Universität Stuttgart

SimTech
Cluster of Excellence

**B. Haasdonk**[a] · **J. Salomon**[b] · **B. Wohlmuth**[c]

# A Reduced Basis Method for Parametrized Variational Inequalities

Stuttgart, May 2011

[a] Institute of Applied Analysis and Numerical Simulation,
University of Stuttgart
Pfaffenwaldring 57, 70569 Stuttgart / Germany
haasdonk@mathematik.uni-stuttgart.de
http://www.ians.uni-stuttgart.de/agh

[b] CEREMADE
Université Paris-Dauphine
pl. de Mal. de Lattre de Tassigny, 75016 Paris / France
salomon@dauphine.fr
http://www.ceremade.dauphine.fr/ salomon

[c] Fakultät Mathematik
Technische Universität München
Boltzmannstrae 3, 85748 Garching / Germany
wohlmuth@ma.tum.de
http://www-m2.ma.tum.de

**Abstract** Reduced basis methods are an efficient tool for significantly reducing the computational complexity of solving parametrized partial differential equations. Originally introduced for elliptic equations, they have been generalized during the last decade to various types of elliptic, parabolic and hyperbolic systems. In this article, we extend the reduction technique to parametrized variational inequalities. Firstly, we propose a reduced basis variational inequality scheme in a saddle-point form and prove existence and uniqueness of the solution. We state some elementary analytical properties of the scheme such as reproduction of solutions, a-priori stability with respect to the data and Lipschitz-continuity with respect to the parameters. Secondly, we provide rigorous a-posteriori error bounds. An offline/online decomposition guarantees an efficient assembling of the reduced scheme, which can be solved by constrained quadratic programming. The reduction scheme is applied to a one-dimensional obstacle problem with a two-dimensional parameter space. The numerical results confirm the theoretical ones and demonstrate the efficiency of the reduction technique.

**Keywords** Model Reduction · Reduced Basis Methods · Variational Inequalities · A Posteriori Error Estimation

**Preprint Series** **Issue No. 2011-16**
**Stuttgart Research Centre for Simulation Technology (SRC SimTech)**
SimTech – Cluster of Excellence
Pfaffenwaldring 7a
70569 Stuttgart
publications@simtech.uni-stuttgart.de
www.simtech.uni-stuttgart.de

# A REDUCED BASIS METHOD FOR PARAMETRIZED VARIATIONAL INEQUALITIES

B. HAASDONK[†], J. SALOMON[‡] AND B. WOHLMUTH[◇]

[†] IANS, Universität Stuttgart, Germany, haasdonk@mathematik.uni-stuttgart.de
[‡] CEREMADE, Université Paris-Dauphine, France, salomon@dauphine.fr
[◇] Fakultät Mathematik, Technische Universität München, Germany, wohlmuth@ma.tum.de

May 24, 2011

**Abstract.** Reduced basis methods are an efficient tool for significantly reducing the computational complexity of solving parametrized partial differential equations. Originally introduced for elliptic equations, they have been generalized during the last decade to various types of elliptic, parabolic and hyperbolic systems. In this article, we extend the reduction technique to parametrized variational inequalities. Firstly, we propose a reduced basis variational inequality scheme in a saddle-point form and prove existence and uniqueness of the solution. We state some elementary analytical properties of the scheme such as reproduction of solutions, a-priori stability with respect to the data and Lipschitz-continuity with respect to the parameters. Secondly, we provide rigorous a-posteriori error bounds. An offline/online decomposition guarantees an efficient assembling of the reduced scheme, which can be solved by constrained quadratic programming. The reduction scheme is applied to a one-dimensional obstacle problem with a two-dimensional parameter space. The numerical results confirm the theoretical ones and demonstrate the efficiency of the reduction technique.

**Key words.** model reduction, reduced basis methods, variational inequalities, a posteriori error estima tion

**AMS subject classifications.** 35J86, 65K15, 65N12, 65N15

**1. Introduction.** We consider efficient solution strategies for parametrized variational inequalities. Such problems can be obtained from variational formulations with additional constraints, e.g., for any given parameter $\boldsymbol{\mu} \in \mathcal{P} \subset \mathbb{R}^p$ we are interested in finding a solution $u(\boldsymbol{\mu})$ of the following minimization problem:

$$\min_{u \in X(\boldsymbol{\mu})} \frac{1}{2} a(u, u; \boldsymbol{\mu}) - f(u; \boldsymbol{\mu}) \tag{1.1}$$

for $X(\boldsymbol{\mu}) \subset V$ a closed convex nonempty set in a separable Hilbert space $V$, $a(\cdot, \cdot; \boldsymbol{\mu})$ a symmetric, continuous and coercive bilinear form, and $f(\cdot; \boldsymbol{\mu})$ a continuous linear form. For $X(\boldsymbol{\mu}) = V$ the above is a standard unconstrained variational optimization problem. Then, the first order optimality condition yields a simple linear system of equations for the solution $u(\boldsymbol{\mu})$. However, if $X(\boldsymbol{\mu})$ is not a subspace of $V$, the solution cannot be obtained from a simple system of equations. Quite often, the convex set can be characterized in terms of a dual cone $M$. Then (1.1) can be reformulated as a saddle point formulation which can be solved by primal-dual active set methods. For a background on variational optimization with constraints and some applications, we refer to the monographs [8, 10, 12, 13, 23, 25, 27] and the references therein. Assume now that the above problem must be solved in a multi-query or real-time context, i.e., the computation of the solution is required to be extremely fast and/or has to be done for many parameters. For standard PDEs in variational form, reduced basis methods [3, 35] provide efficient tools for problem-specific dimensionality reduction. More precisely, instead of the full problem, which is typically infinite or rather high-dimensional, a low-dimensional model is generated which can consequently be solved significantly faster for varying parameters. Many types of partial differential equations have been treated by this reduction technique during the last decade ranging from elliptic [35] to parabolic [14, 33] and hyperbolic equations [16, 17]. So far, all results are restricted to equation systems and

no additional inequality constraints are taken into account. We are interested in adapting these techniques to a class of variational inequality problems. To start we reformulate our variational inequality as a saddle-point problem that has some similar components as the Stokes system which has already been successfully treated with reduced basis methods, e.g., [34, 32]. We refer to [7, 29] for an abstract saddle-point theory and the important role of the supremizer operator.

The article is structured as follows. In the next section, we give the elementary notation and definitions of the full and reduced problem in a saddle-point formulation, its discrete formulation and an offline/online decomposition. In Section 3, we show consistency, boundedness with respect to the data and Lipschitz-continuity with respect to the parameter. Section 4 is devoted to rigorous a-posteriori error estimation based on equality and inequality residuals. We comment in Section 5 on various computational aspects for finite element discretizations with biorthogonal dual bases for the constraints. Finally, in Section 6, we consider as a model problem a one-dimensional obstacle-type inequality. Numerical results illustrate the performance of the proposed method and the influence of the two-dimensional parameter space.

**2. Reduced basis (RB) formulation of a variational inequality.** This section is devoted to the derivation of a general RB formulation for a standard variational inequality.

**2.1. Notations.** We briefly introduce the notation and assumptions which will be used throughout the paper. By $V, W$ we denote two separable Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_V$, $\langle \cdot, \cdot \rangle_W$ and induced norms $\|\cdot\|_V$ , $\|\cdot\|_W$. The set $M \subset W$ is assumed to be a closed convex cone. We assume $a(\cdot, \cdot; \boldsymbol{\mu})$ to be a uniformly continuous and elliptic bilinear form on $V \times V$ for all $\boldsymbol{\mu} \in \mathcal{P}$, where $\mathcal{P} \subset \mathbb{R}^p, p \in \mathbb{N}$ is the parameter domain. More precisely, the parameter-dependent coercivity $\alpha(\boldsymbol{\mu})$ and continuity $\gamma_a(\boldsymbol{\mu})$ constants can be bounded for all $\boldsymbol{\mu} \in \mathcal{P}$ by $0 < \bar{\alpha} \leq \alpha(\boldsymbol{\mu})$ and $\gamma_a(\boldsymbol{\mu}) \leq \bar{\gamma}_a < \infty$, respectively. Moreover, we assume that $a(\cdot, \cdot; \boldsymbol{\mu})$ is Lipschitz-continuous with respect to $\boldsymbol{\mu}$, i.e., for a suitable constant $L_a > 0$ we have $|a(u, v; \boldsymbol{\mu}) - a(u, v; \boldsymbol{\mu}')| \leq L_a \|u\|_V \|v\|_V \|\boldsymbol{\mu} - \boldsymbol{\mu}'\|$ for all $\boldsymbol{\mu}, \boldsymbol{\mu}' \in \mathcal{P}, u, v \in V$. Here $\| \cdot \|$ denotes a norm on $\mathbb{R}^p$, e.g., the Euclidean norm. We assume that the parameter-dependent linear forms $f(\cdot; \boldsymbol{\mu}) \in V'$, $g(\cdot; \boldsymbol{\mu}) \in W'$ are uniformly continuous in $\boldsymbol{\mu}$, i.e., there exist constants $\bar{\gamma}_f, \bar{\gamma}_g > 0$ with $\|f(\cdot; \boldsymbol{\mu})\|_{V'} \leq \bar{\gamma}_f$ and $\|g(\cdot; \boldsymbol{\mu})\|_{W'} \leq \bar{\gamma}_g$ for all $\boldsymbol{\mu} \in \mathcal{P}$. Furthermore, $f(\cdot; \boldsymbol{\mu})$ and $g(\cdot; \boldsymbol{\mu})$ are supposed to be Lipschitz-continuous with respect to $\boldsymbol{\mu}$, i.e., for suitable constants $L_f, L_g > 0$ it holds $\|f(\cdot; \boldsymbol{\mu}) - f(\cdot; \boldsymbol{\mu}')\|_{V'} \leq L_f \|\boldsymbol{\mu} - \boldsymbol{\mu}'\|$ and $\|g(\cdot; \boldsymbol{\mu}) - g(\cdot; \boldsymbol{\mu}')\|_{W'} \leq L_g \|\boldsymbol{\mu} - \boldsymbol{\mu}'\|$ for all $\boldsymbol{\mu}, \boldsymbol{\mu}' \in \mathcal{P}$. Finally, $b(\cdot, \cdot)$ stands for a continuous bilinear form on $V \times W$ with continuity constant $\gamma_b > 0$, which is inf-sup stable, i.e., there exists $\beta > 0$ such that

$$\inf_{\eta \in W} \sup_{v \in V} \frac{b(v, \eta)}{\|v\|_V \|\eta\|_W} \geq \beta > 0.$$

We assume a separable parameter-dependence in $a(\cdot, \cdot; \boldsymbol{\mu})$, $f(\cdot; \boldsymbol{\mu})$ and $g(\cdot; \boldsymbol{\mu})$, i.e., the existence of parameter-dependent scalar functions $\theta_*^q : \mathcal{P} \to \mathbb{R}$ for $* \in \{a, f, g\}$, and of parameter-independent components, i.e., continuous bilinear forms $a^q(\cdot, \cdot)$ and linear functionals $f^q(\cdot) \in V', g^q(\cdot) \in W'$ for $q = 1, \ldots, Q_*$ for reasonably small $Q_a, Q_f, Q_g$, such that

$$a(u, v; \boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \theta_a^q(\boldsymbol{\mu}) a^q(u, v), \tag{2.1}$$

$$f(v; \boldsymbol{\mu}) = \sum_{q=1}^{Q_f} \theta_f^q(\boldsymbol{\mu}) f^q(v), \qquad g(\eta; \boldsymbol{\mu}) = \sum_{q=1}^{Q_g} \theta_g^q(\boldsymbol{\mu}) g^q(\eta). \tag{2.2}$$

As we will see, the case $Q_a = 1$ is of special interest. In that situation, we have $a(u, v; \boldsymbol{\mu}) = \theta_a^1(\boldsymbol{\mu}) a^1(u, v)$ with $\theta_a^1(\boldsymbol{\mu}) > 0$ and $a^1(\cdot, \cdot)$ being symmetric and coercive. Moreover, the coercivity constant $\alpha_{a^1}$ and the continuity constant $\gamma_{a^1}$ of the bilinear form $a^1(\cdot, \cdot)$ satisfy $\alpha_{a^1}/\gamma_{a^1} = \alpha(\boldsymbol{\mu})/\gamma_a(\boldsymbol{\mu})$, and the ratio is independent of $\boldsymbol{\mu}$. We then introduce the operators $A : V' \to V$ and $B, B^1 : W \to V$ by

$$a^1(A\ell, v) = \ell(v), \quad v \in V, \qquad a^1(B^1\eta, v) = b(v, \eta) = \langle B\eta, v\rangle_V, \quad v \in V.$$

**2.2. Detailed problem definition.** We continue with the full parametrized variational problem in a saddle point formulation which is to be approximated by our RB-scheme.

DEFINITION 2.1 (Variational Saddle Point Problem $SP(\boldsymbol{\mu})$).
*Given $\boldsymbol{\mu} \in \mathcal{P}$, find $(u(\boldsymbol{\mu}), \lambda(\boldsymbol{\mu})) \in V \times M$ such that*

$$\begin{aligned} a(u(\boldsymbol{\mu}), v; \boldsymbol{\mu}) + b(v, \lambda(\boldsymbol{\mu})) &= f(v; \boldsymbol{\mu}), && v \in V \\ b(u(\boldsymbol{\mu}), \eta - \lambda(\boldsymbol{\mu})) &\le g(\eta - \lambda(\boldsymbol{\mu}); \boldsymbol{\mu}), && \eta \in M. \end{aligned}$$

This formulation can be obtained from (1.1) by the special structural assumption $X(\boldsymbol{\mu}) = \{v \in V | b(v, \eta) \le g(\eta; \boldsymbol{\mu}), \eta \in M\}$. The proofs of existence and uniqueness are well-known, see, e.g., [12]. For completeness, we give them in the appendix, Sec. A.4. A notable and frequently used property of the solution $(u(\boldsymbol{\mu}), \lambda(\boldsymbol{\mu}))$ is

$$b(u(\boldsymbol{\mu}), \lambda(\boldsymbol{\mu})) = g(\lambda(\boldsymbol{\mu}); \boldsymbol{\mu}), \tag{2.3}$$

which is obtained by using $\eta = 0$ and $\eta = 2\lambda(\boldsymbol{\mu})$ as test functions in $SP(\boldsymbol{\mu})$. Note that the problem $SP(\boldsymbol{\mu})$ can be the analytical problem in infinite dimensional spaces or the discretized problem in finite dimensional spaces of high dimension. In our numerical tests, we use conforming finite elements for $V$ and for $W$, which results in spaces of the same dimension. In the case of finite dimensional spaces $V = \operatorname{span}\{\psi_i, i = 1, \ldots, H_V\}$ and $W = \operatorname{span}\{\chi_i, i = 1, \ldots, H_W\}$, we denote by $H_V$ and $H_W$ the dimension of $V$ and $W$, respectively.

**2.3. Reduced problem definition.** We now derive from the saddle point formulation of Def. 2.1 a corresponding RB-method. Let $S = \{\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{N_S}\} \subset \mathcal{P}$ denote a finite parameter sample set of $N_S$ parameters and $(u(\boldsymbol{\mu}_i), \lambda(\boldsymbol{\mu}_i)) \in V \times M$ the corresponding solutions of $SP(\boldsymbol{\mu}_i)$, the so called *snapshots*. Define $W_N := \operatorname{span}\{\lambda(\boldsymbol{\mu}_i)\}_{i=1}^{N_S} \subset W$ as a reduced basis space with dimension $N_W := \dim W_N$ and basis $\{\xi_i\}_{i=1}^{N_W}$. Then, $M_N := \left\{\sum_{i=1}^{N_S} \alpha_i \lambda(\boldsymbol{\mu}_i) | \alpha_i \ge 0\right\}$ is a closed convex cone and in particular $M_N \subset M$. This choice implies in particular that $\lambda(\boldsymbol{\mu}_i) \in M_N$ for all $i = 1, \ldots, N_S$. Note that we do not explicitly require linear independence of the $\lambda(\boldsymbol{\mu}_i)$, but also accept possibly linearly dependent snapshots. In this case, elements of $\eta \in M_N$ may have multiple different equivalent expansions as linear combination of snapshots.

For formulating the reduced scheme, it remains to give a definition of the reduced space $V_N$ for the primal variable. For this space, we will consider different choices.

1. Pure snapshots: The naive choice for the reduced primal space is given by

$$V_N^{(1)} := \operatorname{span}\{u(\boldsymbol{\mu}_i)\}_{i=1}^{N_S} \subset V. \tag{2.4}$$

2. Enrichment by supremizers: This choice is motivated by [34]

$$V_N^{(2)} := \operatorname{span}\{u(\boldsymbol{\mu}_i), B\xi_j\}_{i,j=1}^{N_S, N_W} \subset V \tag{2.5}$$

3. Enrichment by a-priori solutions: If $Q_a = 1$, an attractive alternative option is

$$V_N^{(3)} := \operatorname{span}\{u(\boldsymbol{\mu}_i), Af^q\}_{i,q=1}^{N_S,Q_f} \subset V. \tag{2.6}$$

4. Enrichment by unconstrained solutions: If $Q_a = 1$, we can also set

$$V_N^{(4)} := \operatorname{span}\{u(\boldsymbol{\mu}_i), \tilde{u}(\boldsymbol{\mu}_j)\}_{i,j=1}^{N_S,N_S} \subset V, \tag{2.7}$$

where $\tilde{u}(\boldsymbol{\mu}_i) \in V$ satisfies $a(\tilde{u}(\boldsymbol{\mu}_i), v; \boldsymbol{\mu}_i) = f(v; \boldsymbol{\mu}_i)$, $v \in V$.

We neglect quite often the upper index $(l)$, $l = 1, 2, 3, 4$, set $N_V := \dim V_N$ and denote the basis of $V_N$ by $\{\varphi_i\}_{i=1}^{N_V}$. We point out that the subscript $N$ does not stand for the dimension, but only indicates that the quantity is connected to a *reduced* problem. We have by construction $\dim W_N \leq N_S$. In the following, we use the following notation: given a vector $w$, we denote its coefficient vectors by $\overline{w}$ when dealing with the reduced basis and by $\underline{w}$ when considering its representation in the high (finite-)dimensional spaces. Moreover from time to time, we omit the $\boldsymbol{\mu}$-dependence, e.g., in the proof of Prop. 3.2.

The goal of a reduced basis scheme is the computation of parameter-dependent solutions $u_N(\boldsymbol{\mu}) \in V_N, \lambda_N(\boldsymbol{\mu}) \in W_N$ by solving a saddle point problem of low complexity.

DEFINITION 2.2 (Reduced Basis Saddle Point Problem $SP_N(\boldsymbol{\mu})$).
*For $\boldsymbol{\mu} \in \mathcal{P}$ find $(u_N(\boldsymbol{\mu}), \lambda_N(\boldsymbol{\mu})) \in V_N \times M_N$ such that*

$$\begin{aligned} a(u_N(\boldsymbol{\mu}), v_N; \boldsymbol{\mu}) + b(v_N, \lambda_N(\boldsymbol{\mu})) &= f(v_N; \boldsymbol{\mu}), \quad v_N \in V_N \\ b(u_N(\boldsymbol{\mu}), \eta_N - \lambda_N(\boldsymbol{\mu})) &\leq g(\eta_N - \lambda_N(\boldsymbol{\mu}); \boldsymbol{\mu}), \quad \eta_N \in M_N. \end{aligned}$$

The following proposition shows that the pairing $(V_N, W_N)$ inherits its inf-sup constant from the pairing $(V, W)$, and thus $SP_N(\boldsymbol{\mu})$ has a unique solution for $V_N = V_N^{(i)}$, $i = 2, 3, 4$. Therefore, these cases are of special interest.

PROPOSITION 2.3 (Existence and Uniqueness of $SP_N(\boldsymbol{\mu})$). *If the reduced primal space $V_N$ is chosen as in* (2.5)–(2.7)*, then the pairing $(V_N, W_N)$ is inf-sup stable with inf-sup constant $\beta_N \geq \beta$ for $V_N = V_N^{(2)}$ and $\beta_N \geq \beta\sqrt{\alpha_{a^1}/\gamma_{a^1}}$ for $V_N = V_N^{(i)}$, $i = 3, 4$.*

*Proof.* We consider the case $V_N = V_N^{(2)}$ and the cases $V_N = V_N^{(3)}$ or $V_N = V_N^{(4)}$ separately. Let $V_N = V_N^{(2)}$, then by construction $B\eta_N \in V_N$ for all $\eta_N \in W_N$ and thus

$$\begin{aligned} \beta_N &:= \inf_{\eta_N \in W_N} \sup_{v_N \in V_N} \frac{b(v_N, \eta_N)}{\|v_N\|_V \|\eta_N\|_W} = \inf_{\eta_N \in W_N} \sup_{v_N \in V_N} \frac{\langle v_N, B\eta_N \rangle_V}{\|v_N\|_V \|\eta_N\|_W} \\ &= \inf_{\eta_N \in W_N} \frac{\langle B\eta_N, B\eta_N \rangle_V}{\|B\eta_N\|_V \|\eta_N\|_W} \geq \inf_{\eta \in W} \frac{\langle B\eta, B\eta \rangle_V}{\|B\eta\|_V \|\eta\|_W} = \inf_{\eta \in W} \sup_{v \in V} \frac{\langle v, B\eta \rangle_V}{\|v\|_V \|\eta\|_W} = \beta > 0. \end{aligned}$$

If $V_N = V_N^{(3)}$ and $Q_a = 1$, we obtain by the fact that $(u(\boldsymbol{\mu}_i), \lambda(\boldsymbol{\mu}_i))$ solves $SP(\boldsymbol{\mu}_i)$

$$a^1(B^1\lambda(\boldsymbol{\mu}_i), v) = b(v, \lambda(\boldsymbol{\mu}_i)) = \theta_a^1(\boldsymbol{\mu}_i) a^1(u(\boldsymbol{\mu}_i), v) - f(v; \boldsymbol{\mu}_i).$$

Then, the definition (2.6) and the decomposition (2.2) yield that

$$B^1\lambda(\boldsymbol{\mu}_i) = \left( \theta_a^1(\boldsymbol{\mu}_i) u(\boldsymbol{\mu}_i) - \sum_{q=1}^{Q_f} \theta_f^q(\boldsymbol{\mu}_i) Af^q \right) \in V_N^{(3)}.$$

If $V_N = V_N^{(4)}$ and $Q_a = 1$, we get that $B^1\lambda(\boldsymbol{\mu}_i) \in V_N^{(4)}$ since

$$a^1(B^1\lambda(\boldsymbol{\mu}_i), v) = b(v, \lambda(\boldsymbol{\mu}_i)) = \theta_a^1(\boldsymbol{\mu}_i) a^1(u(\boldsymbol{\mu}_i), v) - \theta_a^1(\boldsymbol{\mu}_i) a^1(\tilde{u}(\boldsymbol{\mu}_i), v).$$

By linearity, this implies that $B^1\eta_N \in V_N^{(i)}$, $i = 3, 4$ for all $\eta_N \in W_N$. Then, the norm equivalence on $V$ gives

$$\begin{aligned}
\beta_N &:= \inf_{\eta_N \in W_N} \sup_{v_N \in V_N} \frac{b(v_N, \eta_N)}{\|v_N\|_V \|\eta_N\|_W} = \inf_{\eta_N \in W_N} \sup_{v_N \in V_N} \frac{a^1(v_N, B^1\eta_N)}{\|v_N\|_V \|\eta_N\|_W} \\
&\geq \inf_{\eta_N \in W_N} \sup_{v_N \in V_N} \frac{\sqrt{\alpha_{a^1}} a^1(v_N, B^1\eta_N)}{\sqrt{a^1(v_N, v_N)}\, \|\eta_N\|_W} = \inf_{\eta_N \in W_N} \frac{\sqrt{\alpha_{a^1}} a^1(B^1\eta_N, B^1\eta_N)}{\sqrt{a^1(B^1\eta_N, B^1\eta_N)}\, \|\eta_N\|_W} \\
&= \sqrt{\alpha_{a^1}} \inf_{\eta_N \in W_N} \sup_{v \in V} \frac{a^1(v, B^1\eta_N)}{\sqrt{a^1(v, v)}\, \|\eta_N\|_W} \geq \frac{\sqrt{\alpha_{a^1}}}{\sqrt{\gamma_{a^1}}} \inf_{\eta_N \in W_N} \sup_{v \in V} \frac{b(v, \eta_N)}{\|v\|_V \|\eta_N\|_W} \geq \frac{\sqrt{\alpha_{a^1}}}{\sqrt{\gamma_{a^1}}} \beta.
\end{aligned}$$

□

No stability statement can be given for the choice $V_N^{(1)}$ (2.4). One can even explicitly construct pathological cases, where the uniqueness is not valid: Assume a simple example of $N_S = 1$, $\boldsymbol{\mu}_1 \in \mathcal{P}$ and $g(\cdot; \boldsymbol{\mu}_1) = 0$. Then, the solutions $u(\boldsymbol{\mu}_1), \lambda(\boldsymbol{\mu}_1)$ are $b$-orthogonal, i.e., $b(u(\boldsymbol{\mu}_1), \lambda(\boldsymbol{\mu}_1)) = 0$, which is obtained from (2.3). If we assume a definition of reduced spaces without supremizer, i.e., $V_N := \text{span}\{u(\boldsymbol{\mu}_1)\}$, $M_N := \{s\lambda(\boldsymbol{\mu}_1), s \in \mathbb{R}_0^+\} \subset W_N := \text{span}\{\lambda(\boldsymbol{\mu}_1)\}$, this implies two conceptional problems: First, the solution $u_N(\boldsymbol{\mu}_1)$ of $SP_N(\boldsymbol{\mu})$ is not constrained in any way, hence we solve an ordinary unconstrained PDE. Second, we lose the uniqueness of solutions, due to the lack of the inf-sup stability. Any $(u_N, \lambda_N) := (u(\boldsymbol{\mu}_1), s\lambda(\boldsymbol{\mu}_1))$ for $s \in \mathbb{R}_0^+$ is a solution of $SP_N(\boldsymbol{\mu}_1)$. Still, in practice the choice (2.4) may lead to an inf-sup stable scheme, but the inf-sup stability constants can possibly be arbitrarily small.

REMARK 2.4. *If $Q_a = 1$, the choices* (2.6) *and* (2.7) *are possibly computationally attractive alternatives to* (2.5)*. Firstly in the case of $N_W \gg Q_f$, the $Q_f$ a-priori solutions for the enrichment in* (2.6) *are considerably less costly than the $N_W$ supremizer functions in* (2.5)*. Secondly, the constrained solution is often calculated in terms of an iterative solver which uses as initial guess the solution $\tilde{u}(\boldsymbol{\mu}_i)$ of the unconstrained system. Thus no additional cost at all is then required in* (2.7)*. It is easy to see that $\tilde{u}(\boldsymbol{\mu}_i) = u(\boldsymbol{\mu}_i) - \alpha_i B^1 \lambda(\boldsymbol{\mu}_i) \in V_N^{(3)}$ with $\alpha_i^{-1} := \theta_a^1(\boldsymbol{\mu}_i)$ and thus $V_N^{(4)} \subset V_N^{(3)}$. Moreover, if additionally $a^1(\cdot, \cdot) = \sigma\langle \cdot, \cdot \rangle_V$ with $\sigma > 0$, then $B = \sigma B^1$ and $B\lambda(\boldsymbol{\mu}_i) \in V_N^{(3)}$ and thus $V_N^{(2)} \subset V_N^{(3)}$, and the dimension is at most $N_S + Q_f$.*

**2.4. Algebraic formulation.** In order to formulate a discrete matrix inequality problem, we introduce matrices and vectors

$$\overline{A}_N(\boldsymbol{\mu}) := (a(\varphi_j, \varphi_i; \boldsymbol{\mu}))_{i,j=1}^{N_V} \in \mathbb{R}^{N_V \times N_V}, \tag{2.8}$$

$$\overline{B}_N := \big(b(\varphi_i, \lambda(\boldsymbol{\mu}_j))\big)_{i,j=1}^{N_V, N_S} \in \mathbb{R}^{N_V \times N_S}, \tag{2.9}$$

$$\overline{f}_N(\boldsymbol{\mu}) := (f(\varphi_i; \boldsymbol{\mu}))_{i=1}^{N_V} \in \mathbb{R}^{N_V}, \tag{2.10}$$

$$\overline{g}_N(\boldsymbol{\mu}) := (g(\lambda(\boldsymbol{\mu}_i); \boldsymbol{\mu}))_{i=1}^{N_S} \in \mathbb{R}^{N_S}. \tag{2.11}$$

We then obtain the following algebraic form of $SP_N(\boldsymbol{\mu})$:

LEMMA 2.5 (Algebraic Reduced Basis Saddle Point Problem $DSP_N(\boldsymbol{\mu})$). *The solution $(u_N(\boldsymbol{\mu}), \lambda_N(\boldsymbol{\mu}))$ of the reduced saddle point problem $SP_N(\boldsymbol{\mu})$ expanded as $u_N(\boldsymbol{\mu}) := \sum_{i=1}^{N_V} \overline{u}_{N,i}\varphi_i$ and $\lambda_N(\boldsymbol{\mu}) := \sum_{i=1}^{N_S} \overline{\lambda}_{N,i}\lambda(\boldsymbol{\mu}_i)$ where $\overline{u}_N(\boldsymbol{\mu}) := (\overline{u}_{N,i})_{i=1}^{N_V} \in \mathbb{R}^{N_V}$ and*

$\overline{\lambda}_N(\boldsymbol{\mu}) := (\overline{\lambda}_{N,i})_{i=1}^{N_S} \in \mathbb{R}^{N_S}$ *is equivalently characterized by the matrix inequality system*

$$\overline{A}_N(\boldsymbol{\mu})\overline{u}_N(\boldsymbol{\mu}) + \overline{B}_N\overline{\lambda}_N(\boldsymbol{\mu}) = \overline{f}_N(\boldsymbol{\mu}) \tag{2.12}$$

$$\overline{\lambda}_N(\boldsymbol{\mu}) \geq 0 \tag{2.13}$$

$$\overline{g}_N(\boldsymbol{\mu}) - \overline{B}_N^T\overline{u}_N(\boldsymbol{\mu}) \geq 0 \tag{2.14}$$

$$\overline{\lambda}_N(\boldsymbol{\mu})^T(\overline{g}_N(\boldsymbol{\mu}) - \overline{B}_N^T\overline{u}_N(\boldsymbol{\mu})) = 0. \tag{2.15}$$

*Proof.* (2.12)–(2.14) is easy to verify in terms of the definition of the reduced dual cone $M_N$. By setting $\eta = \lambda_N(\boldsymbol{\mu}) \pm \overline{\lambda}_{N,i}\lambda_N(\boldsymbol{\mu}_i)$, we get the component-wise complementarity condition

$$\overline{\lambda}_{N,i}(\overline{g}_N(\boldsymbol{\mu}) - \overline{B}_N^T\overline{u}_N(\boldsymbol{\mu}))_i = 0, \quad i = 1, \ldots, N_S, \tag{2.16}$$

where $(\overline{g}_N(\boldsymbol{\mu}) - \overline{B}_N^T\overline{u}_N(\boldsymbol{\mu}))_i$ is the $i$-th coefficient of the vector $\overline{g}_N(\boldsymbol{\mu}) - \overline{B}_N^T\overline{u}_N(\boldsymbol{\mu})$ and thus (2.15). □

We point out that $\overline{u}_N(\boldsymbol{\mu})$ is unique, whereas the uniqueness of $\overline{\lambda}_N(\boldsymbol{\mu})$ cannot be guaranteed due to the possible linear dependence of the snapshots $\lambda(\boldsymbol{\mu}_i)$. However, all solutions of the coefficient vector $\overline{\lambda}_N(\boldsymbol{\mu})$ represent the same solution function $\lambda_N(\boldsymbol{\mu}) \in M_N$.

**2.5. Offline/online decomposition.** The parameter-dependence of $a(\cdot,\cdot;\boldsymbol{\mu})$, $f(\cdot;\boldsymbol{\mu})$ and $g(\cdot;\boldsymbol{\mu})$ introduced in the previous section transfers into an offline/online decomposition of $DSP_N(\boldsymbol{\mu})$ as follows:

**Offline-phase.** We compute the parameter-independent matrices

$$\begin{aligned}
\overline{A}_N^q &:= (a^q(\varphi_j, \varphi_i))_{i,j=1}^{N_V} \in \mathbb{R}^{N_V \times N_V}, \quad q = 1, \ldots, Q_a \\
\overline{B}_N &:= \big(b(\varphi_i, \lambda(\boldsymbol{\mu}_j))\big)_{i,j=1}^{N_V, N_S} \in \mathbb{R}^{N_V \times N_S}, \\
\overline{f}_N^q &:= (f^q(\varphi_i))_{i=1}^{N_V} \in \mathbb{R}^{N_V}, \quad q = 1, \ldots, Q_f \\
\overline{g}_N^q &:= (g^q(\lambda(\boldsymbol{\mu}_i)))_{i=1}^{N_S} \in \mathbb{R}^{N_S}, \quad q = 1, \ldots, Q_g.
\end{aligned}$$

**Online-phase.** We assemble the parameter-dependent matrices and right hand sides

$$\overline{A}_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \theta_a^q(\boldsymbol{\mu})\overline{A}_N^q, \qquad \overline{f}_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_f} \theta_f^q(\boldsymbol{\mu})\overline{f}_N^q, \qquad \overline{g}_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_g} \theta_g^q(\boldsymbol{\mu})\overline{g}_N^q$$

and solve the discrete reduced problem $DSP_N(\boldsymbol{\mu})$. In particular, we remark that all matrices and vectors involved in the online-phase are low-dimensional. Hence, all operations in the online-phase are independent of the dimensions of $V, W$. Typically, in practice we have $H_W, H_V \gg N_V, N_W, N_S$, hence the reduced online-phase will be considerably faster than the high-dimensional problem solution.

**3. Theoretical results.** In this section, we comment on some analytical aspects, namely a consistency result, stability and Lipschitz-continuity.

**3.1. Reproduction of solutions.** Here we state a consistency result which shows that detailed solutions are recovered by the reduced scheme if the corresponding snapshots are contained in the reduced spaces. This could also be formulated as a corollary of the a-posteriori error analysis of Section 4, but a direct proof is more elegant and elementary.

PROPOSITION 3.1 (Reproduction of Solutions). *Let $V_N$ be given by* (2.5) *or* (2.6) *or* (2.7). *If for some $\boldsymbol{\mu} \in \mathcal{P}$ we have $u(\boldsymbol{\mu}) \in V_N$ and $\lambda(\boldsymbol{\mu}) \in M_N$, then $u_N(\boldsymbol{\mu}) = u(\boldsymbol{\mu})$ and $\lambda_N(\boldsymbol{\mu}) = \lambda(\boldsymbol{\mu})$.*

*Proof.* For $v_N \in V_N$ we directly obtain

$$a(u(\boldsymbol{\mu}), v_N; \boldsymbol{\mu}) + b(v_N, \lambda(\boldsymbol{\mu})) = f(v_N; \boldsymbol{\mu})$$

as $V_N \subset V$ and $(u(\boldsymbol{\mu}), \lambda(\boldsymbol{\mu}))$ solves $SP(\boldsymbol{\mu})$. Similarly, for $\eta_N \in M_N$ we directly obtain

$$b(u(\boldsymbol{\mu}), \eta_N - \lambda(\boldsymbol{\mu})) \leq g(\eta_N - \lambda(\boldsymbol{\mu}); \boldsymbol{\mu})$$

as $M_N \subset M$ and $(u(\boldsymbol{\mu}), \lambda(\boldsymbol{\mu}))$ solves $SP(\boldsymbol{\mu})$. Consequently, $(u(\boldsymbol{\mu}), \lambda(\boldsymbol{\mu}))$ solves $SP_N(\boldsymbol{\mu})$. Due to the uniqueness from Prop. 2.3, we conclude that $(u(\boldsymbol{\mu}), \lambda(\boldsymbol{\mu})) = (u_N(\boldsymbol{\mu}), \lambda_N(\boldsymbol{\mu}))$. □

**3.2. Boundedness.** Next, we show that the solutions are bounded by the data functions of the saddle point problem:

PROPOSITION 3.2 (A-priori Stability Estimates). *Let $V_N = V_N^{(i)}$, $i = 2, 3, 4$ given by* (2.5)–(2.7). *Then, the solution $(u_N(\boldsymbol{\mu}), \lambda_N(\boldsymbol{\mu}))$ of $SP_N(\boldsymbol{\mu})$ is uniformly bounded with respect to $\boldsymbol{\mu}$:*

$$\|u_N(\boldsymbol{\mu})\|_V \leq \frac{1}{2\bar{\alpha}} \left( \bar{\gamma}_f + \frac{\bar{\gamma}_a}{\beta_N} \bar{\gamma}_g \right) + \sqrt{\frac{1}{4\bar{\alpha}^2} \left( \bar{\gamma}_f + \frac{\bar{\gamma}_a}{\beta_N} \bar{\gamma}_g \right)^2 + \frac{\bar{\gamma}_g \bar{\gamma}_f}{\bar{\alpha}\beta_N}} =: \bar{\gamma}_u, \quad (3.1)$$

$$\|\lambda_N(\boldsymbol{\mu})\|_W \leq \frac{1}{\beta_N} \left( \bar{\gamma}_f + \bar{\gamma}_a \bar{\gamma}_u \right) =: \bar{\gamma}_\lambda. \quad (3.2)$$

*Proof.* We start with the proof of (3.2). As $b(\cdot, \cdot)$ is inf-sup stable on $V_N \times W_N$, and $V_N$ is finite dimensional, there exists a $v_{\lambda_N} \in V_N$ such that

$$\begin{aligned} \beta_N \|v_{\lambda_N}\|_V \|\lambda_N\|_W &\leq b(v_{\lambda_N}, \lambda_N) = f(v_{\lambda_N}) - a(u_N, v_{\lambda_N}) \\ &\leq \bar{\gamma}_f \|v_{\lambda_N}\|_V + \bar{\gamma}_a \|u_N\|_V \|v_{\lambda_N}\|_V , \end{aligned}$$

where the last inequality follows from the uniform continuity of $f(\cdot; \boldsymbol{\mu})$ and $a(\cdot, \cdot; \boldsymbol{\mu})$. Hence

$$\|\lambda_N\|_W \leq \frac{1}{\beta_N} \left( \bar{\gamma}_f + \bar{\gamma}_a \|u_N\|_V \right) \quad (3.3)$$

and (3.2) follows as soon as (3.1) is established. The complimentarity (2.3) also holds for the reduced solution, i.e., $b(u_N, \lambda_N) = g(\lambda_N)$. In terms of coercivity and continuity, we get

$$\begin{aligned} \bar{\alpha} \|u_N\|_V^2 \leq a(u_N, u_N) &= f(u_N) - b(u_N, \lambda_N) \\ &= f(u_N) - g(\lambda_N) \leq \bar{\gamma}_f \|u_N\|_V + \bar{\gamma}_g \|\lambda_N\|_W . \end{aligned}$$

Inserting (3.3) and rearranging the terms, we conclude with

$$\|u_N\|_V^2 - \underbrace{\frac{1}{\bar{\alpha}} \left( \bar{\gamma}_f + \frac{\bar{\gamma}_g \bar{\gamma}_a}{\beta_N} \right)}_{=:p} \|u_N\|_V - \underbrace{\frac{\bar{\gamma}_g \bar{\gamma}_f}{\bar{\alpha}\beta_N}}_{=:q} \leq 0.$$

We observe that $p, q \geq 0$, the quadratic equation $x^2 - px - q = 0$ has real roots $x_1 \leq x_2$ and $\|u_N\|_V \in [x_1, x_2]$ such that we finally obtain (3.1):

$$\|u_N\|_V \leq x_2 = \frac{p}{2} + \sqrt{\frac{p^2}{4} + q} = \bar{\gamma}_u.$$

□

**3.3. Lipschitz-continuity.** In this section, we are interested in some regularity results of the solutions with respect to the parameter. As we will show, Lipschitz-continuity of the reduced solutions holds under the assumption of Lipschitz-continuity of the data with respect to $\boldsymbol{\mu}$.

PROPOSITION 3.3 (Lipschitz-continuity with respect to $\boldsymbol{\mu}$). *Let $V_N$ be given by* (2.5) *or* (2.6) *or* (2.7)*. Then, the solution $(u_N(\boldsymbol{\mu}), \lambda_N(\boldsymbol{\mu}))$ of $SP_N(\boldsymbol{\mu})$ is Lipschitz-continuous with respect to $\boldsymbol{\mu}$, i.e., for all $\boldsymbol{\mu}, \boldsymbol{\mu}'$ holds*

$$\|u_N(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu}')\|_V \leq L_u \|\boldsymbol{\mu} - \boldsymbol{\mu}'\|, \tag{3.4}$$
$$\|\lambda_N(\boldsymbol{\mu}) - \lambda_N(\boldsymbol{\mu}')\|_W \leq L_\lambda \|\boldsymbol{\mu} - \boldsymbol{\mu}'\|, \tag{3.5}$$

*with constants independent of $\boldsymbol{\mu}, \boldsymbol{\mu}'$*

$$L_u := C_1 + \sqrt{C_1^2 + C_2}, \qquad L_\lambda := \frac{1}{\beta_N}\left(L_f + L_a\bar{\gamma}_u + \bar{\gamma}_a L_u\right),$$
$$C_1 := \frac{1}{2\bar{\alpha}}\left(\frac{L_g\bar{\gamma}_a}{\beta_N} + L_f + L_a\bar{\gamma}_u\right), \qquad C_2 := \frac{L_g}{\bar{\alpha}\beta_N}\left(L_f + L_a\bar{\gamma}_u\right).$$

*Here $L_a, L_f, L_g$ are the Lipschitz-constants of $a(\cdot,\cdot;\boldsymbol{\mu})$, $f(\cdot;\boldsymbol{\mu})$ and $g(\cdot;\boldsymbol{\mu})$, and $\bar{\gamma}_u$ is defined by (3.1).*

*Proof.* We assume $\boldsymbol{\mu}, \boldsymbol{\mu}' \in \mathcal{P}$ and introduce some abbreviations (with slight abuse of notation) to facilitate the readability: $u := u_N(\boldsymbol{\mu}), u' := u_N(\boldsymbol{\mu}'), \lambda := \lambda_N(\boldsymbol{\mu}), \lambda' := \lambda_N(\boldsymbol{\mu}'), a(\cdot,\cdot) := a(\cdot,\cdot;\boldsymbol{\mu}), a'(\cdot,\cdot) := a(\cdot,\cdot;\boldsymbol{\mu}'), f(\cdot) := f(\cdot;\boldsymbol{\mu}), f'(\cdot) := f(\cdot;\boldsymbol{\mu}'), g(\cdot) := g(\cdot;\boldsymbol{\mu}), g'(\cdot) := g(\cdot;\boldsymbol{\mu}')$. Then, obviously we have

$$a(u,v) + b(v,\lambda) = f(v), \quad v \in V_N, \tag{3.6}$$
$$a'(u',v) + b(v,\lambda') = f'(v), \quad v \in V_N. \tag{3.7}$$

Due to the inf-sup stability for $\lambda - \lambda' \in W_N$ there exists a $v \in V_N$ with

$$\begin{aligned}\beta_N \|v\|_V \|\lambda - \lambda'\|_W &\leq b(v, \lambda - \lambda') = b(v,\lambda) - b(v,\lambda') \\ &= f(v) - a(u,v) - f'(v) + a'(u',v) + a(u',v) - a(u',v) \\ &\leq L_f \|v\|_V \|\boldsymbol{\mu} - \boldsymbol{\mu}'\| + L_a \|u'\|_V \|v\|_V \|\boldsymbol{\mu} - \boldsymbol{\mu}'\| + \bar{\gamma}_a \|u - u'\|_V \|v\|_V .\end{aligned}$$

Using the boundedness of $u'$ due to Prop. 3.2, we obtain

$$\|\lambda - \lambda'\|_W \leq \frac{1}{\beta_N}\left((L_f + L_a\bar{\gamma}_u)\|\boldsymbol{\mu} - \boldsymbol{\mu}'\| + \bar{\gamma}_a \|u - u'\|_V\right). \tag{3.8}$$

Now, the inequality of the saddle point problem yields

$$\begin{aligned}b(u - u', \lambda' - \lambda) &= b(u, \lambda' - \lambda) + b(u', \lambda - \lambda') \\ &\leq g(\lambda' - \lambda) + g'(\lambda - \lambda') \leq L_g \|\lambda' - \lambda\|_W \|\boldsymbol{\mu} - \boldsymbol{\mu}'\|.\end{aligned}$$

Moreover, we find for $v \in V_N$

$$\begin{aligned}a(u - u', v) &= a(u,v) - a(u',v) \\ &= -b(v,\lambda) + f(v) - a(u',v) + a'(u',v) + b(v,\lambda') - f'(v) \\ &\leq b(v, \lambda' - \lambda) + L_f \|v\|_V \|\boldsymbol{\mu} - \boldsymbol{\mu}'\| + L_a \|u'\|_V \|v\|_V \|\boldsymbol{\mu} - \boldsymbol{\mu}'\|.\end{aligned}$$

Then, the coercivity in combination with $v = u - u'$ guarantees

$$\begin{aligned}\bar{\alpha}\,\|u-u'\|_V^2 &\le a(u-u',u-u')\\ &\le L_g\,\|\lambda'-\lambda\|_W\;\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|\\ &\quad +L_f\,\|u-u'\|_V\;\|\boldsymbol{\mu}-\boldsymbol{\mu}'\| + L_a\,\|u'\|_V\;\|u-u'\|_V\;\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|.\end{aligned}$$

Using the boundedness of $u'$ this simplifies to

$$\|u-u'\|_V^2 \le \frac{1}{\bar{\alpha}}\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|\left(L_g\,\|\lambda'-\lambda\|_W + (L_f + L_a\bar{\gamma}_u)\,\|u-u'\|_V\right).$$

Inserting (3.8) and rearranging the terms gives

$$\begin{aligned}\|u-u'\|_V^2 - \frac{1}{\bar{\alpha}}\left(\frac{L_g\bar{\gamma}_a}{\beta_N} + L_f + L_a\bar{\gamma}_u\right)\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|\;\|u-u'\|_V\\ -\left(\frac{L_g}{\bar{\alpha}\beta_N}(L_f + L_a\bar{\gamma}_u)\right)\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|^2 \le 0.\end{aligned}$$

We argue as in the proof of Prop. 3.2: using that the left hand side is of the form $x^2 - 2C_1\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|x - C_2\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|^2$ and has real roots $x_1 \le x_2$, we conclude that $\|u-u'\|_V < x_2$ which proves (3.4). Inserting the last result in (3.8) finally gives (3.5). □

**4. A-posteriori error analysis.** In this section, we focus on the efficient control of the reduction error by a-posteriori error estimators. Adaptive techniques based on a-posteriori error estimators play an important role in enhancing the performance of finite element discretizations, see, e.g., the monographs [1, 19, 31, 36] and the references therein. For abstract variational inequalities in the context of finite elements, we refer to [2, 5, 28], whereas obstacle-type problems are considered in [6, 21, 24, 30]. Also in RB-methods, a-posteriori error bounds can be applied in adaptive basis enrichment schemes, such as the Greedy algorithm [37, 9, 15, 4].

**4.1. Preliminaries.** We start by introducing suitable functionals, which characterize the error of the reduced solution. First, we define the equality residual $r(\cdot;\boldsymbol{\mu}) \in V'$ by

$$r(v;\boldsymbol{\mu}) := f(v;\boldsymbol{\mu}) - a(u_N(\boldsymbol{\mu}),v;\boldsymbol{\mu}) - b(v,\lambda_N(\boldsymbol{\mu})), \quad v \in V. \tag{4.1}$$

Next, we quantify the inequality error by an inequality residual $s(\cdot;\boldsymbol{\mu}) \in W'$ with

$$s(\eta;\boldsymbol{\mu}) := b(u_N(\boldsymbol{\mu}),\eta) - g(\eta;\boldsymbol{\mu}), \quad \eta \in W. \tag{4.2}$$

The residual $r(\cdot;\boldsymbol{\mu})$ represents the right hand side of the error-equation, i.e.,

$$a(u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu}),v;\boldsymbol{\mu}) + b(v,\lambda(\boldsymbol{\mu}) - \lambda_N(\boldsymbol{\mu})) = r(v;\boldsymbol{\mu}), \quad v \in V. \tag{4.3}$$

Equality and inequality residuals can be quantified on $V_N$ and $M_N$ by

$$r(v_N;\boldsymbol{\mu}) = 0, \quad v_N \in V_N \quad \text{and} \quad s(\eta_N;\boldsymbol{\mu}) \le 0, \quad \eta_N \in M_N. \tag{4.4}$$

Moreover, we point out that for the special case of $u_N(\boldsymbol{\mu}) = u(\boldsymbol{\mu})$ and $\lambda_N(\boldsymbol{\mu}) = \lambda(\boldsymbol{\mu})$ we have $r(v;\boldsymbol{\mu}) = 0$ and $s(\eta;\boldsymbol{\mu}) \le 0$ for all $v \in V, \eta \in M$. Hence, the deviation from this equality/inequality gives information about the error and needs to be controlled. In order to quantify the error, we first introduce the Riesz-representers $v_r(\boldsymbol{\mu}) \in V, \eta_s(\boldsymbol{\mu}) \in W$ of the residuals

$$\langle v, v_r(\boldsymbol{\mu})\rangle_V = r(v;\boldsymbol{\mu}), \quad v \in V, \qquad \langle \eta, \eta_s(\boldsymbol{\mu})\rangle_W = s(\eta;\boldsymbol{\mu}), \quad \eta \in W.$$

Additionally, we denote $\tilde{\eta}_s(\boldsymbol{\mu}) \in W$ to be the Riesz-representer of the detailed inequality functional defined by

$$\langle \tilde{\eta}_s(\boldsymbol{\mu}), \eta \rangle_W = b(u(\boldsymbol{\mu}), \eta) - g(\eta; \boldsymbol{\mu}), \quad \eta \in W. \tag{4.5}$$

We require a projection $\pi : W \to M$ which we assume to be an orthogonal projection with respect to a scalar-product $\langle \cdot, \cdot \rangle_\pi$ on $W$ endowed with the induced norm $\|\eta\|_\pi := \sqrt{\langle \eta, \eta \rangle_\pi}$ being equivalent to the $W$-norm via $c_\pi \|\eta\|_W \leq \|\eta\|_\pi \leq C_\pi \|\eta\|_W$ for suitable constants $0 < c_\pi \leq C_\pi$. Moreover, we assume that $\pi$ satisfies the following properties:

$$\langle \eta - \pi(\eta), \eta' \rangle_W \leq 0, \quad \eta \in W, \eta' \in M, \tag{4.6}$$

$$\pi(\tilde{\eta}_s) = 0, \tag{4.7}$$

$$\langle \eta, \tilde{\eta}_s \rangle_\pi \leq 0, \quad \eta \in M. \tag{4.8}$$

For example, these conditions are met by standard orthogonal projections with $\langle \cdot, \cdot \rangle_\pi = \langle \cdot, \cdot \rangle_W$. Other problem specific choices will be given in Section 5. However, note that such a projection operator will in general be non-linear. We state a connection between the primal and dual error, which will be used for the a-posteriori error estimators.

LEMMA 4.1 (Primal/Dual Error Relation). *For any $\boldsymbol{\mu} \in \mathcal{P}$ the dual error can be bounded by the primal error as*

$$\|\lambda(\boldsymbol{\mu}) - \lambda_N(\boldsymbol{\mu})\|_W \leq \frac{1}{\beta} \left( \|r(\cdot; \boldsymbol{\mu})\|_{V'} + \gamma_a(\boldsymbol{\mu}) \|u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_V \right), \tag{4.9}$$

*Proof.* The inf-sup stability of $b(\cdot, \cdot)$ guarantees the existence of an $v \in V, v \neq 0$ such that with (4.3)

$$\begin{aligned} \beta \|v\|_V \|\lambda - \lambda_N\|_W &\leq b(v, \lambda - \lambda_N) = r(v) + a(u_N - u, v) \\ &\leq \|v\|_V \|r\|_{V'} + \gamma_a \|v\|_V \|u - u_N\|_V, \end{aligned}$$

and the result follows. □

**4.2. A-posteriori error estimators.** We can now present a-posteriori error bounds.

PROPOSITION 4.2 (Upper A-Posteriori Error Bound). *For any $\boldsymbol{\mu}$ we define the residual estimators*

$$\delta_r(\boldsymbol{\mu}) := \|r(\cdot; \boldsymbol{\mu})\|_{V'} = \|v_r(\boldsymbol{\mu})\|_V \tag{4.10}$$

$$\delta_{s1}(\boldsymbol{\mu}) := \|\pi(\eta_s(\boldsymbol{\mu}))\|_W \tag{4.11}$$

$$\delta_{s2}(\boldsymbol{\mu}) := \langle \lambda_N(\boldsymbol{\mu}), \pi(\eta_s(\boldsymbol{\mu})) \rangle_W. \tag{4.12}$$

*Then, the reduced basis errors can be bounded by*

$$\|u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_V \leq \Delta_u(\boldsymbol{\mu}) := c_1(\boldsymbol{\mu}) + \sqrt{c_1(\boldsymbol{\mu})^2 + c_2(\boldsymbol{\mu})}, \tag{4.13}$$

$$\|\lambda(\boldsymbol{\mu}) - \lambda_N(\boldsymbol{\mu})\|_W \leq \Delta_\lambda(\boldsymbol{\mu}) := \frac{1}{\beta} \left( \delta_r(\boldsymbol{\mu}) + \gamma_a(\boldsymbol{\mu}) \Delta_u(\boldsymbol{\mu}) \right), \tag{4.14}$$

*with constants*

$$c_1(\boldsymbol{\mu}) := \frac{1}{2\alpha(\boldsymbol{\mu})} \left( \delta_r(\boldsymbol{\mu}) + \frac{\delta_{s1}(\boldsymbol{\mu}) \gamma_a(\boldsymbol{\mu})}{\beta} \right), \quad c_2(\boldsymbol{\mu}) := \frac{1}{\alpha(\boldsymbol{\mu})} \left( \frac{\delta_{s1}(\boldsymbol{\mu}) \delta_r(\boldsymbol{\mu})}{\beta} + \delta_{s2}(\boldsymbol{\mu}) \right).$$

*Proof.* We note that (4.14) is a direct consequence of (4.13) and (4.9). Hence, it remains to show (4.13). Using coercivity, the error-equation (4.3), $b(u, \lambda_N - \lambda) \leq g(\lambda_N - \lambda)$, the definitions of the residuals and $\langle \lambda, \eta_s - \pi(\eta_s) \rangle_W \leq 0$ lead to

$$\begin{aligned}
\alpha \left\| u - u_N \right\|_V^2 &\leq a(u - u_N, u - u_N) = r(u - u_N) - b(u - u_N, \lambda - \lambda_N) \\
&\leq \delta_r \left\| u - u_N \right\|_V + b(u, \lambda_N - \lambda) + b(u_N, \lambda - \lambda_N) \\
&\leq \delta_r \left\| u - u_N \right\|_V + g(\lambda_N - \lambda) + s(\lambda - \lambda_N) + g(\lambda - \lambda_N) \\
&= \delta_r \left\| u - u_N \right\|_V + \langle \lambda, \pi(\eta_s) \rangle_W + \langle \lambda, \eta_s - \pi(\eta_s) \rangle_W \\
&\leq \delta_r \left\| u - u_N \right\|_V + \langle \lambda - \lambda_N, \pi(\eta_s) \rangle_W + \delta_{s2} \\
&\leq \delta_r \left\| u - u_N \right\|_V + \left\| \lambda - \lambda_N \right\|_W \delta_{s1} + \delta_{s2}.
\end{aligned}$$

Inserting (4.9) yields

$$\left\| u - u_N \right\|_V^2 - \frac{1}{\alpha} \left( \delta_r + \frac{\delta_{s1} \gamma_a}{\beta} \right) \left\| u - u_N \right\|_V - \frac{1}{\alpha} \left( \frac{\delta_{s1} \delta_r}{\beta} + \delta_{s2} \right) \leq 0.$$

Using the same argumentation as in previous proofs, i.e., bounding the error by the largest root of the corresponding quadratic polynomial, gives the bound (4.13). □

We briefly comment on the different terms in the upper bound. In the ideal case of $u_N(\boldsymbol{\mu}) = u(\boldsymbol{\mu}), \lambda_N(\boldsymbol{\mu}) = \lambda(\boldsymbol{\mu})$ we obtain $\delta_r(\boldsymbol{\mu}) = \delta_{s1}(\boldsymbol{\mu}) = \delta_{s2}(\boldsymbol{\mu}) = 0$ by (4.7). Then, the error bounds also yield $\Delta_u(\boldsymbol{\mu}) = \Delta_\lambda(\boldsymbol{\mu}) = 0$, identifying exact approximation a-posteriori, i.e., in the online-phase.

Let us assume an unconstrained case of $\lambda_N(\boldsymbol{\mu}) = 0$ and $b(u_N(\boldsymbol{\mu}), \eta) \leq g(\eta; \boldsymbol{\mu})$ for all $\eta \in M$. Then, it is easy to see that $\delta_{s1}(\boldsymbol{\mu}) = \delta_{s2}(\boldsymbol{\mu}) = 0$, and we perfectly reproduce the tight a-posteriori bound for elliptic equations of [35]: $\Delta_u(\boldsymbol{\mu}) = \frac{\|r\|_{V'}}{\alpha(\boldsymbol{\mu})}$.

Let us now comment on a difference of the above procedure to RB-error estimation for the Stokes problem [32]. There, the residual $s(\cdot; \boldsymbol{\mu})$ is involved in an equality constraint, and hence, the deviation can simply be computed by taking $\|s\|_{W'}$ as an error estimator component. In our case we cannot do that. We obviously would correctly penalize if $s(\eta; \boldsymbol{\mu}) > 0$ for some $\eta \in M$ as desired, but we would also penalize $s(\eta; \boldsymbol{\mu}) < 0$ which is not necessary. Our approach involving the projection operator is one way of adequately measuring the reduced basis inequality violation.

An interesting fact is that the equality residual can be observed to vanish in certain situations.

PROPOSITION 4.3 (Vanishing Equality Residual). *Let $Q_a = 1$ and the reduced primal space $V_N$ be chosen as* (2.6). *Then, we obtain $r(\cdot; \boldsymbol{\mu}) = 0$.*

*Proof.* Recalling $a(\cdot, \cdot; \boldsymbol{\mu}) = \theta_a^1(\boldsymbol{\mu}) a^1(\cdot, \cdot)$, then the definition of the equality residual and of the operators $A$ and $B^1$ yield

$$r(v; \boldsymbol{\mu}) = \sum_{q=1}^{Q_f} \theta_f^q(\boldsymbol{\mu}) a^1(A f^q, v) - \theta_a^1(\boldsymbol{\mu}) a^1(u_N(\boldsymbol{\mu}), v) - a^1(B^1 \lambda_N(\boldsymbol{\mu}), v) = a^1(z, v),$$

with $z := \sum_{q=1}^{Q_f} \theta_f^q(\boldsymbol{\mu}) A f^q - \theta_a^1(\boldsymbol{\mu}) u_N(\boldsymbol{\mu}) - B^1 \lambda_N(\boldsymbol{\mu}) \in V_N$. As $r(\cdot; \boldsymbol{\mu})$ vanishes on $V_N$, we obtain

$$0 = r(z; \boldsymbol{\mu}) = a^1(z, z; \boldsymbol{\mu}) \geq \alpha(\boldsymbol{\mu}) \|z\|^2.$$

Therefore, $z = 0$, and consequently $r(v) = a^1(0, v) = 0, v \in V$. □

The residuals do not only provide an upper bound for the error in Prop. 4.2 but also yield lower bounds.

PROPOSITION 4.4 (Lower A-Posteriori Error Bounds). *For any $\boldsymbol{\mu} \in \mathcal{P}$ the following lower bounds for the reduction error hold:*

$$\delta_r(\boldsymbol{\mu}) \leq \gamma_a(\boldsymbol{\mu}) \left\| u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu}) \right\|_V + \gamma_b \left\| \lambda(\boldsymbol{\mu}) - \lambda_N(\boldsymbol{\mu}) \right\|_W , \tag{4.15}$$

$$\delta_{s1}(\boldsymbol{\mu}) \leq \frac{\gamma_b C_\pi}{c_\pi} \left\| u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu}) \right\|_V , \tag{4.16}$$

$$\delta_{s2}(\boldsymbol{\mu}) \leq \frac{\bar{\gamma}_\lambda \gamma_b C_\pi}{c_\pi} \left\| u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu}) \right\|_V . \tag{4.17}$$

*Proof.* Thanks to the error-equation (4.3), we obtain with the Riesz-representation $v_r \in V$ of $r \in V'$

$$\begin{aligned} \gamma_a(\boldsymbol{\mu}) \left\| u - u_N \right\|_V \left\| v_r \right\|_V + \gamma_b \left\| \lambda - \lambda_N \right\|_W \left\| v_r \right\|_V &\geq a(u - u_N, v_r) + b(v_r, \lambda - \lambda_N) \\ &= r(v_r) = \langle v_r, v_r \rangle_V = \left\| v_r \right\|_V^2 , \end{aligned}$$

which gives (4.15).

We note that orthogonal projections on convex sets have Lipschitz-constant one. Thus (4.7) and the norm-equivalence on $V$ guarantee

$$\delta_{s1} = \|\pi(\eta_s) - \pi(\tilde{\eta}_s)\|_W \leq \frac{1}{c_\pi} \|\pi(\eta_s) - \pi(\tilde{\eta}_s)\|_\pi \leq \frac{1}{c_\pi} \|\eta_s - \tilde{\eta}_s\|_\pi \leq \frac{C_\pi}{c_\pi} \|\eta_s - \tilde{\eta}_s\|_W .$$

For the last term we continue with

$$\begin{aligned} \|\eta_s - \tilde{\eta}_s\|_W &= \sup_{\eta \in W} \frac{b(u_N, \eta) - g(\eta) - b(u, \eta) + g(\eta)}{\|\eta\|_W} = \sup_{\eta \in W} \frac{b(u_N - u, \eta)}{\|\eta\|_W} \\ &\leq \sup_{\eta \in W} \frac{\gamma_b \|u_N - u\|_V \|\eta\|_W}{\|\eta\|_W} \leq \gamma_b \|u_N - u\|_V \end{aligned} \tag{4.18}$$

from which we can conclude (4.16).

The bound (4.17) follows from (4.16) and the observation that $\delta_{s2} \leq \bar{\gamma}_\lambda \|\pi(\eta_s)\|_W$. □

REMARK 4.5. *A closer look on the lower and upper bounds given in Propositions 4.4 and 4.2 reveals a gap. In the upper bound the term $\sqrt{\delta_{s2}}$ enters, whereas in the lower bound the term $\delta_{s2}$ appears. This results from the variational inequality setting.*

REMARK 4.6. *We point out that the lower bounds strongly depend on the constants $c_\pi$ and $C_\pi$. If chosing $\langle \cdot, \cdot \rangle_\pi := \langle \cdot, \cdot \rangle_W$ these are independent of the discretization $C_\pi = c_\pi = 1$, but then the evaluation of $\pi$ has the same complexity as the original problem. To reduce the computational cost, alternative scalar products can be selected, but then $C_\pi / c_\pi$ possibly depends on the high dimension $H_W$ and is possibly quite large. In that case, the lower bounds are not very informative.*

**5. Implementational aspects.** In this section, we comment on computational aspects that are required for the realization of the reduced scheme and our experiments. Quite often $W$ is the dual space of a finite dimensional $V$ which simplifies some computations, as the dimensions coincide, and the inner products are related.

**5.1. Solution of the detailed and reduced problem.** In practice it is quite common to choose the basis of $W$ and the bilinear form $b$ so that the matrix associated to $b(\cdot, \cdot)$ has a simple diagonal form, this property is often referred as biorthogonality. This makes the

use of a primal-dual active set strategy for the high-dimensional problem associated with the snapshots computationally attractive cf. [18, 20, 26].

For the reduced problem (2.12–2.15), different solvers such as interior point methods, SQP and penalty techniques can be applied. Here, we do not require a special biorthogonalization of the basis of $V_N$ with respect to the vectors spanning $M_N$. Hence, we accept that the matrix $\overline{B}_N$ is now possibly dense. It is easy to show that the solution of $DSP_N(\boldsymbol{\mu})$ is equivalent to the solution of a constrained convex quadratic optimization problem, cf. the variational minimization problem (1.1).

REMARK 5.1 (Discrete Quadratic Program $DQP_N(\boldsymbol{\mu})$). *The solution vectors* $(\overline{u}_N(\boldsymbol{\mu}), \overline{\lambda}_N(\boldsymbol{\mu}))$ *of* $DSP_N(\boldsymbol{\mu})$ *are equivalently obtained as solution of a constrained convex quadratic optimization problem. In particular,* $\overline{u}_N(\boldsymbol{\mu})$ *is the unique minimizer of*

$$\min \frac{1}{2}\overline{v}_N^T \overline{A}_N(\boldsymbol{\mu})\overline{v}_N - \overline{f}_N(\boldsymbol{\mu})^T \overline{v}_N \tag{5.1}$$

$$s.t. \qquad \overline{B}_N^T \overline{v}_N \leq \overline{g}_N(\boldsymbol{\mu}), \tag{5.2}$$

*and* $\overline{\lambda}_N(\boldsymbol{\mu})$ *is a non-negative vector of Lagrange multipliers in the optimum, i.e.,* $\overline{A}_N\overline{u}_N + \overline{B}_N\overline{\lambda}_N = \overline{f}_N$ *and* $\overline{\lambda}_{N,i} = 0$ *if* $(\overline{B}_N^T\overline{u}_N - \overline{g}_N)_i < 0$.

Hence, any off-the-shelf quadratic optimization routine using Lagrange multipliers can be used to compute $\overline{u}_N(\boldsymbol{\mu})$ and $\overline{\lambda}_N(\boldsymbol{\mu})$. Note that we did not assume linear independence of the dual snapshots in the construction of the reduced cone. Therefore, the inequality system may have linearly dependent rows. As mentioned above, this may lead to a non-unique vector $\overline{\lambda}_N$, but a unique function $\lambda_N(\boldsymbol{\mu})$.

**5.2. Choice of the projection for the case $W = V'$.** For the special case of $W = V'$ we could use the $\langle\cdot,\cdot\rangle_W$ inner product and the corresponding orthogonal projection for the error estimators. Here, for computational simplicity we use a different projection. First, we make some further assumptions on the problem. We suppose that $W$ is endowed with a basis $\{\chi_i\}_i$ such that $M$ can be written as:

$$M = \left\{ \sum_{i=1}^{H_W} \underline{\beta}_i \chi_i \,|\, \underline{\beta}_i \geq 0 \right\}.$$

For obstacle type problems this is natural to hold in the finite element setting, since $\{\chi_i\}_{i=1}^{H_W}$ can be chosen as a dual basis of $\{\psi_i\}_{i=1}^{H_V}$. We then have $H_V = H_W =: H$ with the inner product matrices $\underline{M}^V := (\langle\psi_i,\psi_j\rangle_V)_{i,j=1}^H$ and $\underline{M}^W := (\langle\chi_i,\chi_j\rangle_W)_{i,j=1}^H$. These matrices allow to compute inner products and norms as required in the a-posteriori error bounds. For instance, for any $\eta, \eta' \in W$ with coefficient vectors $\underline{\eta}, \underline{\eta}' \in \mathbb{R}^H$ we have $\langle\eta,\eta'\rangle_W = \underline{\eta}^T \underline{M}^W \underline{\eta}'$. In the case of $W = V'$ one can even verify, that $\underline{M}^W = (\underline{M}^V)^{-1}$. If $\underline{M}^V$ is an $M$-matrix, which is typical for finite element discretization spaces, one can even guarantee, that $\underline{M}^W$ has non-negative entries. Details on these aspects are provided in Section A.2. We define $\pi : W \to M$ as follows:

$$\pi(\eta) = \sum_{i=1}^{H} \underline{\pi}_i \chi_i, \quad \underline{\pi} = (\underline{\pi}_i)_{i=1}^H := (\underline{M}^W)^{-1}[\underline{M}^W \underline{\eta}]_+ \tag{5.3}$$

with $[\cdot]_+$ denoting the component-wise positive part of a vector. An element $\eta \in W$ is in $M$ if and only if $\underline{\eta}_i \geq 0$ for all indices $i$. Hence, with $\underline{M}^V$ being an $M$-matrix, we obtain $\underline{M}^W \underline{\eta} \geq 0$ for $\underline{\eta} \geq 0$. Thus for $\eta \in M$, we have $[\underline{M}^W \underline{\eta}]_+ = \underline{M}^W \underline{\eta}$ and $\pi(\eta) = \eta$. We

define an alternative inner product on $W$ by

$$\langle \eta, \eta' \rangle_\pi := \underline{\eta}^T (\underline{M}^W)^2 \underline{\eta}'. \tag{5.4}$$

Symmetry, bilinearity and positive definiteness are obviously inherited from the inner product of $W$. One can verify, that $\pi$ from (5.3) is the orthogonal projection on $M$ with respect to this inner product (5.4), and that it satisfies the assumptions (4.6)–(4.8). Details on this are given in the appendix in Section A.6. Therefore, $\pi$ can be applied in the a-posteriori error bounds. We note that the dense matrix $\underline{M}^W$ which enters formally in the definition of $\pi$ is not required to evaluate $\delta_{s1}$ and $\delta_{s2}$. By definition (4.11), (4.12), we have

$$\begin{aligned}
\delta_{s1}^2 &= \left( (\underline{M}^W)^{-1} [\underline{M}^W \underline{\eta}_s(\boldsymbol{\mu})]_+ \right)^T \underline{M}^W (\underline{M}^W)^{-1} [\underline{M}^W \underline{\eta}_s(\boldsymbol{\mu})]_+ \\
&= [\underline{B}^T \underline{u}_N(\boldsymbol{\mu}) - \underline{g}(\boldsymbol{\mu})]_+^T \underline{M}^V [\underline{B}^T \underline{u}_N(\boldsymbol{\mu}) - \underline{g}(\boldsymbol{\mu})]_+, \\
\delta_{s2} &= \underline{\lambda}_N(\boldsymbol{\mu})^T \underline{M}^W (\underline{M}^W)^{-1} [\underline{M}^W \underline{\eta}_s(\boldsymbol{\mu})]_+ \\
&= \underline{\lambda}_N(\boldsymbol{\mu})^T [\underline{B}^T \underline{u}_N(\boldsymbol{\mu}) - \underline{g}(\boldsymbol{\mu})]_+.
\end{aligned}$$

We observe that this equivalent representation of $\delta_{s1}$ and $\delta_{s2}$ shows directly that these contributions are equal to zero if the reduced solution $u_N(\boldsymbol{\mu}) \in X(\boldsymbol{\mu})$, and thus these terms can be regarded as a measure for the violation of the constraint.

**5.3. Offline/online decomposition of error estimators.** We now comment on the offline/online decomposition of the a-posteriori error estimators $\Delta_u(\boldsymbol{\mu}), \Delta_\lambda(\boldsymbol{\mu})$. In contrast to other RB-methods, the error estimators are currently not yet fully decomposable in an offline/online fashion due to the non-linear projection. But parts of the required components can be decomposed. We assume the parameter-dependent constants $\gamma_a(\boldsymbol{\mu})$, $\alpha(\boldsymbol{\mu})$ and the inf-sup constant $\beta$ or upper/lower bounds thereof to be available (e.g., by the uniform bounds $\bar{\gamma}_a$, $\bar{\alpha}$) or computable (e.g., by the successive constraint method [22]) in the online-phase. The remaining relevant quantities, which then require an offline/online decomposition are the bounds $\delta_r(\boldsymbol{\mu}), \delta_{s1}(\boldsymbol{\mu})$ and $\delta_{s2}(\boldsymbol{\mu})$.

The equality residual bound is $\delta_r(\boldsymbol{\mu}) = \|r(\cdot, \boldsymbol{\mu})\|_{V'}$. Recalling the separable parameter representations of $a(\cdot, \cdot; \boldsymbol{\mu})$ and $f(\cdot; \boldsymbol{\mu})$ from (2.1), (2.2), we compute in the offline-phase the Riesz-representers $v_a^{q,i}, v_f^q, v_b^i \in V$ of the parameter-independent linear forms $a^q(\varphi_i, \cdot)$, $f^q(\cdot)$, $b(\cdot, \lambda(\boldsymbol{\mu}_i))$ for all $q, i$ in suitable ranges. Then, in view of the definition of the residual (4.1) and the coefficient representations $u_N(\boldsymbol{\mu}) = \sum_{i=1}^{N_V} \overline{u}_{N,i} \varphi_i$, $\lambda_N(\boldsymbol{\mu}) = \sum_{i=1}^{N_S} \overline{\lambda}_{N,i} \lambda(\boldsymbol{\mu}_i)$, we see that the Riesz-representative $v_r \in V$ of $r(\cdot; \boldsymbol{\mu})$, satisfies

$$v_r(\boldsymbol{\mu}) = \sum_{q=1}^{Q_f} \theta_f^q(\boldsymbol{\mu}) v_f^q - \sum_{q=1}^{Q_a} \sum_{i=1}^{N_V} \theta_a^q(\boldsymbol{\mu}) \overline{u}_{N,i} v_a^{q,i} - \sum_{i=1}^{N_S} \overline{\lambda}_{N,i} v_b^i =: \sum_{q=1}^{Q_r} \theta_r^q(\boldsymbol{\mu}) v_r^q \tag{5.5}$$

by choosing an arbitrary enumeration $v_r^1, \ldots, v_r^{Q_r}$ of the vectors $v_a^{q,i}, v_f^q, v_b^i$ with suitable coefficient functions $\theta_r^q(\boldsymbol{\mu})$. In particular we set $Q_r := Q_f + N_V Q_a + N_S$. Therefore, in the offline-phase, we compute the inner-product matrix

$$\overline{K}_r := \left( \left\langle v_r^q, v_r^{q'} \right\rangle_V \right)_{q,q'=1}^{Q_r} \in \mathbb{R}^{Q_r \times Q_r}.$$

In the online-phase we determine the vector $\overline{\theta}_r(\boldsymbol{\mu}) = (\theta_r^1(\boldsymbol{\mu}), \ldots, \theta_r^{Q_r}(\boldsymbol{\mu}))^T$ and obtain the desired residual norm as

$$\delta_r(\boldsymbol{\mu}) = \|v_r(\boldsymbol{\mu})\|_V = \sqrt{\overline{\theta}_r(\boldsymbol{\mu})^T \overline{K}_r \overline{\theta}_r(\boldsymbol{\mu})}.$$

Here, the estimator component $\delta_r(\boldsymbol{\mu})$ is fully decomposed into an offline- and online-phase. In particular, for any new $\boldsymbol{\mu}$, this quantity can be computed in $O(Q_r^2)$ and is independent of the detailed dimensions $H_V$ or $H_W$.

Concerning the inequality residual components $\delta_{s1}$ and $\delta_{s2}$, we cannot provide a full offline/online decomposition. In particular, the online-phase will still depend on operations of the dimensionality $H_W$. However, a partial decomposition is possible and in the special case of Subsection 5.2, we observe that $\underline{B}^T \underline{u}_N(\boldsymbol{\mu}) - \underline{g}(\boldsymbol{\mu})$ can be computed as a trivial embedding of a linear operator of reduced cost and some sparse high dimensional matrix vector operations.

**6. Experiments.** In this section, we test our approach on some obstacle type examples in 1D. We consider an "elastic" rope hanging over a surface that may cause contact. Our setting is as follows: The domain $\Omega = (0, 1)$ is discretized with a uniform mesh of step size $\Delta x := 1/K$ for $K \in \mathbb{N}$. For the discrete function space $V$, we use standard conforming nodal first order finite elements $V := \{v \in H_0^1(\Omega) | v_{|[x_k, x_{k+1}]} \in P_1, k = 0, \ldots, K-1\}$ of dimension $H_V = H_W = H := K - 1 = 200$ with $x_k := k\Delta x$. We associate the basis function $\psi_i \in V$ with its Lagrange node $x_i \in \Omega$, i.e., $\psi_i(x_j) = \delta_{ij}, i, j = 1, \ldots, H$. The discretization of the Lagrange multipliers is performed using a dual finite element basis of $W = V'$. The cone $M$ is defined by:

$$M = \left\{ \sum_{i=1}^{H} \underline{\eta}_i \chi_i,\ \underline{\eta}_i \geq 0 \right\}.$$

**6.1. Two-dimensional parameter space.** In a first example, the parameter domain is $\mathcal{P} := [10, 50] \times [-0.05, 0.5] \subset \mathbb{R}^2$ and the parameter vector consists of a pair of parameters $\boldsymbol{\mu} = (\mu_1, \mu_2)$. The bilinear forms $a$ and $b$ are given by:

$$\begin{aligned} a(u, v; \boldsymbol{\mu}) &:= \int_\Omega \nu(\boldsymbol{\mu})(x) \nabla u(x) \cdot \nabla v(x) dx\ , \quad v, u \in V \\ b(u, \eta) &:= -\eta(u), \quad u \in V, \eta \in W \end{aligned}$$

with $\nu(\boldsymbol{\mu})(x) = \mu_1 Ind_{[0,1/2]}(x) + \nu_0 Ind_{[1/2,1]}(x)$ which characterizes the "elasticity" of the rope. Here we denote by $Ind_\Gamma$ the characteristic function of an interval $\Gamma$. We use the value $\nu_0 = 30$.

As we use a dual finite element basis, the matrix $\underline{B} = (b(\psi_i, \chi_j))_{i,j=1}^{H,H}$ corresponding to the bilinear form $b(\cdot, \cdot)$ is a multiple of the identity. The right-hand side functional $f$ corresponds to gravity and is defined by:

$$f(v; \boldsymbol{\mu}) = \int_\Omega \gamma(x; \boldsymbol{\mu}) v(x) dx, \quad v \in V$$

where $\gamma(x; \boldsymbol{\mu}) := \gamma_0 = -1$. The obstacle is given as $g(\eta; \boldsymbol{\mu}) = \sum_{i=1}^{H} \underline{\eta}_i h(x_i; \boldsymbol{\mu})$ for $\eta = \sum_{i=1}^{H} \underline{\eta}_i \chi_i$ with a parameter-dependent barrier function

$$h(x; \boldsymbol{\mu}) = -0.2\,(\sin(\pi x) - \sin(3\pi x)) - 0.5 + \mu_2 x.$$

Examples of solutions are represented in Fig. 6.1.

In the offline-phase, we compute the snapshots, i.e., a set of detailed solutions of our obstacle problem corresponding to various values $\boldsymbol{\mu}_i$ of the parameter $\boldsymbol{\mu}$. The reduced basis $\{\varphi_i\}_{i=1}^{N_V}$ is taken as an orthonormal family of $V_N = V_N^{(2)}$ given by Eq. (2.5). In our test, we consider $N_S = 25$ values of $\boldsymbol{\mu}$ taken on a $5 \times 5$-grid composed of uniformly distributed

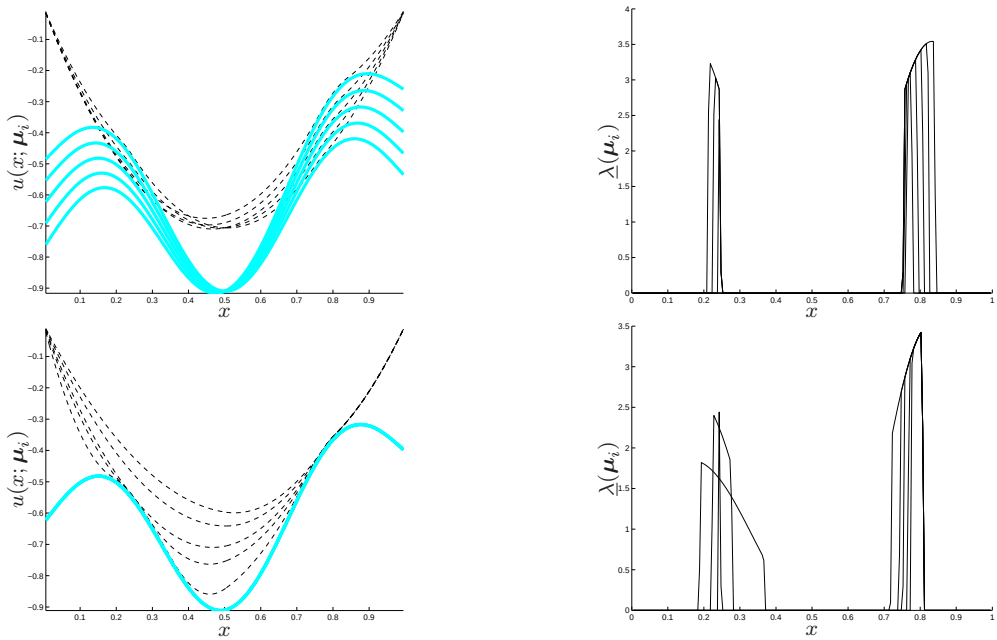


FIGURE 6.1. *Snapshots and obstacles corresponding to the 1D rope example. Top: the RB-parameters are associated to the obstacle with $\mu_2$ values uniformly distributed in $[-0.05, 0.5]$ and constant elasticity $\mu_1 = 30$. Bottom: the RB-parameters are associated to the elasticity with $\mu_1$ values uniformly distributed in $[10, 50]$ and fixed obstacle for $\mu_2 = 0.225$. Left: primal solution, Right: corresponding Lagrange multipliers. The $u$ snapshots are represented with black dashed lines, the obstacles with solid light blue lines, the $\lambda$ snapshots with solid black lines.*

points. Thanks to a standard singular value decomposition routine, we extract from the family $\{u(\boldsymbol{\mu}_i), B\lambda(\boldsymbol{\mu}_i)\}$ an orthonormal basis using the eigenvectors of the corresponding correlation matrix associated to eigenvalues larger than a tolerance $Tol = 10^{-8}$. As a consequence, $N_V = 29$ vectors are considered as reduced basis for the primal variable, see Fig. 6.2. An

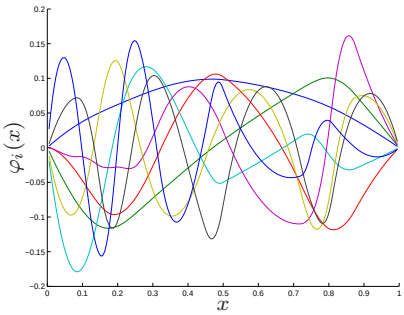


FIGURE 6.2. *Eight first vectors of the reduced basis $\{\varphi_i\}_{i=1}^{N_V}$ forming $V_N$.*

example of the exact and reduced solutions corresponding to the (non-snapshot)-parameter $\boldsymbol{\mu}^\star = (21.7157, 0.1111)^T$ is depicted in Fig. 6.3. We see that the reduced and exact primal solutions $u_N(\boldsymbol{\mu})$ and $u(\boldsymbol{\mu})$ show almost no difference. However, we observe a difference between the reduced and exact dual solution $\lambda(\boldsymbol{\mu})$ and $\lambda_N(\boldsymbol{\mu})$. This is confirmed quantitatively as $\|u(\boldsymbol{\mu}^*) - u_N(\boldsymbol{\mu}^*)\|_V = 0.012961$ and $\|\lambda(\boldsymbol{\mu}^*) - \lambda_N(\boldsymbol{\mu}^*)\|_W = 0.295079$ corresponding to the norms $\|u(\boldsymbol{\mu}^*)\|_V = 1.662406$ and $\|\lambda(\boldsymbol{\mu}^*)\|_W = 18.814451$. Note, that in particular with

$\alpha(\boldsymbol{\mu}^*) = 21.715729$, $\beta_N = 1$, $\gamma(\boldsymbol{\mu}^*) = 30.000000$ and $\|r\|_V = 0.001111$ the inequality (4.9) can be numerically verified, despite the visual contrary impression.

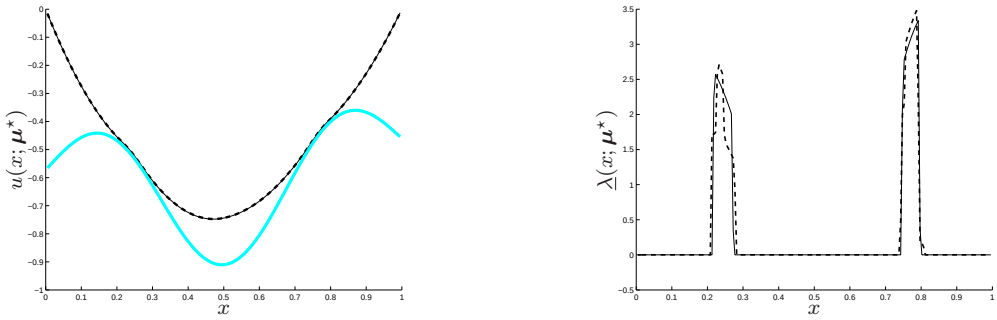


FIGURE 6.3. *Exact and reduced solutions for a non-snapshot parameter* $\boldsymbol{\mu}^* = (21.7157, 0.1111)^T$ . *Left: primal solutions. Right: dual solutions. Solid line: exact solutions, dashed line: reduced solutions.*

**6.2. Efficiency of the method.** Our second example aims at quantifying the efficiency of our method. For a more expressive demonstration of the effects, we keep the previous model but consider a simpler example with a fixed barrier function $h(x) = 5x - 10$. The parameter vector is a scalar parameter $\boldsymbol{\mu} = (\mu)$ that parameterizes the elasticity $\nu(\boldsymbol{\mu}) = \mu$ through the relation

$$\nu(\boldsymbol{\mu})(x) = \mu.$$

The parameter domain is $\mathcal{P} := [0.2, 2] \subset \mathbb{R}$. We keep the functionals $f(\cdot;\boldsymbol{\mu})$ and $b(\cdot,\cdot)$ as before. We first investigate the error decay with growing number of snapshots $N_S$. For this, we construct reduced bases corresponding to an equidistant choice of $N_S$ parameters from $\mathcal{P}$. For the primal basis we include the supremizers, hence compute $V_N = V_N^{(2)}$ according to (2.5). For each reduced model, we determine the maximum error over a test-set of parameters given as the vertices of a uniform 10x10 grid. The results are depicted in Fig. 6.4. This numerical example exhibits asymptotically an exponential decrease with respect to the number of snapshots used to build the reduced basis.

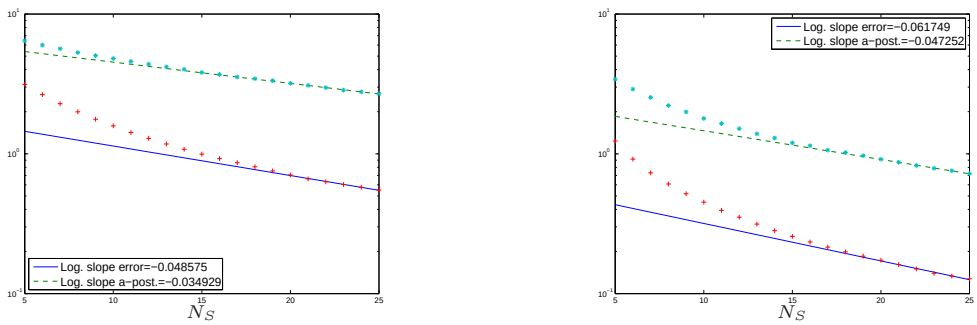


FIGURE 6.4. *Semilog plot of the errors* $\max_{\boldsymbol{\mu}\in\mathcal{F}} \left(\|u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_V\right)$ *(left) and* $\max_{\boldsymbol{\mu}\in\mathcal{F}} \left(\|\lambda(\boldsymbol{\mu}) - \lambda_N(\boldsymbol{\mu})\|_W\right)$ *(right) and of the a-posteriori estimators* $\max_{\boldsymbol{\mu}\in\mathcal{F}} \Delta_u(\boldsymbol{\mu})$ *and* $\max_{\boldsymbol{\mu}\in\mathcal{F}} \Delta_\lambda(\boldsymbol{\mu})$ *with respect to the number of snapshots* $N_S$. *The numerical values are represented by crosses for the errors and by stars for the estimators. A linear regression of the last 5 values is represented by the solid line for the error and by the dashed line for the a-posteriori estimator. The slopes of the regressions are indicated in the boxes on the top of the picture.*

| $N_S$ | $\beta_N$ with supremizers | $log_{10}(\beta_N)$ without supremizers |
|---|---|---|
| 5 | 1.000000 | -2.568645 |
| 10 | 1.000000 | -5.631370 |
| 15 | 1.000000 | -8.623529 |
| 20 | 1.000000 | -11.515277 |
| 25 | 1.000000 | -14.094935 |

TABLE 6.1
*Inf-sup constant for various values of $N_S$.*

To highlight the importance of adding the supremizers in the definition of $V_N$, we compare our results with $V_N = V_N^{(1)} = \mathrm{span}\{u(\mu_i)\}$, i.e., the space of snapshots without enrichment. In a first test, we compute for both settings the inf-sup constant $\beta_N$ for various values of $N_S$, see the appendix for details on these computations. In contrast to the previous example, $a(\cdot, \cdot; \boldsymbol{\mu})$ is proportional to the scalar product on $V$ and $Q_a = 1$. From Remark 2.4 it then follows that (2.5)–(2.7) yields the same space and that $V_N$ compared to the naive choice is at most enriched by one element. The results are given in Table 6.1. In this example, the supremizers greatly improve the stability of the method, and we verify the theoretical findings of $\beta_N \geq \beta = 1$ for $V_N = V_N^{(2)}$.

Secondly, we compare the computational cost of the solver. In our case, we use the open-source octave code *'qsolve.m'* [1], based on a standard active set method as described in [11]. Numerically, the number of iterations required to solve the reduced problem during the online-phase increases linearly with respect to $N$ in both cases, see Fig. 6.5. However, in the case of (2.4), the slope is roughly two times bigger than in the case of (2.5).

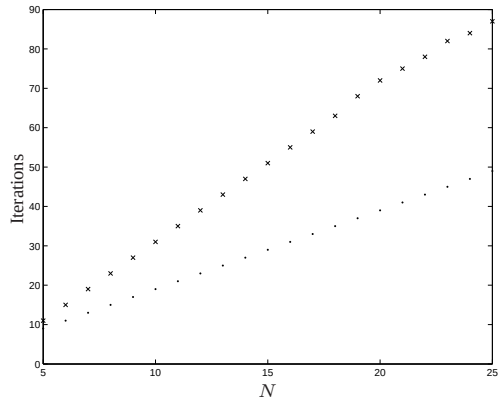


FIGURE 6.5. *Effect of the inclusion of supremizers. Number of iterations required to solve the reduced problem. Dots: $V_N = V_N^{(2)}$ with supremizers, cf.* (2.5)*; crosses: $V_N = V_N^{(1)}$ without supremizers, cf.* (2.4)*.*

Let us emphasize that in this example the inclusion of the supremizer functions does not improve the accuracy of the reduced solutions. But as indicated by Table 6.1 and in Fig. 6.5, the inf-sup constants and the computational time indeed do improve by this space extension.

In a third test, we illustrate the performance of our reduced basis scheme compared to a parameter-wise computation of the detailed solution. We evaluate the actual acceleration due to our reduced-basis method, we measure the respective computation times of the offline- and

[1] `http://docs.ufrmd.dauphine.fr/optiNum/mat/qsolve.txt`.

online-phases and compare them to the time required by a standard method. More precisely, we define:

- $t_1$: Computation time for one fine scale solution, i.e., time required to determine $(u(\mu_i), \lambda(\mu_i))$. In our case, this time is closely related to the performance of the primal-dual active set strategy and depends on the dimension of $V$ and $W$.
- $t_2$: Offline-phase computation time, i.e., time corresponding to the snapshot computation, determination of the supremizers, the orthogonal basis $(\varphi_i)_{i=1}^{N_V}$, the matrix components $\overline{A}_N^q$ and $\overline{B}_N$ and the vector components $\overline{f}_N^q$ and $\overline{g}_N^q$.
- $t_3$: Online-phase computation time, i.e., time required to assemble the matrix $\overline{A}_N(\boldsymbol{\mu})$, vectors $\overline{f}_N(\boldsymbol{\mu}), \overline{g}_N(\boldsymbol{\mu})$ and to solve the reduced saddle point problem in order to determine $(u_N(\mu), \lambda_N(\mu))$. In our case, this time is closely related to the performance of the applied quadratic optimization routine.

For the example we consider in this section, we choose $N_S = 10$ as the number of snapshots that we use to build the reduced basis. One finds the mean values (over approximately 150 measurements obtained with an Intel Core2 Duo 2.6 GHz processor): $t_1 = 0.095283$, $t_2 = 0.956744$, $t_3 = 0.005902$. Suppose now that one has to solve the problem $L$ times. The times required by our reduced-basis method scales as $t_2 + L \cdot t_3$, whereas a standard method of $L$ detailed simulations requires a time that scales as $L \cdot t_1$. Asymptotically, we observe an acceleration factor of about 16.

**6.3. Reduced basis generation with the Greedy algorithm.** In a last test, we use the Greedy algorithm [37, 9, 15, 4] for adaptive basis generation based on the a-posteriori estimators $\Delta_u$ and $\Delta_\lambda$ described in Section 4 to compute a relevant basis. In this procedure, we sequentially enrich the current basis with the snapshot corresponding to the parameter value that maximizes an error indicator $\Delta(\boldsymbol{\mu})$, e.g., $\Delta(\boldsymbol{\mu}) := \omega_u \Delta_u + \omega_\lambda \Delta_\lambda$ with some weights $\omega_u, \omega_\lambda \in \mathbb{R}^+$. The precise algorithm is the following:

ALGORITHM 1. *Input: $N_{max} \in \mathbb{N}$, accuracy $\varepsilon_{tol}$, initial sample $\boldsymbol{\mu}_1$, training set $\mathcal{T} \subset \mathcal{P}$.*

1. *set $k := 0$*
2. *repeat*
   (a) *set k:= k+1*
   (b) *Compute the snapshot $(u(\boldsymbol{\mu}_k), \lambda(\boldsymbol{\mu}_k))$.*
   (c) *define $V_N$ and $M_N$ corresponding to the snapshots $\{(u(\boldsymbol{\mu}_i), \lambda(\boldsymbol{\mu}_i))\}_{i=1}^k$.*
   (d) *define $\varepsilon := \max_{\boldsymbol{\mu} \in \mathcal{T}} \Delta(\boldsymbol{\mu})$ and $\boldsymbol{\mu}_{k+1} := \arg\max_{\boldsymbol{\mu} \in \mathcal{T}} \Delta(\boldsymbol{\mu})$*
3. *until $\varepsilon < \varepsilon_{tol}$ or $k \geq N_{max}$*
4. *set $N_S := k$ and return reduced basis $\{\varphi_i\}_{i=1}^{N_V}$ and snapshots $\{\lambda(\boldsymbol{\mu}_i)\}_{i=1}^{N_S}$*

Only $N_S$ detailed problems are solved (during step 2b); all other solutions deal with reduced problems of small dimension. Additional manipulations of high dimensional vectors are done only $N_S$ times in Step 2c) to compute the orthogonal basis and in Step 2d) when computing the projection $\pi$ required to evaluate the a-posteriori estimator $\Delta_u(\boldsymbol{\mu}) + \Delta_\lambda(\boldsymbol{\mu})$. In our case $V' = W$ and this projection consists in applying component-wise the cheap function $\pi_c(x) = \min(0, x)$, cf. Section 5.3.

In our test, we use the example of Section 6.1. As initial value, we set $\boldsymbol{\mu}_1 = (1.1, -0.1)$ which is the center value of the parameter interval. We set the weights for the error estimator $\Delta(\boldsymbol{\mu})$ as $\omega_u = \omega_\lambda = 1$. We compare the following different basis generation procedures. First, $\mathcal{B}_u$ denotes the reduced basis associated with the coarse grid $\mathcal{C}$ of $5 \times 5 = 25$ uniformly distributed points in $\mathcal{P}$. Second, $\mathcal{B}_\Delta$ stands for the reduced basis resulting from the Greedy algorithm Algorithm 1 using the a-posteriori error estimator $\Delta(\boldsymbol{\mu})$ with $N_{max} = 25$. The training grid $\mathcal{T}$ we use is composed of $17 \times 17 = 289$ uniformly distributed points. Finally, $\mathcal{B}_e$ denotes the basis obtained by the Greedy algorithm using the true error as selection measure, i.e. $\Delta(\boldsymbol{\mu})$ is replaced with $\varepsilon_N(\boldsymbol{\mu}) := \omega_u e_u(\boldsymbol{\mu}) + \omega_\lambda e_\lambda(\boldsymbol{\mu})$ with $e_u = \|u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_V$

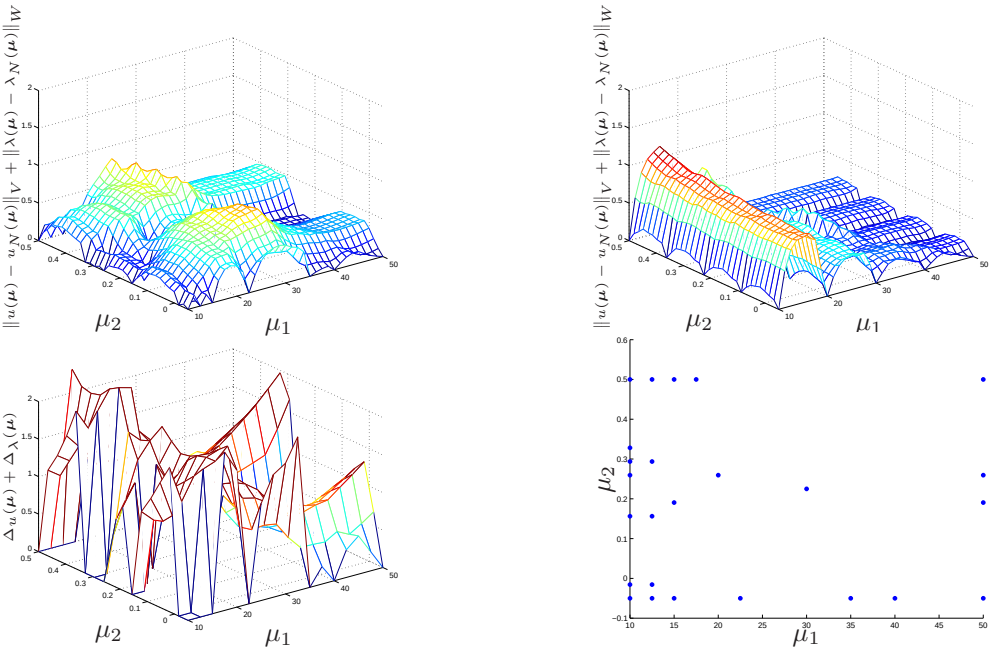


FIGURE 6.6. *Top: numerical values of the error* $\varepsilon_N = \|u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_V + \|\lambda(\boldsymbol{\mu}) - \lambda_N(\boldsymbol{\mu})\|_W$ *when using the reduced basis obtained by Algorithm 1 (left) and the reduced basis obtained taking uniform coarse grid (right). Both plots are using the fine test grid* $\mathcal{F}$*. Bottom: numerical values of the a-posteriori estimator* $\Delta_u(\boldsymbol{\mu}) + \Delta_\lambda(\boldsymbol{\mu})$ *obtained by Algorithm 1 plotted on the train grid* $\mathcal{T}$ *(left) and selected values* $\boldsymbol{\mu}_1, \cdots, \boldsymbol{\mu}_{N_S}$ *in* $\mathcal{T}$ *(right). Note that both errors and a-posteriori estimates are zero when* $\boldsymbol{\mu}$ *corresponds to a snapshot's parameter value, see Prop. 3.1 and Prop. 4.4.*

and $e_\lambda = \|\lambda(\boldsymbol{\mu}) - \lambda_N(\boldsymbol{\mu})\|_W$ The latter basis generation procedure is computationally very expensive and not practical, but it is included as a reference method. In all cases we obtain $N_S = 25$ snapshots and determine the error $\varepsilon_N$ on a fine test grid $\mathcal{F}$ of $33 \times 33 = 1089$ uniformly distributed points. The results are depicted in Fig. 6.6. The primal part $\{\varphi_i\}_{i=1}^{N_V}$ of the reduced basis is of dimensions $N_V = 29$ for $\mathcal{B}_u$ and of dimension $N_V = 40$ for $\mathcal{B}_\Delta$. Hence we have a small overcost of the greedy algorithm. Nevertheless, the maximal error is significantly reduced when using the greedy algorithm. This is shown in Table 6.2, where we report the maximum of the component errors $e_u$, $e_\lambda$, the error $\varepsilon_N$. In this example, the

| | $\max_{\boldsymbol{\mu}\in\mathcal{F}}\{e_u\}$ | $\max_{\boldsymbol{\mu}\in\mathcal{F}}\{e_\lambda\}$ | $\max_{\boldsymbol{\mu}\in\mathcal{F}}\{\varepsilon_N\}$ | $N_V$ |
|---|---|---|---|---|
| $\mathcal{B}_u$ | 0.059944 | 1.092896 | 1.146332 | 29 |
| $\mathcal{B}_\Delta$ | 0.028021 | 0.836206 | 0.864084 | 40 |
| $\mathcal{B}_e$ | 0.031795 | 0.485100 | 0.502336 | 42 |

TABLE 6.2

*Size of basis and maximal test errors obtained when using the Greedy algorithm, or a uniform grid for basis generation.*

accuracy $\varepsilon_N = 1.146332$ is obtained using the uniform grid $\mathcal{C}$ with $N_S = 25$ snapshots and a reduced primal space of dimension $N_V = 29$. When using the greedy algorithm with the a-posteriori error estimators, the same accuracy is obtained with roughly half the number $N_S = 14$ of required snapshots, and the corresponding basis $\mathcal{B}_\Delta$ contains in this case $N_V = 26$ vectors. The Greedy algorithm used with the true error as selection criteria, once again with $N_S = 25$ snapshots, results in a basis $\mathcal{B}_e$ of $N_V = 42$ vectors. The decrease of

the error $\varepsilon_N$ and the a-posteriori estimate $\max_{\boldsymbol{\mu}\in\mathcal{T}}\{\Delta_u(\boldsymbol{\mu}) + \Delta_\lambda(\boldsymbol{\mu})\}$ along the iterations of the construction of $\mathcal{B}_\Delta$ is shown in Fig. 6.7. The results indicate that the greedy algorithm

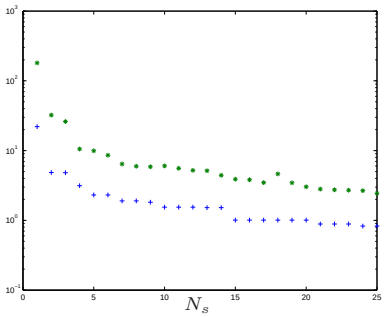


FIGURE 6.7. *Evolution of* $\max_{\boldsymbol{\mu}\in\mathcal{T}}\{\varepsilon_N\}$ *(blue crosses) and* $\max_{\boldsymbol{\mu}\in\mathcal{T}}\{\omega_u\Delta_u(\boldsymbol{\mu}) + \omega_\lambda\Delta_\lambda(\boldsymbol{\mu})\}$ *(green stars) with respect to the number of snapshots* $N_S$ *used to assemble the reduced problem.*

indeed is a procedure leading to compact bases, also in our case of variational inequalities. The a-posteriori error estimators indeed seem to be good substitutes for the true errors in the greedy algorithm as the dimensions and accuracies of $\mathcal{B}_\Delta$ and $\mathcal{B}_e$ do not deviate too much.

**7. Conclusion.** We have presented a reduced basis scheme for parametrized variational inequalities. We tackled the conceptional difficulty of including inequality constraints in the RB-scheme and analysis. This required ingredients from VI-theory for a-posteriori error analysis by non-linear projection operators. The full offline-online decomposition of the reduced scheme could be obtained as in the unconstrained case. The a-posteriori error estimation also yields rigorous and efficient lower and upper bounds. In contrast to the unconstrained case, the a-posteriori error estimators are not yet fully decomposable in an offline-online fashion, as the online-phase still slightly depends on the high dimension. This is subject to further investigations.

Development of RB-methods is widely driven by solving more and more complex parametrized PDEs. When viewed as parametrized variational minimization problems, another class of possible extensions is adding more complex constraints to the optimization problem. Such extensions and reduction schemes may very likely leave the comfortable realm of Galerkin projections, but gain conceptional advantages. We only mention the possibilities of Reduced Basis schemes with respect to conservation of physical properties, such as mass and non-negativity. Such formulations are expected to be interesting future research directions.

REFERENCES


[1] M. Ainsworth and J. Oden. *A posteriori error estimation in finite element analysis*. Chichester: Wiley, 2000.

[2] M. Ainsworth, J. Oden, and C. Lee. Local a posteriori error estimators for variational inequalities. *Numer. Methods Partial Differ. Equations*, 9:23–33, 1993.

[3] M. Barrault, Y. Maday, N.C. Nguyen, and A.T. Patera. An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *C. R. Math. Acad. Sci. Paris Series I*, 339:667–672, 2004.

[4] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence rates for greedy algorithms in reduced basis methods. *IGPM Report, RWTH Aachen*, 310, 2010.

[5] V. Bostan, W. Han, and B. Reddy. A posteriori error estimation and adaptive solution of elliptic variational inequalities of the second kind. *Appl. Numer. Math.*, 52:13–38, 2005.

[6] D. Braess. A posteriori error estimators for obstacle problems – another look. *Numer. Math.*, 101:523–549, 2005.

[7] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. Springer, New York, 1991.

[8] G. Duvaut and J. Lions. *Inequalities in mechanics and physics, Translated form the French (1972) by C.W. John*. Springer, Berlin, 1976.

[9] J. L. Eftang, A. T. Patera, and E. M. Rønquist. An hp certified reduced basis method for parametrized elliptic partial differential equations. *SIAM J. Sci Comp*, 32(6):3170–3200, 2010.

[10] I. Ekeland and R. Temam. *Convex analysis and variational problems (Studies in mathematics and its applications; 1)*. North-Holland Publ. Comp., 1976.

[11] P.E. Gill, W. Murray, and M.H. Wright. *Numerical Linear Algebra and Optimization, Vol. 1*. Addison Wesley, 1991.

[12] R. Glowinski. *Numerical methods for nonlinear variational problems*. Springer-Verlag, Berlin Heidelberg, 1984.

[13] R. Glowinski, J. Lions, and R. Trémolières. *Numerical analysis of variational inequalities*. North-Holland, Amsterdam, 1981.

[14] M.A. Grepl and A.T. Patera. A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations. *M2AN, Math. Model. Numer. Anal.*, 39(1):157–181, 2005.

[15] B. Haasdonk, M. Dihlmann, and M. Ohlberger. A training set and multiple bases generation approach for parametrized model reduction based on adaptive grids in parameter space. Technical report, University of Stuttgart, 2010. Accepted by MCMDS.

[16] B. Haasdonk and M. Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *M2AN, Math. Model. Numer. Anal.*, 42(2):277–302, 2008.

[17] B. Haasdonk and M. Ohlberger. Reduced basis method for explicit finite volume approximations of nonlinear conservation laws. In *Hyperbolic problems: theory, numerics and applications*, pages 605–614, 2009.

[18] C. Hager and B. Wohlmuth. Semismooth Newton methods for variational problems with inequality constraints. *GAMM-Mitt.*, 33:8–24, 2010.

[19] W. Han. *A posteriori error analysis via duality theory. With applications in modeling and numerical approximations*. Springer, New York, 2005.

[20] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semi-smooth Newton method. *SIAM Journal on Optimization*, 13:865–888, 2002.

[21] R. Hoppe and R. Kornhuber. Adaptive multilevel methods for obstacle problems. *SIAM J. Numer. Anal.*, 31:301–323, 1994.

[22] D.B.P. Huynh, G. Rozza, S. Sen, and A.T. Patera. A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants. *C. R. Math. Acad. Sci. Paris Series I*, 345:473–478, 2007.

[23] K. Ito and K. Kunisch. *Lagrange multiplier approach to variational problems and applications*. Society for Industrial and Applied Mathematics (SIAM), Philadelpia, PA, 2008.

[24] C. Johnson. Adaptive finite element methods for the obstacle problem. *Math. Models Methods Appl. Sci.*, 2:483–487, 1992.

[25] C. Kanzow and C. Geiger. *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer, 2002.

[26] T. Karkkainnen, K. Kunisch, and P. Tarvainen. Primal-dual active set methods for obstacle problems. *J. Optimization Theory and Appl.*, 119:499–533, 2003.

[27] D. Kinderlehrer and G. Stampacchia. *An introduction to variational inequalities and their applications*. SIAM, Philadelphia, 2000.

[28] W. Liu and N. Yan. A posteriori error estimators for a class of variational inequalities. *J. Sci. Comput.*, 15:361–393, 2000.

[29] R. Nicolaides. Existence, uniqueness and approximation for generalized saddle point problems. *SIAM J. Numer. Anal.*, 19:349–357, 1982.

[30] R. Nochetto, K. Siebert, and A. Veeser. Pointwise a posteriori error control for elliptic obstacle problems. *Numer. Math.*, 95:163–195, 2003.

[31] S. Repin. *A posteriori estimates for partial differential equations*. Radon Series on Computational and Applied Mathematics. Walter de Gruyter GmbH & Co. KG, 2008.

[32] D.V. Rovas. *Reduced-Basis Output Bound Methods for Parametrized Partial Differential Equations*. PhD thesis, MIT, Cambridge, MA, 2003.

[33] D.V. Rovas, L. Machiels, and Y. Maday. Reduced basis output bound methods for parabolic problems. *IMA J. Numer. Anal.*, 26(3):423–445, 2006.

[34] G. Rozza. *Shape design by optimal flow control and reduced basis techniques: Applications to bypass configurations in haemodynamics*. PhD thesis, École Polytechnique Fédérale de Lausanne, November 2005.

[35] G. Rozza, D.B.P. Huynh, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: application to transport and continuum mechanics. *Arch. Comput. Meth. Eng.*, 15(3):229–275, 2008.

[36] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Wiley–Teubner Series Advances in Numerical Mathematics. Wiley–Teubner, Chichester, Stuttgart, 1996.

[37] 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. In *In Proceedings of 16th AIAA computational fluid dynamics conference*, 2003. Paper 2003-3847.

**Appendix A. Details and Proofs.**

In this addendum we give details on various aspects and proofs for certain statements. Some of the following results can be found in standard-text-books on variational inequalities, as listed in the introduction. For completeness, we reproduce them in the finite dimensional context. These insights were very helpful for us, and may also be of use for the interested reader.

**A.1. Inf-Sup Implications.** The following characterization of inf-sup stability in finite dimensional spaces can be stated:

LEMMA A.1 (Inf-sup implications). *Let $b(v,\eta)$ be inf-sup-stable on $V \times W$, i.e.,*

$$\inf_{\eta\neq 0}\sup_{v\neq 0} b(v,\eta)/(\|v\|_V \|\eta\|_W) \geq \beta > 0.$$

*We can define the closed nonempty set $V_0 := \{v \in V | b(v,\eta) = 0 \forall \eta \in W\}$. We define two sets of equation systems for $f \in V'$ and $g \in W'$, respectively:*

$$b(v,\eta) = f(v), \quad v \in V \tag{A.1}$$

$$b(v,\eta) = g(\eta), \quad \eta \in W \tag{A.2}$$

*Then holds*

*i) Eqn.* (A.1) *has at most one solution $\eta \in W$.*
*ii) Eqn.* (A.2) *has at least one solution $v \in V$.*
*iii) Eqn.* (A.1) *has a unique solution $\eta \in W$, if $f(v_0) = 0$ for $v_0 \in V_0$*
*iv) Eqn.* (A.2) *has a unique solution $v \in V_0^\perp$.*

*Proof.* The bilinear form $b$ can be represented by a linear continuous operator $B : W \to V$ as $b(v,\eta) = \langle v, B\eta\rangle_V$. (This can be obtained by using Riesz-representation theorem for fixed $\eta$, which uniquely defines each $B\eta$ and then showing, that $B$ is linear and continuous.) For i): Let $\eta, \eta' \in W$ be two different solutions, then we have

$$b(v,\eta-\eta') = f(v) - f(v) = 0$$

hence, $\sup_v b(v,\eta-\eta') = 0$ which is a contradiction to the inf-sup stability condition as $\eta - \eta' \neq 0$. In particular this implies that $B$ is injective.

For ii): Let $B^{ad} : V \to W$ be the adjoint of $B$ and $g_r \in W$ be the Riesz-representative of $g$. ¿From i) we know, that $B$ is injective. Then $B^{ad}$ is surjective, such that for $g_r \in W$ there exists a $v \in V$ with $B^{ad}v = g_r$. This satisfies

$$b(v,\eta) = \langle v, B\eta\rangle_V = \left\langle B^{ad}v, \eta\right\rangle_W = \langle g_r, \eta\rangle_W = g(\eta).$$

For iii): It remains to show existence, uniqueness then follows from i). First we note, that

$$V_0 = \{v \in V : \langle v, B\eta\rangle_V = 0 \forall \eta \in W\} = R(B)^\perp$$

where we denote the range of $B$ as $R(B)$. The condition that $f$ vanishes on $V_0$ implies for its Riesz-representative $f_r$:

$$0 = f(v_0) = \langle f_r, v_0\rangle_V, \quad \forall v_0 \in V_0.$$

Hence $f_r \in V_0^\perp = (R(B)^\perp)^\perp = R(B)$. Therefore there exists $\eta \in W$ with $B\eta = f_r$. This satisfies for all $v \in V$

$$b(v,\eta) = \langle v, B\eta\rangle_V = \langle v, f_r\rangle_V = f(v)$$

For iv): Let $\bar{v} \in V$ be a solution (existence due to ii)) and $P : V \to V_0$ be the orthogonal projection, which exists, as $V_0$ is nonempty and closed. Then we define $v := \bar{v} - P\bar{v}$. We then first note, that also $v$ is a solution, as for all $\eta \in W$ we obtain

$$b(v, \eta) = b(\bar{v}, \eta) - b(P\bar{v}, \eta) = g(\eta) - 0$$

as $P\bar{v} \in V_0$. As $v$ is the projection error of $\bar{v}$, it is orthogonal to the projection space, hence $v \in V_0^\perp$. Uniqueness follows by considering a second solution $\tilde{v} \in V_0^\perp$, hence $v - \tilde{v} \in V_0^\perp$. Let $\hat{v} \in V_0^\perp = R(B)$ then there exists $\hat{\eta} \in W$ such that $B\hat{\eta} = \hat{v}$. Then we observe

$$\langle v - \tilde{v}, \hat{v} \rangle_V = \langle v - \tilde{v}, B\hat{\eta} \rangle_V = b(v - \tilde{v}, \hat{\eta}) = 0$$

as both $v$ and $\tilde{v}$ are solutions. Hence $v - \tilde{v} \in V_0$. Since $v - \tilde{v} \in V_0^\perp$ we conclude $v = \tilde{v}$. □

**A.2. Computation of scalar products for $W = V'$.** The scalar product on $W$ plays a crucial role in the computation of the Riesz representers or of the residual norms required in the a-posteriori error terms. Here, we give details on the facts that were mentioned in Section 5.2. Recall that we have denoted a basis $\{\psi_i\}_{i=1}^{H_V}$ of $V$ and $\{\chi_i\}_{i=1}^{H_W}$ of $W$ where $H_V$ and $H_W$ are the dimensions of the high- but finite-dimensional discrete spaces, e.g., standard finite element spaces. Then, any function $v = \sum_{i=1}^{H_V} \underline{v}_i \psi_i \in V$ is characterized by its coefficient vector $\underline{v} = (\underline{v}_i)_{i=1}^{H_V}$. The inner product matrix $\underline{M}^V := (\langle \psi_i, \psi_j \rangle_V)_{i,j=1}^{H_V}$ then allows to compute scalar products between any pairs of elements $v, u \in V$ by $\langle v, u \rangle_V = \underline{v}^T \underline{M}^V \underline{u}$. If the space $W$ is independent of $V$, one can similarly obtain and compute scalar products in $W$ by the given inner product matrix $\underline{M}^W$ and computing $\langle \eta, \eta' \rangle_W = \underline{\eta}^T \underline{M}^W \underline{\eta}'$ with $\underline{\eta} \in \mathbb{R}^{H_W}$ being the coefficient vector of the function $\eta \in W$. In practice, $W = V'$ is frequently chosen with $\{\chi_i\}_{i=1}^{H_W}$ being a dual basis. This implies equal dimension $H_W = H_V =: H$ and $\chi_i(\psi_j) = \delta_{ij}$ for all basis functions $\psi_j \in V, \chi_i \in W$ for $i, j = 1, \ldots, H$. Then, for any given $\eta = \sum_{i=1}^{H} \underline{\eta}_i \chi_i \in W$ with coefficient vector $\underline{\eta} = (\underline{\eta}_i)_{i=1}^{H} \in \mathbb{R}^H$ its Riesz representative $R(\eta) = \sum_{i=1}^{H} \underline{r}_i \psi_i \in V$ has the coefficient vector $\underline{r} = (\underline{r}_i)_{i=1}^{H} = (\underline{M}^V)^{-1} \underline{\eta}$, as for any $v = \sum_{i=1}^{H} \underline{v}_i \psi_i \in V$ there holds

$$\underline{r}^T \underline{M}^V \underline{v} = \langle R(\eta), v \rangle_V = \eta(v) = \underline{\eta}^T \underline{v} = ((\underline{M}^V)^{-1} \underline{\eta})^T \underline{M}^V \underline{v}.$$

The Riesz theorem then allows to compute scalar products for any $\eta_1, \eta_2 \in W$ by

$$\langle \eta_1, \eta_2 \rangle_W = \langle R(\eta_1), R(\eta_2) \rangle_V = ((\underline{M}^V)^{-1} \underline{\eta}_1)^T \underline{M}^V (\underline{M}^V)^{-1} \underline{\eta}_2 = \underline{\eta}_1^T (\underline{M}^V)^{-1} \underline{\eta}_2.$$

Hence, we obtain the particular relation of the inner product matrices of $V$ and $W$: $\underline{M}^W = (\underline{M}^V)^{-1}$, which can be used in practice. Consequently, if $\underline{M}^V$ is an $M$-matrix which is the case for a low order conforming finite element stiffness matrix for the Laplace operator, $\underline{M}^W$ has non-negative entries, and thus $\langle \chi_i, \chi_j \rangle_W \geq 0, \quad i, j = 1, \ldots, H_W$.

**A.3. Computation of the Inf-Sup Constant.** We now provide a characterization of the inf-sup constant, which was used for its computation in the experiments section.

LEMMA A.2 (Characterization of $\beta_N$). *The inf-sup constant $\beta_N$ of the reduced problem $SP_N(\boldsymbol{\mu})$ is the smallest singular value of the matrix*

$$\widetilde{B}_N := (\overline{M}_N^V)^{-1/2} \overline{B}_N (\overline{M}_N^W)^{-1/2},$$

*where $\overline{B}_N$ is given in* (2.9) *and we define*

$$\overline{M}_N^V := \left( \langle \varphi_i, \varphi_j \rangle_V \right)_{i,j=1}^{N_V} \tag{A.3}$$

$$\overline{M}_N^W := \left( \langle \xi_i, \xi_j \rangle_W \right)_{i,j=1}^{N_W}. \tag{A.4}$$

*Proof.* We define the operator $B_N : W_N \to V_N$ by pointwise application of the Riesz-representation theorem $\langle v_N, B_N \eta_N \rangle_V = b(v_N, \eta_N)$ for all $v_N \in V_N, \eta_N \in W_N$. We denote $\overline{v}_N = (v_{N,i})_{i=1}^{N_V} \in \mathbb{R}^{N_V}$ and $\overline{\eta}_N = (\eta_{N,i})_{i=1}^{N_W} \in \mathbb{R}^{N_W}$ the coordinate vector of an element $v_N = \sum_{i=1}^{N_V} v_{N,i}\varphi_i \in V_N$ and $\eta_N = \sum_{i=1}^{N_W} \eta_{N,i}\xi_i \in W_N$. We set $\widetilde{v}_N = (\overline{M}_N^V)^{1/2})\overline{v}_N$ and $\widetilde{\eta}_N = (\overline{M}_N^W)^{1/2}\overline{\eta}_N$ as suitably coordinate transformed vectors. We then obtain

$$\begin{aligned}
\beta_N &= \inf_{\eta_N \in W_N} \sup_{v_N \in V_N} \frac{b(v_N, \eta_N)}{\|v_N\|_V \|\eta_N\|_W} \\
&= \inf_{\overline{\eta}_N \in \mathbb{R}^{N_W}} \sup_{\overline{v}_N \in \mathbb{R}^{N_V}} \frac{\overline{v}_N^T \overline{B}_N \overline{\eta}_N}{\sqrt{\overline{v}_N^T \overline{M}_N^V \overline{v}_N}\sqrt{\overline{\eta}_N^T \overline{M}_N^W \overline{\eta}_N}} \\
&= \inf_{\widetilde{\eta}_N \in \mathbb{R}^{N_W}} \sup_{\widetilde{v}_N \in \mathbb{R}^{N_V}} \frac{\widetilde{v}_N^T ((\overline{M}_N^V)^{-1/2})^T \overline{B}_N (\overline{M}_N^W)^{-1/2} \widetilde{\eta}_N}{\sqrt{\widetilde{v}_N^T \widetilde{v}_N}\sqrt{\widetilde{\eta}_N^T \widetilde{\eta}_N}} \\
&= \inf_{\widetilde{\eta}_N \in \mathbb{R}^{N_W}} \sup_{\widetilde{v}_N \in \mathbb{R}^{N_V}} \frac{\widetilde{v}_N^T \widetilde{B}_N \widetilde{\eta}_N}{\sqrt{\widetilde{v}_N^T \widetilde{v}_N}\sqrt{\widetilde{\eta}_N^T \widetilde{\eta}_N}}
\end{aligned}$$

For fixed $\widetilde{\eta}_N$ the supremum of the quotient is obtained with Cauchy-Schwartz choosing the vector $\widetilde{v}_N := \widetilde{B}_N \widetilde{\eta}_N$. Hence

$$\beta_N = \inf_{\widetilde{\eta}_N \in \mathbb{R}^{N_W}} \frac{(\widetilde{B}_N \widetilde{\eta}_N)^T \widetilde{B}_N \widetilde{\eta}_N}{\sqrt{(\widetilde{B}_N \widetilde{\eta}_N)^T \widetilde{B}_N \widetilde{\eta}_N}\sqrt{\widetilde{\eta}_N^T \widetilde{\eta}_N}} = \inf_{\widetilde{\eta}_N \in \mathbb{R}^{N_W}} \frac{\sqrt{(\widetilde{B}_N \widetilde{\eta}_N)^T \widetilde{B}_N \widetilde{\eta}_N}}{\sqrt{\widetilde{\eta}_N^T \widetilde{\eta}_N}}.$$

We obtain the Rayleigh-quotient

$$\beta_N^2 = \inf_{\widetilde{\eta}_N \in \mathbb{R}^{N_W}} \frac{\widetilde{\eta}_N^T (\widetilde{B}_N^T \widetilde{B}_N) \widetilde{\eta}_N}{\widetilde{\eta}_N^T \widetilde{\eta}_N},$$

which implies, that $\beta_N^2$ is the smallest eigenvalue of $\widetilde{B}_N^T \widetilde{B}_N$ or, equivalently, $\beta_N$ is the smallest singular value of $\widetilde{B}_N$. □

**A.4. Existence and Uniqueness of Saddle Point Problem.** We first define an alternative formulation:

DEFINITION A.3 (Weak Form of Parametrized Variational Inequality $VI(\boldsymbol{\mu})$). *For $\boldsymbol{\mu} \in \mathcal{P}$ find $u(\boldsymbol{\mu}) \in X$ such that*

$$a(u, v - u; \boldsymbol{\mu}) \geq f(v - u; \boldsymbol{\mu}), \quad v \in X$$

Under the above assumptions, due to the finite dimensionality, [25] is applicable and ensures existence of a unique solution.

The following proposition states the equivalence of $VI(\boldsymbol{\mu})$ and $SP(\boldsymbol{\mu})$. The existence and uniqueness of solutions for $SP(\boldsymbol{\mu})$ then follows from the existence and uniqueness of $VI(\boldsymbol{\mu})$.

LEMMA A.4 (Equivalence of $SP(\boldsymbol{\mu})$ and $VI(\boldsymbol{\mu})$). *Let $VI(\boldsymbol{\mu})$ and $SP(\boldsymbol{\mu})$ be given with $X = X(\boldsymbol{\mu}) = \{v \in V : b(v, \eta) \leq g(\eta; \boldsymbol{\mu}), \eta \in M\}$.*

*i) If $u$ is the solution of $VI(\boldsymbol{\mu})$, then there exists a unique $\lambda \in M$ such that $(u, \lambda)$ is the solution of $SP(\boldsymbol{\mu})$.*

*ii) If $(u, \lambda)$ is the solution of $SP(\boldsymbol{\mu})$, then $u$ is the solution of $VI(\boldsymbol{\mu})$.*

*Proof.* Concerning part ii): Assume $(u, \lambda)$ solves $SP(\boldsymbol{\mu})$, then for $\tau \in M$ we find that $\eta = \lambda + \tau \in M$, hence the inequality of $SP(\boldsymbol{\mu})$ yields $b(u, \tau) \leq g(\tau; \boldsymbol{\mu})$ which guarantees $u \in X$. For $\eta = 2\lambda \in M$ and $\eta = 0 \in M$ we get

$$b(u, \lambda) \leq g(\lambda; \boldsymbol{\mu}) \qquad b(u, -\lambda) \leq g(-\lambda; \boldsymbol{\mu})$$

and hence $b(u, \lambda) = g(\lambda; \boldsymbol{\mu})$. Choose $v \in X$, then $v - u \in V$ is an allowed test function in the first equation of $SP(\boldsymbol{\mu})$, which gives

$$\begin{aligned} a(u, v - u; \boldsymbol{\mu}) &= f(v - u; \boldsymbol{\mu}) - b(v - u, \lambda) \\ &= f(v - u; \boldsymbol{\mu}) - b(v, \lambda) + g(\lambda; \boldsymbol{\mu}) \\ &\geq f(v - u; \boldsymbol{\mu}). \end{aligned}$$

The last line follows, as $\lambda \in M$ and $v \in X$.

Concerning part i):

Let $u \in X$ be the unique solution of $VI(\boldsymbol{\mu})$ then $\lambda \in W$ is chosen as the unique solution of the variational equation

$$b(v, \lambda) = f(v; \boldsymbol{\mu}) - a(u, v; \boldsymbol{\mu}), \quad v \in V \tag{A.5}$$

Existence and uniqueness follow from Lemma A.1 iii) as $b$ is inf-sup stable, $f(\cdot; \boldsymbol{\mu}) - a(u, \cdot; \boldsymbol{\mu}) \in V'$ and for all $v_0 \in V_0 = \{v \in V : b(v, \eta) = 0 \forall \eta \in W\}$ holds $f(v_0; \boldsymbol{\mu}) - a(u, v_0; \boldsymbol{\mu}) = 0$. The latter is easy to see as for such $v_0$ we find that $v := u \pm v_0 \in X$ and using these as test functions we conclude from $VI(\boldsymbol{\mu})$ that $a(u, v_0; \boldsymbol{\mu}) = f(v_0; \boldsymbol{\mu})$.

We now briefly prove the following equivalence:

$$M = M^*(\boldsymbol{\mu}) = \{\eta \in W : b(v, \eta) \leq g(\eta; \boldsymbol{\mu}), v \in X\} \tag{A.6}$$

Assume $\eta \in M$ then $b(v, \eta) \leq g(\eta; \boldsymbol{\mu})$ for all $v \in X$ due to the definition. Hence $\eta \in M^*(\boldsymbol{\mu})$ and we get $M \subset M^*(\boldsymbol{\mu})$. For the other direction, we assume $\eta \notin M$ and lead that to a contradiction. As $M$ is a closed convex cone, the separation-theorem implies the existence of a $h \in W'$ with $h(\eta) > \sup_{\eta' \in M} h(\eta') =: \alpha$. As $M$ is a cone, $\alpha = 0$. As $b$ is inf-sup stable on $V \times W$ and $g + h \in W'$, Lemma A.1 ii) gives the existence of a solution $x \in V$ such that

$$b(x, \eta') = h(\eta') + g(\eta'; \boldsymbol{\mu}) \quad , \eta' \in W.$$

In particular for $\eta' \in M$ we get $b(x, \eta') \leq 0 + g(\eta'; \boldsymbol{\mu})$ hence $x \in X$. But for the special choice of $\eta$ above, we have $b(x, \eta) = h(\eta) + g(\eta; \boldsymbol{\mu}) > 0 + g(\eta; \boldsymbol{\mu})$ and therefore $\eta \notin M^*(\boldsymbol{\mu})$. This proves $M^*(\boldsymbol{\mu}) \subset M$ which completes statement (A.6)

Now, let $v \in X$, which gives

$$\begin{aligned} b(v, \lambda) &= f(v; \boldsymbol{\mu}) - a(u, v; \boldsymbol{\mu}) \\ &= f(v - u; \boldsymbol{\mu}) - a(u, v - u; \boldsymbol{\mu}) + f(u; \boldsymbol{\mu}) - a(u, u; \boldsymbol{\mu}) \\ &\leq 0 + b(u, \lambda) \\ &\leq g(\lambda; \boldsymbol{\mu}) \end{aligned}$$

Where the second-last line follows from $u$ solving $VI(\boldsymbol{\mu})$ and (A.5), and the last line from the fact that $u \in X$. Consequently $\lambda \in M$ due to (A.6).

According to Lemma A.1 iv) there exists a unique solution $w \in V_0^{\perp}$ satisfying

$$b(w, \eta) = g(\eta; \boldsymbol{\mu}) \quad \eta \in W$$

hence in particular $w \in X$ and also $2u - w \in X$ since

$$b(2u - w, \eta) = 2b(u, \eta) - b(w, \eta) \leq 2g(\eta; \boldsymbol{\mu}) - g(\eta; \boldsymbol{\mu}) = g(\eta; \boldsymbol{\mu})$$

Using $w$ and $2u - w$ as test-functions in $VI(\boldsymbol{\mu})$ yields

$$a(u, w - u; \boldsymbol{\mu}) = f(w - u; \boldsymbol{\mu}).$$

The definition of $\lambda$ then yields

$$b(w - u, \lambda) = 0,$$

and thus

$$b(u, \lambda) = b(w, \lambda) = g(\lambda; \boldsymbol{\mu}).$$

For $\eta \in M$, $u$ satisfies

$$b(u, \eta - \lambda) = b(u, \eta) - g(\lambda; \boldsymbol{\mu}) \leq g(\eta - \lambda; \boldsymbol{\mu})$$

and thus $(u, \lambda)$ solves the saddle point problem.
□

**A.5. Equivalence of $SP_N(\boldsymbol{\mu})$ and $DSP_N(\boldsymbol{\mu})$.** We provide detailed arguments for the equivalence statement of Lemma 2.5. Note, that the same reasoning can be applied for reformulating the original high-dimensional FEM problem $SP(\boldsymbol{\mu})$ into a an algebraic formulation.

*Proof.* $SP_N(\boldsymbol{\mu}) \Rightarrow DSP_N(\boldsymbol{\mu})$ : Let $(u_N, \lambda_N) \in V_N \times M_N$ be the unique solution of $SP_N(\boldsymbol{\mu})$ with coefficient vectors $\underline{u}_N, \underline{\lambda}_N$. Then, (2.13) is obviously satisfied by definition of the convex cone $M_N$. Using $\varphi_i$ as test function in $SP_N(\boldsymbol{\mu})$ yields the $i$-th line of (2.12):

$$\begin{aligned}
(\overline{A}_N)_{(i,\cdot)}\overline{u}_N + (\overline{B}_N)_{(i,\cdot)}\overline{\lambda}_N &= \sum_{j=1}^{N_V} a(\varphi_j, \varphi_i; \boldsymbol{\mu}) u_{N,j} + \sum_{j=1}^{N_W} b(\varphi_i, \lambda(\boldsymbol{\mu}_j))\lambda_{N,j} \\
&= a\left(\sum_{j=1}^{N_V} u_{N,j}\varphi_j, \varphi_i; \boldsymbol{\mu}\right) + b\left(\varphi_i, \sum_{j=1}^{N_W} \lambda_{N,j}\right)\lambda(\boldsymbol{\mu}_j) \\
&= a(u_N, \varphi_i; \boldsymbol{\mu}) + b(\varphi_i, \lambda_N) = f(\varphi_i; \boldsymbol{\mu}).
\end{aligned}$$

For proving (2.14) we use $\eta_N := s\lambda(\boldsymbol{\mu}_i), s \in \mathbb{R}^+$ as test function in the inequality of $SP_N(\boldsymbol{\mu})$:

$$b(u_N, s\lambda(\boldsymbol{\mu}_i) - \lambda_N) \leq g(s\lambda(\boldsymbol{\mu}_i) - \lambda_N; \boldsymbol{\mu}).$$

This can be rearranged to

$$s(b(u_N, \lambda(\boldsymbol{\mu}_i)) - g(\lambda(\boldsymbol{\mu}_i); \boldsymbol{\mu})) \leq b(u_N, \lambda_N) - g(\lambda_N; \boldsymbol{\mu}) \tag{A.7}$$

The right hand side (RHS) is independent of $s$, therefore, the left hand side (LHS) cannot be positive. If the LHS was positive, for sufficiently large $s$ it would exceed the RHS, which would be a contradiction. Hence, we have

$$\begin{aligned}
0 \geq b(u_N, \lambda(\boldsymbol{\mu}_i)) - g(\lambda(\boldsymbol{\mu}_i); \boldsymbol{\mu}) &= \sum_{j=1}^{N_V} b(\varphi_j, \lambda(\boldsymbol{\mu}_i))u_{N,j} - g(\lambda(\boldsymbol{\mu}_i)) \\
= \overline{u}_N^T(\overline{B}_N)_{(\cdot,i)} - g(\lambda(\boldsymbol{\mu}_i)) &= (\overline{B}_N)_{(i,\cdot)}\overline{u}_N - g(\lambda(\boldsymbol{\mu}_i))
\end{aligned}$$

which is the $i$-th row of (2.14).

Identical reasoning as above with $\eta_N := s\lambda_N, s \in \mathbb{R}^+$ implies, that $b(u_N, \lambda_N) - g(\lambda_N; \boldsymbol{\mu}) \leq 0$. But setting $\eta_N = 0$ as test-function in the inequality of $SP_N(\boldsymbol{\mu})$ gives $b(u_N, \lambda_N) - g(\lambda_N; \boldsymbol{\mu}) \geq 0$. Hence it must be equal 0 and we conclude

$$0 = b(u_N, \lambda_N) - g(\lambda_N; \boldsymbol{\mu}) = \overline{u}_N^T \overline{B}_N \overline{\lambda}_N - \overline{g}_N^T \overline{\lambda}_N = (\overline{u}_N^T \overline{B}_N - \overline{g}_N^T)\overline{\lambda}_N$$

which is the last equation (2.15) of $DSP_N(\boldsymbol{\mu})$.

$DSP_N(\boldsymbol{\mu}) \Rightarrow SP_N(\boldsymbol{\mu})$ : Let $(\overline{u}_N, \overline{\lambda}_N)$ be a solution of $DSP_N(\boldsymbol{\mu})$ and $(u_N, \lambda_N) \in V \times W$ the corresponding functions obtained by the corresponding linear combinations. Then obviously $\lambda_N \in M_N$ as it is a positive combination by (2.13). Eqn. (2.12) is equivalent to

$$a(u_N, \varphi_i; \boldsymbol{\mu}) + b(v_i, \lambda_N) = f(\varphi_i; \boldsymbol{\mu}) \quad \forall i.$$

Due to the linearity this also holds for arbitrary $v_N = \sum_{i=1}^{N_V} \alpha_i \varphi_i$, i.e.,

$$\begin{aligned} a(u_N, v_N; \boldsymbol{\mu}) + b(v_N, \lambda_N) &= \sum_i \alpha_i (a(u_N, \varphi_i; \boldsymbol{\mu}) + b(\varphi_i, \lambda_N)) \\ &= \sum_i \alpha_i f(\varphi_i; \boldsymbol{\mu}) = f(v_N; \boldsymbol{\mu}), \end{aligned}$$

which is the equality of $SP_N(\boldsymbol{\mu})$. As before, we rewrite (2.14) as

$$b(u_N, \lambda(\boldsymbol{\mu}_i)) - g(\lambda(\boldsymbol{\mu}_i); \boldsymbol{\mu}) \leq 0 \quad \forall i. \tag{A.8}$$

Hence we obtain

$$b(u_N, \eta_N) - g(\eta_N; \boldsymbol{\mu}) \leq 0 = b(u_N, \lambda_N) - g(\lambda_N; \boldsymbol{\mu}), \quad \forall \eta_N \in M_N,$$

where the first inequality follows from positive linear combinations of (A.8) and the second equality is due to (2.15). Resorting terms yields

$$b(u_N, \eta_N - \lambda_N) \leq g(\eta_N - \lambda_N; \boldsymbol{\mu}), \quad \forall \eta_N \in M_N,$$

which is the inequality of $SP_N(\boldsymbol{\mu})$. □

**A.6. Properties of the projection for the case $W = V'$.** In Section 5.2 we introduced an alternative projection $\pi$. For use in a-posteriori error estimation, different properties must be checked. We first verify that $\pi$ is the orthogonal projection with respect to the inner product $\langle \cdot, \cdot \rangle_\pi$ by showing the characterizing property

$$\langle \eta - \pi(\eta), \eta' - \pi(\eta) \rangle_\pi \leq 0, \quad \eta \in W, \eta' \in M.$$

This can be seen by realizing that

$$\begin{aligned} \langle \eta - \pi(\eta), \eta' - \pi(\eta) \rangle_\pi &= (\underline{\eta} - \underline{\pi})^T (\underline{M}^W)^2 (\underline{\eta}' - \underline{\pi}) \\ &= (\underline{M}^W \underline{\eta} - [\underline{M}^W \underline{\eta}]_+)^T (\underline{M}^W \underline{\eta}' - [\underline{M}^W \underline{\eta}]_+) \\ &= \sum_{i=1}^{H_W} (\underline{M}^W \underline{\eta} - [\underline{M}^W \underline{\eta}]_+)_i (\underline{M}^W \underline{\eta}' - [\underline{M}^W \underline{\eta}]_+)_i. \end{aligned}$$

We always have $(\underline{M}^W \underline{\eta} - [\underline{M}^W \underline{\eta}]_+)_i \leq 0$ by definition. If $(\underline{M}^W \underline{\eta} - [\underline{M}^W \underline{\eta}]_+)_i < 0$, then $(\underline{M}^W \underline{\eta})_i < 0$ and hence $\left([\underline{M}^W \underline{\eta}]_+\right)_i = 0$ implying $(\underline{M}^W \underline{\eta}' - [\underline{M}^W \underline{\eta}]_+)_i \geq 0$ as $\eta' \in M$

and $\underline{M}^V$ being an $M$-matrix. Consequently, $\pi$ is an orthogonal projection. Also, the norm induced by $\langle \cdot, \cdot \rangle_\pi$ is trivially equivalent to the $W$-norm due to the finite dimensionality. Next, we argue on the validity of the assumptions (4.6)-(4.8). For (4.6) we obtain with $\eta \in W, \eta' \in M$:

$$\begin{aligned}\langle \eta - \pi(\eta), \eta' \rangle_W &= (\underline{\eta} - \underline{\pi})^T \underline{M}^W \underline{\eta}' = (\underline{M}^W)^{-1} (\underline{M}^W \underline{\eta} - [\underline{M}^W \underline{\eta}]_+)^T \underline{M}^W \underline{\eta}' \\ &= (\underline{M}^W \underline{\eta} - [\underline{M}^W \underline{\eta}]_+)^T \underline{\eta}' \leq 0,\end{aligned}$$

as $\eta' \in M$ implies $\underline{\eta}' \geq 0$, while $\underline{M}^W \underline{\eta} - [\underline{M}^W \underline{\eta}]_+ \leq 0$. Recalling the definition of the detailed inequality residual (4.5), we obtain for all $\eta' \in M$

$$0 \geq b(u(\boldsymbol{\mu}), \eta') - g(\eta') = \langle \eta', \tilde{\eta}_s \rangle_W = (\underline{\eta}')^T \underline{M}^W \underline{\tilde{\eta}}_s.$$

This is equivalent to

$$\underline{M}^W \underline{\tilde{\eta}}_s \leq 0 \Leftrightarrow \pi(\tilde{\eta}_s) = 0,$$

which proves (4.7). For $\eta \in M$ we finally obtain

$$\langle \eta, \tilde{\eta}_s \rangle_\pi = \underline{\eta} (\underline{M}^W)^2 \underline{\tilde{\eta}}_s = \underline{\tilde{\eta}} \underline{M}^W \underline{\tilde{\eta}}_s = \langle \tilde{\eta}, \tilde{\eta}_s \rangle_W$$

by setting $\underline{\tilde{\eta}} := \underline{M}^W \underline{\eta}$ and $\tilde{\eta} := \sum_i \underline{\tilde{\eta}}_i \chi_i$. Hence, for obtaining the non-positivity (4.8) it remains to show that $\tilde{\eta} \in M$. Indeed, $\underline{\eta} \geq 0$ and by non-negativity of $\underline{M}^W$ we get $\underline{M}^W \underline{\eta} \geq 0$, which means $\tilde{\eta} \in M$. So the projection $\pi$ indeed satisfies all properties required for its use in the a-posteriori error estimation.