, and secondly, to solve online quadratic partial differential equations — natural convection with Grashof number up to 10^7 — and provide fast and reliable error bounds, with gains in term of online resources — a notebook against a 16 processors cluster — and in terms of CPU wall-time — two orders of magnitude.

Future work will be to show that even in the multi-parameter case our approach is feasible, i.e., the offline CPU time is acceptable and the online performance is not affected (preliminary tests with Grashof and Prandl numbers as parameters are promising), including in presence of geometrical changes (where, e.g., the aspect ratio of the cavity is a further parameter).

We have presumed that the (parameter-dependent) solution lays on an isolated branch. The natural norm approach should also allow to consider bifurcations since with respect to the "local" norm the inf-sup constant is of order one. Nevertheless, the natural norm degenerates near the bifurcation because of a singular mode. This mode has probably to be treated by deflation and the eigenvalue solver should take into account the norm degeneration.

Appendix. Dual approximation.

In addition to our "truth" primal problem, we may also consider a "truth" dual (or adjoint) problem [23, 14, 3, 21, 33, 32] associated with our particular output functional to improve the convergence of the output.

In this section we consider the case of quadratic, parameter independent output

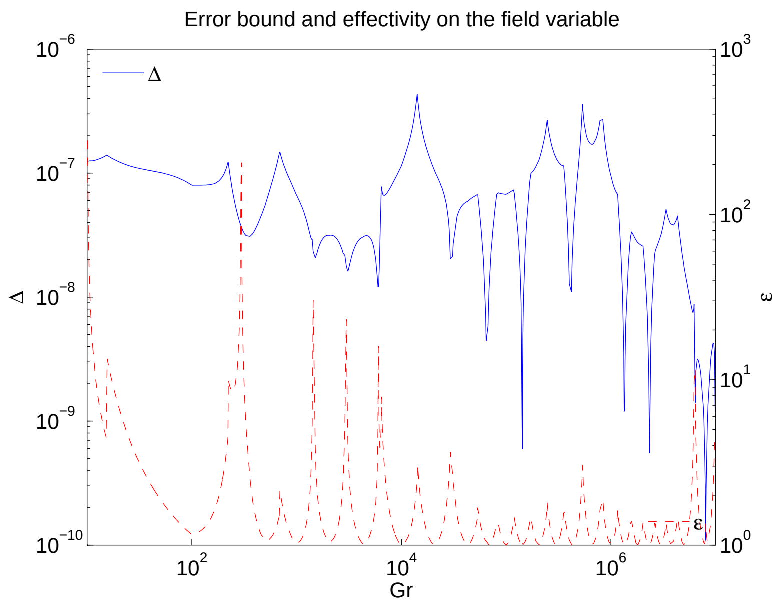

FIG. 9. *Error bound and effectivities* $\Delta_N/\|u-u_N\|_{\overline{\mu}}$ *over 1000 random samples,* $N = N_{\max} = 2{+}4{\cdot}19$.

(the extension to the linear or the affine dependent cases is straight forward),

$$\ell(\boldsymbol{v}) = \ell_0(\boldsymbol{v}) - \ell_1(\boldsymbol{v},\boldsymbol{v}),$$

where ℓ_0 is continuous and linear and ℓ_1 continuous, symmetric, and bilinear.

For a given approximation $\boldsymbol{u}_N(\boldsymbol{\mu})$ in W_N to $\boldsymbol{u}(\boldsymbol{\mu})$, the dual problem reads: find $\boldsymbol{\psi}^N$ in X such that

$$da(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})(\boldsymbol{v},\boldsymbol{\psi}^N(\boldsymbol{\mu})) \;= -\ell_0(\boldsymbol{v}) - 2\ell_1(\boldsymbol{u}_N(\boldsymbol{\mu}),\boldsymbol{v}) \quad \forall \boldsymbol{v}\in X. \tag{A.1}$$

We apply our reduce basis method to solve the dual problem (A.1) and solve for the reduced basis solution $\boldsymbol{\psi}^N_N(\boldsymbol{\mu})$ in W^{du}_N, such that

$$da(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})(\boldsymbol{v},\boldsymbol{\psi}^N_N(\boldsymbol{\mu})) \;= -\ell_0(\boldsymbol{v}) - 2\ell_1(\boldsymbol{u}_N(\boldsymbol{\mu}),\boldsymbol{v}) \quad \forall \boldsymbol{v}\in W^{du}_N,$$

where W^{du}_N (in general different from W_N) is a reduced basis space of dimension N (for simplicity in this presentation, we consider the same dimension for the reduced primal and dual spaces). We denote the primal and dual residuals as

$$\begin{aligned} r(\boldsymbol{v};\boldsymbol{\mu}) &= a(\boldsymbol{u}(\boldsymbol{\mu}),\boldsymbol{v};\mu) - a(\boldsymbol{u}_N(\boldsymbol{\mu}),\boldsymbol{v};\boldsymbol{\mu}) = f(\boldsymbol{v}) - a(\boldsymbol{u}_N(\boldsymbol{\mu}),\boldsymbol{v};\boldsymbol{\mu}) \quad \forall \boldsymbol{v}\in X,\\ r^{N,du}(\boldsymbol{v};\boldsymbol{\mu}) &= da(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})(\boldsymbol{v},\boldsymbol{\psi}^N(\boldsymbol{\mu})) - da(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})(\boldsymbol{v},\boldsymbol{\psi}^N_N(\boldsymbol{\mu}))\\ &= -\ell_0(\boldsymbol{v}) - 2\ell_1(\boldsymbol{u}_N(\boldsymbol{\mu}),\boldsymbol{v}) - da(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})(\boldsymbol{v},\boldsymbol{\psi}^N_N(\boldsymbol{\mu})) \quad \forall \boldsymbol{v}\in X, \end{aligned}$$

and the primal error as $\boldsymbol{e}(\boldsymbol{\mu}) = \boldsymbol{u}(\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})$. We define the dual corrected output as

$$s^{N,du}_N(\boldsymbol{\mu}) = \ell(\boldsymbol{u}_N(\boldsymbol{\mu})) - r(\boldsymbol{\psi}^N_N(\boldsymbol{\mu});\boldsymbol{\mu}).$$

The following lemma holds.

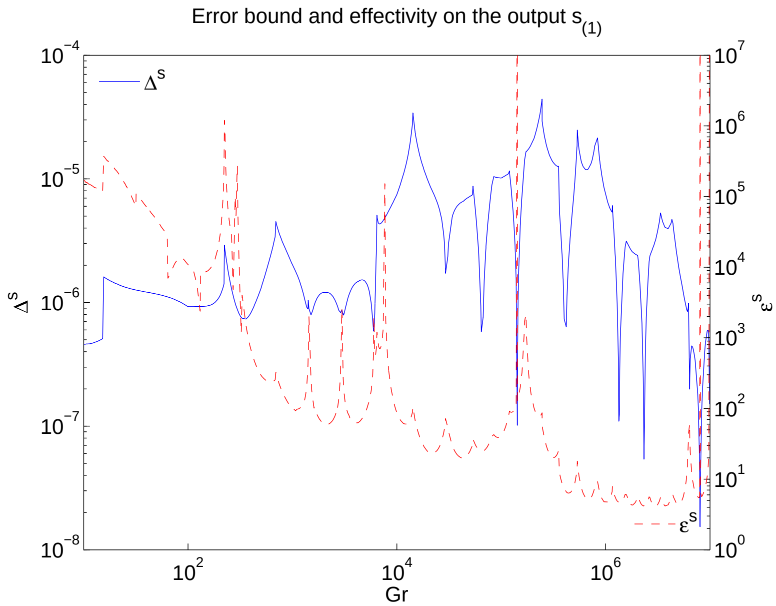

FIG. 10. *Error bound and effectivities* $\Delta_N^s/|s-s_N|$ *over 1000 random samples for output* $s_{(1)}$, $N = N_{\max} = 2{+}4{\cdot}19$.

LEMMA A.1.

$$s(\boldsymbol{\mu}) - s_N^{N,du}(\boldsymbol{\mu}) =$$
$$\ell_1(\boldsymbol{e}(\boldsymbol{\mu}),\boldsymbol{e}(\boldsymbol{\mu})) - r^{N,du}(\boldsymbol{e}(\boldsymbol{\mu});\boldsymbol{\mu}) + \frac{1}{2}\Big[da(\boldsymbol{u}(\boldsymbol{\mu});\boldsymbol{\mu}) - da(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})\Big](\boldsymbol{e}(\boldsymbol{\mu}),\boldsymbol{\psi}_N^N(\boldsymbol{\mu})). \tag{A.2}$$

Proof.

$$\begin{aligned}
s(\mu) - s_N^{N,du}(\mu) \;=\;& \ell_0(\boldsymbol{u}(\boldsymbol{\mu})) - \ell_0(\boldsymbol{u}_N(\boldsymbol{\mu})) + \ell_1\left(\boldsymbol{u}(\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu}), \boldsymbol{u}(\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu})\right)\\
&+ 2\,\ell_1\left(\boldsymbol{u}(\boldsymbol{\mu}) - \boldsymbol{u}_N(\boldsymbol{\mu}), \boldsymbol{u}_N(\boldsymbol{\mu})\right) + r(\boldsymbol{\psi}_N^N(\boldsymbol{\mu});\boldsymbol{\mu})\\
=&\ell_0(\boldsymbol{e}(\boldsymbol{\mu})) + \ell_1\left(\boldsymbol{e}(\boldsymbol{\mu}),\boldsymbol{e}(\boldsymbol{\mu})\right) + 2\,\ell_1\left(\boldsymbol{u}_N(\boldsymbol{\mu}),\boldsymbol{e}(\boldsymbol{\mu})\right)\\
&+ a(\boldsymbol{u}(\boldsymbol{\mu}),\boldsymbol{\psi}_N^N(\boldsymbol{\mu});\boldsymbol{\mu}) - a(\boldsymbol{u}_N(\boldsymbol{\mu}),\boldsymbol{\psi}_N^N(\boldsymbol{\mu});\boldsymbol{\mu})\\
=&\ell_1(\boldsymbol{e}(\boldsymbol{\mu}),\boldsymbol{e}(\boldsymbol{\mu})) - r^{N,du}(\boldsymbol{e}(\boldsymbol{\mu});\boldsymbol{\mu}) - da(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})\left(\boldsymbol{e}(\boldsymbol{\mu}),\boldsymbol{\psi}_N^N(\boldsymbol{\mu})\right)\\
&+ \frac{1}{2}\Big[da(\boldsymbol{u}(\boldsymbol{\mu});\boldsymbol{\mu}) + da(\boldsymbol{u}_N(\boldsymbol{\mu});\boldsymbol{\mu})\Big](\boldsymbol{e}(\boldsymbol{\mu}),\boldsymbol{\psi}_N^N(\boldsymbol{\mu})).
\end{aligned}$$

□

We are now able to compute an upper bound for the dual corrected output. Let

$$|||\ell_1|||_{\overline{\boldsymbol{\mu}\boldsymbol{\mu}}} \equiv \sup_{\boldsymbol{v}\in X} \frac{|\ell_1(\boldsymbol{v},\boldsymbol{v})|}{|||\boldsymbol{v}|||_{\overline{\boldsymbol{\mu}}}^2}, \tag{A.3}$$
$$\|r^{N,du}(\cdot;\boldsymbol{\mu})\|_{\overline{\boldsymbol{\mu}}'} \equiv \sup_{\boldsymbol{v}\in X} \frac{|r^{N,du}(\boldsymbol{v};\boldsymbol{\mu})|}{|||\boldsymbol{v}|||_{\overline{\boldsymbol{\mu}}}},$$

and

$$\Delta^s_{N,\overline{\mu}}(\mu) \equiv \|r^{N,du}(\cdot;\boldsymbol{\mu})\|_{\overline{\boldsymbol{\mu}}'}\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}) + \left(\frac{1}{2}\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})\|\boldsymbol{\psi}^N_N(\boldsymbol{\mu})\|_X + |||\ell_1|||_{\overline{\boldsymbol{\mu}\boldsymbol{\mu}}}\right)\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})^2. \tag{A.4}$$

Then, from (2.8), (3.14), (A.2), and (A.3), the following output error bound holds

THEOREM A.2.

$$|s(\boldsymbol{\mu}) - s^{N,du}_N(\boldsymbol{\mu})| \leq \Delta^s_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu}).$$

As we expect from the dual approach, the error bound is quadratic with respect to the dual norm of the residuals r and $r^{N,du}$. The main advantage of (A.4) is that there is no dependence on $\beta(\overline{\boldsymbol{\mu}})$, which is the square root of the worse mode of $T^{\overline{\boldsymbol{\mu}}}_N$; however, this mode may be hidden behind other terms like the natural dual norm of the dual residual, or $\rho_{\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$ (both explicitly), or in $\Delta_{N,\overline{\boldsymbol{\mu}}}(\boldsymbol{\mu})$.

Note that to simplify the offline work, the term $\|r^{N,du}(\cdot;\boldsymbol{\mu})\|_{\overline{\boldsymbol{\mu}}'}$ may also be replaced using

$$\|r^{N,du}(\cdot;\boldsymbol{\mu})\|_{\overline{\boldsymbol{\mu}}'} \leq \frac{1}{\beta(\overline{\boldsymbol{\mu}})}\|r^{N,du}(\cdot;\boldsymbol{\mu})\|_{X'}$$

at detriment of the independency on $\beta(\overline{\boldsymbol{\mu}})$.

Acknowledgements. I thank Professor Anthony T. Patera for his valuable comments and many contributions. I am grateful to Drs. Karen Veroy–Grepl and Gianluigi Rozza for many discussions on the subject. This work was supported by DARPA and AFOSR under Grant FA9550-05-1-0114 and by the Singapore-MIT Alliance.

REFERENCES

[1] B. O. Almroth, P. Stern, and F. A. Brogan, *Automatic choice of global shape functions in structural analysis*, AIAA Journal, 16 (1978), pp. 525–528.

[2] E. Balmes, *Parametric families of reduced finite element models: Theory and applications*, Mechanical Systems and Signal Processing, 10 (1996), pp. 381–394.

[3] R. Becker and R. Rannacher, *An optimal control approach to a posteriori error estimation in finite element methods*, Acta Numer., 10 (2001), pp. 1–102.

[4] F. Brezzi, J. Rappaz, and P.A. Raviart, *Finite dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions*, Numerische Mathematik, 36 (1980), pp. 1–25.

[5] H. C. Elman, David J. S., and A. J. Wathen, *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2005.

[6] J. P. Fink and W. C. Rheinboldt, *On the error behavior of the reduced basis technique for nonlinear finite element approximations*, Z. Angew. Math. Mech., 63 (1983), pp. 21–28.

[7] M. A. Grepl, N. C. Nguyen, K. Veroy, A. T. Patera, and G. R. Liu, *Certified rapid solution of partial differential equations for real-time parameter estimation and optimization*, in Proceedings of the 2nd Sandia Workshop of PDE-Constrained Optimization: Towards Real-Time and On-Line PDE-Constrained Optimization, SIAM Computational Science and Engineering Book Series, 2007.

[8] K. Ito and S. S. Ravindran, *A reduced basis method for control problems governed by PDEs*, in Control and Estimation of Distributed Parameter Systems, W. Desch, F. Kappel, and K. Kunisch, eds., Birkhäuser, 1998, pp. 153–168.

[9] ———, *A reduced-order method for simulation and control of fluid flows*, Journal of Computational Physics, 143 (1998), pp. 403–425.

[10] A. V. Knyazev, *Toward the optimal preconditioned eigensolver: locally optimal block preconditioned conjugate gradient method*, SIAM J. Sci. Comput., 23 (2001), pp. 517–541 (electronic). Copper Mountain Conference (2000).

[11] A.E. Løvgren, Y. Maday, and E.M. Rønquist, *A reduced basis element method for the steady Stokes problem*, M2AN, 40 (2006), pp. 529–552.

[12] ———, *The reduced basis element method for fluid flows*, in Analysis and Simulation of Fluid Dynamics, C. Calgaro, J.-F.Coulombel, and T. Goudon, eds., vol. VII of Adv. Math. Fluid Mech., Birkhäuser, 2007, pp. 129–154.

[13] L. Machiels, Y. Maday, I. B. Oliveira, A. T. Patera, and D. V. Rovas, *Output bounds for reduced-basis approximations of symmetric positive definite eigenvalue problems*, C. R. Acad. Sci. Paris, Série I, 331 (2000), pp. 153–158.

[14] Y. Maday, A. T. Patera, and D. V. Rovas, *A blackbox reduced-basis output bound method for noncoercive linear problems*, in Nonlinear Partial Differential Equations and Their Applications, Collége de France Seminar Volume XIV, D. Cioranescu and J.-L. Lions, eds., Elsevier Science B.V., 2002, pp. 533–569.

[15] Y. Maday, A. T. Patera, and G. Turinici, *Global* a priori *convergence theory for reduced-basis approximation of single-parameter symmetric coercive elliptic partial differential equations*, C. R. Acad. Sci. Paris, Série I, 335 (2002), pp. 289–294.

[16] N. C. Nguyen, *Reduced-Basis Approximation and* A Posteriori *Error Bounds for Nonaffine and Nonlinear Partial Differential Equations: Application to Inverse Analysis*, PhD thesis, Singapore-MIT Alliance, National University of Singapore, July 2005.

[17] N. C. Nguyen, K. Veroy, and A. T. Patera, *Certified real-time solution of parametrized partial differential equations*, in Handbook of Materials Modeling, S. Yip, ed., Springer, 2005, pp. 1523–1558.

[18] A. K. Noor and J. M. Peters, *Reduced basis technique for nonlinear analysis of structures*, AIAA Journal, 18 (1980), pp. 455–462.

[19] A.T. Patera, *A fixed point algorithm for the Sobolev embedding constant*. Private communication.

[20] J. S. Peterson, *The reduced basis method for incompressible viscous flow calculations*, SIAM J. Sci. Stat. Comput., 10 (1989), pp. 777–786.

[21] N.A. Pierce and M. B. Giles, *Adjoint recovery of superconvergent functionals from PDE approximations*, SIAM Review, 42 (2000), pp. 247–264.

[22] T. A. Porsching, *Estimation of the error in the reduced basis method solution of nonlinear equations*, Mathematics of Computation, 45 (1985), pp. 487–496.

[23] C. Prud'homme, D. Rovas, K. Veroy, Y. Maday, A. T. Patera, and G. Turinici, *Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods*, Journal of Fluids Engineering, 124 (2002), pp. 70–80.

[24] A. Quarteroni and G. Rozza, *Numerical solution of parametrized Navier-Stokes equations by reduced basis methods*, Numer. Methods Partial Differential Equations, 23 (2007), pp. 923–948.

[25] A. Quarteroni, G. Rozza, and A. Quaini, *Reduced basis methods for optimal control of advection-diffusion problem*, in Adv. in Num. Math,, W. Fitzgibbon, R. Hoppe, J. Periaux, O. Pironneau, and Y. Vassilevski, eds., Moscow, Institute of Numerical Mathematics, Russian Academy of Sciences and Houston, Department of Mathematics, University of Houston, 2006, pp. 193–216.

[26] G. Rozza, *Optimization, control and shape design for an arterial bypass*, Internat. J. Numer. Methods Fluids, 47 (2005), pp. 1411–1419.

[27] G. Rozza and K. Veroy, *On the stability of the reduced basis method for Stokes equations in parametrized domains*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 1244–1260.

[28] M. Sala, *Domain Decomposition Preconditioners: Theoretical Properties, Application to the Compressible Euler Equations, Parallel Aspects.*, PhD thesis, EPFL, 2003.

[29] M. Sala and M. Heroux, *Robust algebraic preconditioners with IFPACK 3.0*, Tech. Report SAND-0662, Sandia National Laboratories, 2005.

[30] M. Sala, K. Stanley, and M. Heroux, *Amesos: A set of general interfaces to sparse direct solver libraries*, in Proceedings of PARA'06 Conference, Umea, Sweden, 2006.

[31] ———, *On the design of interfaces to sparse direct solvers*, submitted, (2006).

[32] S. Sen, K. Veroy, D.B.P. Huynh, S. Deparis, N.C. Nguyen, and A.T. Patera, *"Natural norm" a posteriori error estimators for reduced basis approximations*, Journal of Computational Physics, 217 (2006), pp. 37–62.

[33] 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*, International Journal for Numerical Methods in Fluids, 47 (2005), pp. 773–788.

[34] K. Veroy, C. Prud'homme, and A. T. Patera, *Reduced-basis approximation of the viscous Burgers equation: Rigorous* a posteriori *error bounds*, C. R. Acad. Sci. Paris, Série I, 337 (2003), pp. 619–624.

[35] K. Veroy, C. Prud'homme, D. V. Rovas, and A. T. Patera, A posteriori *error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations (AIAA Paper 2003-3847)*, in Proceedings of the 16th AIAA Computational Fluid Dynamics Conference, June 2003.