

MATHICSE Technical Report

Nr. 08.2014

February 2014 (NEW march 2015)



Heuristic strategies for the approximation of stability factors in quadratically nonlinear parametrized PDEs

Andrea Manzoni, Federico Negri

Heuristic strategies for the approximation of stability factors in quadratically nonlinear parametrized PDEs

Andrea Manzoni · Federico Negri

Received: date / Accepted: date

Abstract In this paper we present some heuristic strategies to compute rapid and reliable approximations to stability factors in nonlinear, inf-sup stable parametrized PDEs. The efficient evaluation of these quantities is crucial for the rapid construction of a posteriori error estimates to reduced basis approximations. In this context, stability factors depend on the problem's solution, and in particular on its reduced basis approximation. Their evaluation becomes therefore very expensive and cannot be performed prior to (and independently of) the construction of the reduced space. As a remedy, we first propose a linearized, heuristic version of the Successive Constraint Method (SCM), providing a suitable estimate – rather than a rigorous lower bound as in the original SCM – of the stability factor. Moreover, for the sake of computational efficiency, we develop an alternative heuristic strategy, which combines a radial basis interpolant, suitable criteria to ensure its positive-ness, and an adaptive choice of interpolation points through a greedy procedure. We provide some theoretical results to support the proposed strategies, which are then applied to a set of test cases dealing with parametrized Navier-Stokes equations. Finally, we show that the interpolation strategy is inexpensive to apply and robust even in the proximity of bifurcation points, where the estimate of stability factors is particularly critical.

Keywords Stability factors · nonlinear parametrized PDEs · reduced basis methods · Brezzi-Rappaz-Raviart theory · radial basis functions

Mathematics Subject Classification (2000) 65N12 · 65N30 · 76D05 · 78M34

A. Manzoni
CMCS-MATHICSE-SB, Ecole Polytechnique Fédérale de Lausanne,
Station 8, CH-1015 Lausanne, Switzerland
E-mail: andrea.manzoni@epfl.ch

F. Negri
CMCS-MATHICSE-SB, Ecole Polytechnique Fédérale de Lausanne,
Station 8, CH-1015 Lausanne, Switzerland
E-mail: federico.negri@epfl.ch

1 Introduction

Stability factors of differential operators are relevant for the well-posedness analysis of problems governed by PDEs and enter in the (a posteriori) error estimates of any numerical approximation method. Their rapid and reliable evaluation is thus crucial, especially when dealing with nonlinear parametrized PDEs, i.e. nonlinear PDEs depending on a vector of $p \geq 1$ input parameters $\boldsymbol{\mu} \in \mathcal{D} \subset \mathbb{R}^p$, related to either physical properties or geometrical features.

In this paper we focus on (quadratically) nonlinear parametrized PDEs, which can be written in the following general form:

$$\mathcal{E}(\boldsymbol{\mu}, u) = 0 \quad \text{in } V', \quad (1)$$

being V a suitable Hilbert space, V' its dual and $\mathcal{E} : \mathcal{D} \times V \rightarrow V'$ a nonlinear, inf-sup stable, parametrized operator. A meaningful example is represented by the steady Navier-Stokes equations in a backward facing step domain, parametrized with respect to the Reynolds number and/or the domain aspect ratio.

Our long-term goal is to compute, in a very efficient way, a numerical approximation of the solution $u(\boldsymbol{\mu})$ for any $\boldsymbol{\mu} \in \mathcal{D}$. To this end, we rely on the Reduced Basis (RB) method, which allows to compute a reduced approximation $u_N(\boldsymbol{\mu}) \in V_N$ of the PDE solution $u(\boldsymbol{\mu}) \in V$, for any $\boldsymbol{\mu} \in \mathcal{D}$, as a linear combination of *snapshots* corresponding to a small set of sampled parameter values $\boldsymbol{\mu}^1, \dots, \boldsymbol{\mu}^N$. This can be made through a Galerkin projection in the low-dimensional subspace $V_N = \text{span}\{u_h(\boldsymbol{\mu}^1), \dots, u_h(\boldsymbol{\mu}^N)\}$, being $u_h(\boldsymbol{\mu}^i) \in V_h$, $i = 1, \dots, N$. Here V_h is a high-fidelity approximation space of dimension $\mathcal{N}_h \gg N$ and $u_h(\boldsymbol{\mu}) \in V_h$ is the high-fidelity approximation to $u(\boldsymbol{\mu})$, obtained by any kind of numerical discretization technique. Moreover, we aim at providing an *a posteriori* error bound, which usually takes the form [5, 32]

$$\|u_h(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_V \leq \frac{2}{\beta_h(u_N(\boldsymbol{\mu}))} \|\mathcal{E}(\boldsymbol{\mu}, u_N(\boldsymbol{\mu}))\|_{V'}, \quad (2)$$

at least for N sufficiently large, see Sect. 3.2. Here $\beta_h(u_N(\boldsymbol{\mu}))$ denotes the stability factor related with the (discrete, high-fidelity approximation of the) differential operator. The computation of $\beta_h(u_N(\boldsymbol{\mu}))$, for any $\boldsymbol{\mu} \in \mathcal{D}$, requires the solution of a generalized eigenproblem of dimension \mathcal{N}_h , thus preventing both the Offline and Online efficiency of the RB approximation.

To overcome this computational bottleneck, the so-called Successive Constraint Method (SCM) has been first introduced in [18] (see also [7, 8]). A general version using the so-called *natural norm* [30] has been analyzed in [16], while a recent application to Stokes equations is given in [28]. This method has been developed for linear parametrized operators and provides a parametric lower bound to their stability factor. Since in the linear case the latter is independent of $u_N(\boldsymbol{\mu})$, the procedure admits an Offline-Online computational treatment for which the Online cost is independent of \mathcal{N}_h , and the Offline computations are performed prior to the RB space construction.

In the nonlinear case, since the stability factor depends on the RB solution $u_N(\boldsymbol{\mu})$, the construction of suitable lower bounds can not be performed prior to (and independently of) the construction of the reduced space. To overcome this bottleneck, we propose to approximate $\beta_h(u_N(\boldsymbol{\mu}))$ by $\beta_h(u_h(\boldsymbol{\mu}))$, i.e. by the

stability factor evaluated with respect to the high-fidelity solution $u_h(\boldsymbol{\mu})$. Indeed, thanks to the approximation property of the RB space V_N , we can prove that the error $|\beta_h(u_N(\boldsymbol{\mu})) - \beta_h(u_h(\boldsymbol{\mu}))|$ vanishes as $N \rightarrow \mathcal{N}_h$. Then, we propose two different strategies to construct, prior to the generation of the RB space V_N , an estimate to the stability factor $\beta_h(u_h(\boldsymbol{\mu}))$.

In particular, we first develop a *linearized* version of SCM. The proposed algorithm is mechanically similar to the original SCM, but is different in spirit in that it provides an estimate, rather than a lower bound, of the stability factor. Indeed, we forgo the rigor of the original SCM to enhance computational efficiency. Nevertheless, although this procedure enables a very rapid Online evaluation of the stability factor, it still entails a quite expensive Offline stage (especially when dealing with $p \geq 3$ parameters), which may jeopardize the efficiency of the whole reduction process, as shown by the numerical test cases of Sect. 6.

For this reason, we then propose some inexpensive, heuristic strategies to directly approximate the stability factor. These strategies combine a radial basis interpolant to the stability factor, suitable criteria to ensure its positiveness, and an adaptive choice of interpolation points through a greedy procedure. In this way, it is possible to obtain a reliable approximation of the stability factor, whose Offline construction and Online evaluation prove to be much faster than in the case of the linearized SCM algorithm.

We test the efficacy of these procedures by considering different flow problems which depend on both physical and geometrical parameters. Moreover, in order to assess the robustness of the adaptive interpolation, we also consider a numerical test case whose solution features a bifurcation point – where the estimate of stability factors is critical. Hence, we also show that our heuristic technique proves to be effective when aiming at the detection of bifurcation points.

The structure of the paper is as follows. In Sect. 2 we provide some theoretical results dealing with the Lipschitz-continuity of affine, parametrized nonlinear operators, in order to support the proposed algorithms. Despite some ideas were anticipated in [22] and, more recently, in [33,34], in Appendix A we provide a detailed proof. In Sect. 3 both the high-fidelity and the reduced approximation of this class of problems are introduced. In Sect. 4 we present our linearized extension of the SCM algorithm to the case of nonlinear inf-sup stable operators. Then, in Sect. 5 we present some heuristic strategies to circumvent the computational burden entailed by the SCM algorithm. In the end, we show in Sect. 6 some numerical test cases dealing with steady parametrized Navier-Stokes equations. Finally, in Sect. 7 we test the robustness of the interpolation procedure in the proximity of a bifurcation point.

2 Stability factors for nonlinear, inf-sup stable parametrized PDEs

Given a regular spatial domain $\Omega \subset \mathbb{R}^n$, let $V = V(\Omega)$ be a Hilbert space with inner product $(\cdot, \cdot)_V$ and induced norm $\|v\|_V = \sqrt{(v, v)_V}$. Moreover, let $\mathcal{D} \subset \mathbb{R}^p$ be a bounded closed parameter domain, whose elements are denoted by $\boldsymbol{\mu} \in \mathcal{D}$.

Although the techniques proposed in this paper are suitable also for more general nonlinear problems, here we deepen the analysis in the case of stationary, quadratically nonlinear parametrized operators, for which our problem of interest

can be expressed as follows: given $\boldsymbol{\mu} \in \mathcal{D}$, find $u = u(\boldsymbol{\mu}) \in V$ s.t.

$$A(u(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = a(u(\boldsymbol{\mu}), v; \boldsymbol{\mu}) + c(u(\boldsymbol{\mu}), u(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = f(v; \boldsymbol{\mu}) \quad \forall v \in V. \quad (3)$$

Here $a(\cdot, \cdot; \boldsymbol{\mu})$ is a continuous, inf-sup stable bilinear form over $V \times V$ and $c(\cdot, \cdot, \cdot; \boldsymbol{\mu})$ is a continuous trilinear form over $V \times V \times V$. Moreover, the right-hand side is a parametrized linear form $f(\cdot; \boldsymbol{\mu}) : V \rightarrow \mathbb{R}$, given by

$$f(v; \boldsymbol{\mu}) = {}_{V'} \langle F(\boldsymbol{\mu}), v \rangle_V,$$

being $F(\boldsymbol{\mu}) \in V'$ and $V' = \mathcal{L}(V; \mathbb{R})$ the dual space of V . According to the general Brezzi-Rappaz-Raviart (BRR) theory [3], problem (3) is well posed if and only if the following continuity and (Babuška) inf-sup conditions hold:

$$\gamma(u_h(\boldsymbol{\mu})) := \sup_{v \in V} \sup_{w \in V} \frac{dA(u(\boldsymbol{\mu}); \boldsymbol{\mu})(v, w)}{\|v\|_V \|w\|_V} < +\infty, \quad \forall \boldsymbol{\mu} \in \mathcal{D}, \quad (4)$$

$$\exists \beta^0(\boldsymbol{\mu}) > 0 : \beta(u_h(\boldsymbol{\mu})) := \inf_{v \in V} \sup_{w \in V} \frac{dA(u(\boldsymbol{\mu}); \boldsymbol{\mu})(v, w)}{\|v\|_V \|w\|_V} \geq \beta^0(\boldsymbol{\mu}), \quad \forall \boldsymbol{\mu} \in \mathcal{D}. \quad (5)$$

In fact, these conditions ensure the existence of a *local branch of non-singular solutions* [13], see Proposition 1. Here $dA(u(\boldsymbol{\mu}); \boldsymbol{\mu})(\cdot, \cdot)$ denotes the Fréchet derivative of $A(\cdot, \cdot; \boldsymbol{\mu})$ with respect to the first variable, which is given, at $z \in V$, by

$$dA(z; \boldsymbol{\mu})(w, v) = a(w, v; \boldsymbol{\mu}) + c(z, w, v; \boldsymbol{\mu}) + c(w, z, v; \boldsymbol{\mu}) \quad \forall v, w, z \in V; \quad (6)$$

from now on, we denote $d(z; \boldsymbol{\mu})(w, v) = c(z, w, v; \boldsymbol{\mu}) + c(w, z, v; \boldsymbol{\mu})$. Furthermore,

$$\gamma^a(\boldsymbol{\mu}) = \sup_{v \in V} \sup_{w \in V} \frac{a(v, w; \boldsymbol{\mu})}{\|v\|_V \|w\|_V} < +\infty, \quad \gamma^c(\boldsymbol{\mu}) = \sup_{u \in V} \sup_{v \in V} \sup_{w \in V} \frac{c(u, v, w; \boldsymbol{\mu})}{\|u\|_V \|v\|_V \|w\|_V} < +\infty$$

denote the continuity constants of $a(\cdot, \cdot; \boldsymbol{\mu})$ and $c(\cdot, \cdot, \cdot; \boldsymbol{\mu})$, respectively, while

$$\gamma^d(\boldsymbol{\mu}) = \sup_{v \in V} \sup_{w \in V} \frac{d(u(\boldsymbol{\mu}); \boldsymbol{\mu})(v, w)}{\|v\|_V \|w\|_V} < +\infty$$

denotes the continuity constant of $d(u; \boldsymbol{\mu})(\cdot, \cdot)$. The stability factor $\beta(\boldsymbol{\mu})$ we want to estimate obviously depends on the nonlinear form on the left-hand side of (3), and thus, through $u(\boldsymbol{\mu})$, on the right-hand side, too. This makes the accurate estimate of $\beta(\boldsymbol{\mu})$ much more involved than in the linear case.

In the following subsections we introduce some definitions and provide some basic results on the continuity and the regularity of the map $\boldsymbol{\mu} \mapsto u(\boldsymbol{\mu})$ that will be used in the sequel.

2.1 Supremizer operator, norms and parametric dependence

We introduce the parametrized linear operator $T^\mu : V \rightarrow V$ such that, for any $\mu \in \mathcal{D}$, $v \in V$,

$$(T^\mu v, w)_V = dA(u(\mu); \mu)(v, w) \quad \forall w \in V; \quad (7)$$

equivalently, by Riesz theorem,

$$T^\mu v = \arg \sup_{w \in V} \frac{dA(u(\mu); \mu)(v, w)}{\|w\|_V}, \quad \forall v \in V. \quad (8)$$

Because of (8), T^μ is called *supremizer operator*. It follows that (4) and (5) can be equivalently expressed as

$$\gamma(\mu) = \sup_{w \in V} \frac{\|T^\mu w\|_V}{\|w\|_V}, \quad \beta(\mu) = \inf_{w \in V} \frac{\|T^\mu w\|_V}{\|w\|_V}. \quad (9)$$

Assuming that $0 < \beta^0(\mu) \leq \beta(\mu)$ and $\gamma(\mu) < \infty$ for each $\mu \in \mathcal{D}$, implies that

$$\|w\|_\mu := \|T^\mu w\|_V \quad \forall w \in V, \quad (10)$$

defines a norm, usually referred to [30,10] as *natural norm*. Thanks to (9), this latter is equivalent to the V-norm,

$$\frac{1}{\gamma(\mu)} \|T^\mu w\|_V \leq \|w\|_V \leq \frac{1}{\beta(\mu)} \|T^\mu w\|_V, \quad \forall w \in V. \quad (11)$$

In order to develop an Offline/Online strategy, we assume that the forms appearing in (3) fulfill the following *parameter separability* – also called *affine parameter dependence* – property: for any $u, v, w \in V$,

$$a(u, v; \mu) = \sum_{q=1}^{Q_a} \Theta_q^a(\mu) a_q(u, v), \quad c(u, v, w; \mu) = \sum_{q=1}^{Q_c} \Theta_q^c(\mu) c_q(u, v, w) \quad (12)$$

for some integers Q_a, Q_c , where $\Theta_q^a, \Theta_q^c \in C^1(\mathcal{D})$ and $a_q(\cdot, \cdot), c_q(\cdot, \cdot, \cdot)$ are continuous bilinear (trilinear) forms over $V \times V$ ($V \times V \times V$), respectively; moreover, we set $Q_A = Q_a + Q_c$. The requirement that Θ_q^a, Θ_q^c are of class $C^1(\mathcal{D})$ is essential to ensure that $\mu \mapsto u(\mu)$ is a regular map, as we will see in Sect. 2.2.

Moreover, we denote the (now μ -independent) continuity constants of $a_q(\cdot, \cdot)$ and $c_q(\cdot, \cdot, \cdot)$ by

$$\gamma_q^a = \sup_{v \in V} \sup_{w \in V} \frac{a_q(v, w)}{\|v\|_V \|w\|_V} < +\infty, \quad \gamma_q^c = \sup_{u \in V} \sup_{v \in V} \sup_{w \in V} \frac{c_q(u, v, w)}{\|u\|_V \|v\|_V \|w\|_V} < +\infty, \quad (13)$$

respectively. In the same way, we set

$$d(z; \mu)(w, v) = \sum_{q=1}^{Q_c} \Theta_q^c(\mu) (c_q(z, w, v) + c_q(w, z, v)) = \sum_{q=1}^{Q_c} \Theta_q^c(\mu) d_q(z)(v, w),$$

and denote by

$$\gamma_q^d(u) = \sup_{v \in V} \sup_{w \in V} \frac{d_q(u)(v, w)}{\|v\|_V \|w\|_V} < +\infty \quad (14)$$

the (u -dependent) continuity constants of $d_q(u)(\cdot, \cdot)$.

2.2 Fréchet derivatives of operators and regularity of solutions

Let us show some theoretical results required to ensure the well-posedness of the linearized SCM procedure. For the sake of generality, let us cast problem (3) under the form (1), where the operator $\mathcal{E} : \mathcal{D} \times V \rightarrow V'$ is defined as:

$${}_{V'}\langle \mathcal{E}(\boldsymbol{\mu}, z), w \rangle_V = A(z, w; \boldsymbol{\mu}) - f(w; \boldsymbol{\mu}). \quad (15)$$

Let us denote by $d_u \mathcal{E}(\boldsymbol{\mu}, z) : V \rightarrow V'$ and $d_{\boldsymbol{\mu}} \mathcal{E}(\boldsymbol{\mu}, z) : \mathcal{D} \rightarrow V'$ the (partial) Fréchet derivatives of \mathcal{E} at $(\boldsymbol{\mu}, z) \in \mathcal{D} \times V$. Moreover, we denote by $B_r(\boldsymbol{\mu}) \subset \mathcal{D}$ the open ball with radius $r > 0$ and center $\boldsymbol{\mu} \in \mathcal{D}$. First of all, we can state a general result ensuring that $\boldsymbol{\mu} \mapsto u(\boldsymbol{\mu})$ is a regular map:

Proposition 1 *For the parametrized operator $A(\cdot, \cdot; \boldsymbol{\mu}) : V \times V \rightarrow \mathbb{R}$ defined in (3), suppose that:*

1. *the continuity and the inf-sup conditions (4)–(5) hold;*
2. *the parameter separability assumption (12) holds, being $\Theta_q^a, \Theta_{q'}^c : \mathcal{D} \rightarrow \mathbb{R}$, $q = 1, \dots, Q_a$, $q' = 1, \dots, Q_c$, prescribed C^1 functions.*

Moreover assume that $\mathcal{E}(\boldsymbol{\mu}_0, u_0) = 0$ for some $\boldsymbol{\mu}_0 \in \mathcal{D}$, $u_0 \in V$. Then, there exist $r_0, r > 0$ and a unique $u(\boldsymbol{\mu}) \in B_r(u_0) \cap V$ such that

$$\mathcal{E}(\boldsymbol{\mu}, u(\boldsymbol{\mu})) = 0 \quad \forall \boldsymbol{\mu} \in B_{r_0}(\boldsymbol{\mu}_0) \cap \mathcal{D}.$$

Furthermore, the map $\boldsymbol{\mu} \mapsto u(\boldsymbol{\mu})$ is Lipschitz continuous and

$$u'(\boldsymbol{\mu}) = -(d_u \mathcal{E}(\boldsymbol{\mu}, u(\boldsymbol{\mu})))^{-1} d_{\boldsymbol{\mu}} \mathcal{E}(\boldsymbol{\mu}, u(\boldsymbol{\mu})).$$

Proof The proof is a direct consequence of the Implicit Function Theorem: we refer here to the version stated by Hildebrandt and Graves [15], see also [35]. A very general version providing further insights on the Lipschitz constant of the map $\boldsymbol{\mu} \mapsto u(\boldsymbol{\mu})$ can be found, e.g., in [5, 19]. Provided that the continuity condition (4) holds, $d_u \mathcal{E}$ is continuous at each point $(\boldsymbol{\mu}_0, u_0) \in \mathcal{D} \times V$; its inverse is a continuous linear operator thanks to the inf-sup condition (5) – in other words, $d_u \mathcal{E}$ is an isomorphism, for any $(\boldsymbol{\mu}_0, u_0) \in \mathcal{D} \times V$. Furthermore, if the parameter separability assumption (12) holds, for suitable C^1 functions $\Theta_q^a, \Theta_{q'}^c : \mathcal{D} \rightarrow \mathbb{R}$, $q = 1, \dots, Q_a$, $q' = 1, \dots, Q_c$, then \mathcal{E} is a C^1 map. Then, the Implicit Function Theorem ensures the existence of $r_0, r > 0$ and of a (unique) C^1 map $\boldsymbol{\mu} \mapsto u(\boldsymbol{\mu})$ such that, for every $\boldsymbol{\mu} \in B_{r_0}(\boldsymbol{\mu}_0) \cap \mathcal{D}$, $\mathcal{E}(\boldsymbol{\mu}, u(\boldsymbol{\mu})) = 0$. \square

Exploiting the result above, we can show that also the Fréchet derivative of $A(\cdot, \cdot; \boldsymbol{\mu})$ defines a regular map, provided that suitable *a priori* (or energy) estimates on $\|u(\boldsymbol{\mu})\|_V$ hold. In the same way, also the supremizer operator defines a Lipschitz-continuous map with respect to parameter variations (see Appendix A.1 for the proof).

Proposition 2 *Under the assumptions of Proposition 1, and by assuming that*

$$\exists K_u > 0 \quad \text{s.t.} \quad \|u(\boldsymbol{\mu})\|_V \leq K_u \quad \forall \boldsymbol{\mu} \in \mathcal{D}, \quad (16)$$

there exists a positive constant $C > 0$ such that, for any $\boldsymbol{\mu}, \boldsymbol{\mu}^ \in \mathcal{D}$, $v, w \in V$*

$$|dA(u(\boldsymbol{\mu}); \boldsymbol{\mu})(v, w) - dA(u(\boldsymbol{\mu}^*); \boldsymbol{\mu}^*)(v, w)| \leq C \|\boldsymbol{\mu} - \boldsymbol{\mu}^*\| \|v\|_V \|w\|_V. \quad (17)$$

Furthermore, the following estimate holds:

$$\|T^\mu w - T^{\mu^*} w\|_V \leq \frac{C}{\beta(\mu^*)} |\mu - \mu^*| \|T^{\mu^*} w\|_V \quad \forall w \in V. \quad (18)$$

Remark 1 In the Navier-Stokes case, an *a priori* estimate like (16) follows by the coercivity of the bilinear form $a(\cdot, \cdot; \mu)$ and the skew-symmetry (with respect to the last two arguments) of the trilinear form $c(\cdot, \cdot, \cdot; \mu)$; see e.g. [31], Sect. 2.1.

3 High-fidelity and reduced approximation

In this section we introduce the high-fidelity approximation of problem (3), based on a Galerkin-Finite Element (FE) method, and then a lower-fidelity approximation based on the RB method. Moreover, we discuss some stability issues related with these two approximation strategies, as well as a general *a posteriori* error estimate, where the role of stability factors is highlighted.

3.1 Finite element approximation

Let us denote by $V_h \subset V$ a FE approximation space of dimension \mathcal{N}_h , with inherited inner product $(v, w)_{V_h} = (v, w)_V$ and norm $\|v\|_{V_h} = \|v\|_V$. The Galerkin-FE approximation of (3) reads as follows: given $\mu \in \mathcal{D}$, find $u_h(\mu) \in V_h$ s.t.

$$A(u_h(\mu), v_h; \mu) = f(v_h; \mu) \quad \forall v_h \in V_h. \quad (19)$$

Problem (19) is equivalent to the following algebraic nonlinear system: given $\mu \in \mathcal{D}$, find $\mathbf{u}_h(\mu) \in \mathbb{R}^{\mathcal{N}_h}$ such that

$$(\mathbb{K}(\mu) + \mathbb{C}(\mathbf{u}_h(\mu); \mu))\mathbf{u}_h(\mu) = \mathbf{f}(\mu) \quad \text{in } \mathbb{R}^{\mathcal{N}_h}. \quad (20)$$

Here $\mathbf{u}_h(\mu) \in \mathbb{R}^{\mathcal{N}_h}$ is the vector representation of $u_h(\mu) \in V_h$ over a Lagrangian basis $\{\varphi_j^h\}_{j=1}^{\mathcal{N}_h}$ of V_h , i.e. given a set of Lagrange nodes $\{\mathbf{x}_j\}_{j=1}^{\mathcal{N}_h}$, $\mathbf{v}_h^{(j)} = v_h(\mathbf{x}_j)$ for any $v_h \in V_h$. Moreover,

$$(\mathbb{K}(\mu))_{ij} = a(\varphi_j^h, \varphi_i^h; \mu), \quad (\mathbb{C}(\mathbf{w}_h; \mu))_{ij} = c(w_h, \varphi_i^h, \varphi_j^h; \mu), \quad \mathbf{f}^{(j)}(\mu) = f(\varphi_j^h; \mu)$$

are the matrices corresponding to the linear and the nonlinear term, and the vector corresponding to the source term, respectively ($i, j = 1, \dots, \mathcal{N}_h$).

Concerning the stability of the approximation (19), we rely on the Brezzi-Rappaz-Raviart theory [3, 5]. As for the original problem, let us assume that¹:

$$\gamma_h(u_h(\mu)) := \sup_{v_h \in V_h} \sup_{w_h \in V_h} \frac{dA(u_h(\mu); \mu)(v_h, w_h)}{\|v_h\|_V \|w_h\|_V} < +\infty, \quad \forall \mu \in \mathcal{D}, \quad (21)$$

$$\exists \beta_h^0(\mu) > 0 : \beta_h(u_h(\mu)) := \inf_{v_h \in V_h} \sup_{w_h \in V_h} \frac{dA(u_h(\mu); \mu)(v_h, w_h)}{\|v_h\|_V \|w_h\|_V} \geq \beta_h^0(\mu), \quad \forall \mu \in \mathcal{D}. \quad (22)$$

¹ In the following we denote $\beta_h(u_h(\mu))$ by $\beta_h(\mu)$, i.e. we omit the dependence on the solution, wherever it is clear from the context.

Then, if V_h is chosen accordingly to these conditions – which are, in fact, the discrete version of (4)–(5) – problem (19) admits a unique solution.

Concerning the regularity of the solution with respect to $\boldsymbol{\mu}$, a result similar to that of Proposition 1 can be proved if we consider a Galerkin approximation, i.e. find $u_h \in V_h$ s.t.

$${}_V \langle \mathcal{E}(\boldsymbol{\mu}, u_h), w_h \rangle_V = 0 \quad \forall w_h \in V_h, \quad (23)$$

where V_h is such that (21) and (22) hold (see e.g. [5, Chapter 12 and Remark 13.2]). For instance, Taylor-Hood elements [13] allow to meet these requirements in the Navier-Stokes case.

By introducing the discrete supremizer operator² $T^\mu : V_h \rightarrow V_h$ s.t.

$$(T^\mu v_h, w_h)_V = dA(u_h(\boldsymbol{\mu}); \boldsymbol{\mu})(v_h, w_h) \quad \forall v_h, w_h \in V_h, \quad (24)$$

we have that

$$(\beta_h(u_h(\boldsymbol{\mu})))^2 = \left(\inf_{v \in V_h} \frac{dA(u_h(\boldsymbol{\mu}); \boldsymbol{\mu})(v, T^\mu v)}{\|v\|_V \|T^\mu v\|_V} \right)^2 = \inf_{v \in V_h} \frac{\|T^\mu v\|_V^2}{\|v\|_V^2}. \quad (25)$$

The algebraic counterpart of (25) can be obtained by introducing the matrix norm

$$\mathbb{X}_{ij} = (\varphi_j^h, \varphi_i^h)_V, \quad i, j = 1, \dots, \mathcal{N}_h \quad (26)$$

of V_h , so that $\|v_h\|_V^2 = \mathbf{v}_h^T \mathbb{X} \mathbf{v}_h$ for any $v_h \in V_h$. Moreover, denoting by \mathbf{t}_h the vector of components $\mathbf{t}_h^{(i)} = (T^\mu v_h)(x_i)$, we have that

$$\mathbf{w}_h^T \mathbb{X} \mathbf{t}_h = \mathbf{w}_h^T \mathbb{F}(\boldsymbol{\mu}) \mathbf{v}_h, \quad (27)$$

being $\mathbb{F}(\boldsymbol{\mu}) \in \mathbb{R}^{\mathcal{N}_h \times \mathcal{N}_h}$ the matrix corresponding to the FE discretization of the Fréchet derivative, i.e. $(\mathbb{F}(\boldsymbol{\mu}))_{ij} = dA(u_h(\boldsymbol{\mu}); \boldsymbol{\mu})(\varphi_i^h, \varphi_j^h)$, $i, j = 1, \dots, \mathcal{N}_h$. Then,

$$\beta_h(u_h(\boldsymbol{\mu})) = \inf_{\mathbf{v}_h \in \mathbb{R}^{\mathcal{N}_h}} \frac{\mathbf{v}_h^T \mathbb{F}^T(\boldsymbol{\mu}) \mathbb{X}^{-1} \mathbb{F}(\boldsymbol{\mu}) \mathbf{v}_h}{\mathbf{v}_h^T \mathbb{X} \mathbf{v}_h},$$

from (25) and (27), so that $\beta_h(u_h(\boldsymbol{\mu})) = (\lambda_{\min}(\boldsymbol{\mu}))^{1/2}$, where $\lambda_{\min}(\boldsymbol{\mu})$ is the smallest eigenvalue $\lambda(\boldsymbol{\mu})$ such that $(\lambda(\boldsymbol{\mu}), \mathbf{v}_h) \in \mathbb{R}_+ \times V_h$, $\mathbf{v}_h \neq 0$, satisfy

$$\mathbb{F}^T(\boldsymbol{\mu}) \mathbb{X}^{-1} \mathbb{F}(\boldsymbol{\mu}) \mathbf{v}_h = \lambda(\boldsymbol{\mu}) \mathbb{X} \mathbf{v}_h. \quad (28)$$

Thus, the evaluation of the stability factor $\beta_h(u_h(\boldsymbol{\mu}))$, for any $\boldsymbol{\mu} \in \mathcal{D}$, entails the solution of both (20) and (28).

² For the sake of notation, we denote by T^μ the discrete supremize operator, too.

3.2 Reduced basis approximation

Our final goal is to compute, for any $\boldsymbol{\mu} \in \mathcal{D}$, a RB approximation $u_N(\boldsymbol{\mu}) \in V_N$ to $u_h(\boldsymbol{\mu})$, where

$$V_N = \text{span}\{u_h(\boldsymbol{\mu}^1), \dots, u_h(\boldsymbol{\mu}^N)\} \subset V_h \quad (29)$$

is a reduced space, made by $N \ll \mathcal{N}_h$ solutions to problem (19) computed for properly chosen parameter values $\boldsymbol{\mu}^1, \dots, \boldsymbol{\mu}^N$. Then, we perform the Gram-Schmidt procedure on the snapshots to obtain an orthonormal basis $\{\phi_1, \dots, \phi_N\}$, so that we have $u_N(\boldsymbol{\mu}) = \sum_{j=1}^N u_N^{(j)}(\boldsymbol{\mu}) \phi_j$, where the components $\{u_N^{(j)}\}_{j=1}^N$ are computed through a Galerkin³ projection of (3) over V_N : find $u_N(\boldsymbol{\mu}) \in V_N$ s.t.

$$A(u_N(\boldsymbol{\mu}), v_N; \boldsymbol{\mu}) = f(v_N; \boldsymbol{\mu}) \quad \forall v_N \in V_N. \quad (30)$$

Proving the stability of (30) would demand to prove that an inf-sup condition holds. This however could be challenging (see, e.g., [23]).

Here, we rather look for an estimate of the discrete inf-sup stability factor

$$\beta_h(u_N(\boldsymbol{\mu})) = \inf_{v \in V_h} \sup_{w \in V_h} \frac{dA(u_N(\boldsymbol{\mu}); \boldsymbol{\mu})(v, w)}{\|v\|_V \|w\|_V} \quad \forall \boldsymbol{\mu} \in \mathcal{D}, \quad (31)$$

which enters in the following *a posteriori* error bound: for any $N \geq N^*(\boldsymbol{\mu})$,

$$\|u_h(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_V \leq \frac{2}{\beta_h(u_N(\boldsymbol{\mu}))} \|r(\cdot; \boldsymbol{\mu})\|_{V_h'} \quad \forall \boldsymbol{\mu} \in \mathcal{D}, \quad (32)$$

where $r(w; \boldsymbol{\mu}) := A(u_N(\boldsymbol{\mu}), w; \boldsymbol{\mu}) - f(w; \boldsymbol{\mu}) \forall w \in V_h$ is the residual, $N^*(\boldsymbol{\mu})$ the smallest N such that $\tau_N(\boldsymbol{\mu}) < 1$, for all $N \geq N^*(\boldsymbol{\mu})$, $\tau_N(\boldsymbol{\mu})$ is defined as

$$\tau_N(\boldsymbol{\mu}) = \frac{4\gamma_h^c(\boldsymbol{\mu}) \|r(\cdot; \boldsymbol{\mu})\|_{V_h'}}{\beta_h(u_N(\boldsymbol{\mu}))^2},$$

and $\gamma_h^c(\boldsymbol{\mu})$ is the discrete continuity constant of $c(\cdot, \cdot; \boldsymbol{\mu})$. See e.g. [32, 10, 23] for further details and proofs in the Navier-Stokes case, or [6, 27] for recent applications to nonlinear advection/diffusion problems.

Since $\beta_h(u_N(\boldsymbol{\mu}))$ depends on the RB solution, computing a parametric lower bound for the stability factor *before* assembling the reduced space is *infeasible*, unless these two procedures are run simultaneously, as shown e.g. in [33]. This latter option however suffers from two crucial limitations: it requires a problem-specific and intrusive implementation of the two procedures, and worsen the computational complexity of the construction and evaluation of the lower bound by making it dependent on N (see [33] for further details). We avoid this extra burden by looking for a convenient approximation of the stability factor $\beta_h(u_h(\boldsymbol{\mu}))$ defined in (22), rather than seeking a lower bound to the stability factor $\beta_h(u_N(\boldsymbol{\mu}))$. In fact, the former quantity provides an *asymptotically* good approximation to the latter, thanks to

Proposition 3 *The following relation holds:*

$$|\beta_h(u_h(\boldsymbol{\mu})) - \beta_h(u_N(\boldsymbol{\mu}))| \leq 2\gamma_h^c(\boldsymbol{\mu}) \|u_h(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_V \quad \forall \boldsymbol{\mu} \in \mathcal{D}. \quad (33)$$

³ For the sake of simplicity, here we restrict ourselves to the case of Galerkin projection, although sometimes a more general Petrov-Galerkin method is used. See e.g. [25, 29] for a detailed review of RB methods.

Proof By exploiting the trilinearity and the continuity of $c(\cdot, \cdot; \boldsymbol{\mu})$, we have

$$\begin{aligned}
\beta_h(u_1) &= \inf_{v \in V_h} \sup_{w \in V_h} \left(\frac{a(v, w; \boldsymbol{\mu}) + c(u_2, v, w; \boldsymbol{\mu}) + c(v, u_2, w; \boldsymbol{\mu})}{\|v\|_V \|w\|_V} \right. \\
&\quad \left. + \frac{c(u_1 - u_2, v, w; \boldsymbol{\mu}) + c(v, u_1 - u_2, w; \boldsymbol{\mu})}{\|v\|_V \|w\|_V} \right) \\
&= \inf_{v \in V_h} \sup_{w \in V_h} \frac{dA(u_2)(v, w; \boldsymbol{\mu}) + c(u_1 - u_2, v, w; \boldsymbol{\mu}) + c(v, u_1 - u_2, w; \boldsymbol{\mu})}{\|v\|_V \|w\|_V} \\
&\leq \inf_{v \in V_h} \sup_{w \in V_h} \frac{dA(u_2)(v, w; \boldsymbol{\mu})}{\|v\|_V \|w\|_V} + 2\gamma_h^c(\boldsymbol{\mu}) \|u_1 - u_2\|_V \quad (34) \\
&= \beta_h(u_2) + 2\gamma_h^c(\boldsymbol{\mu}) \|u_1 - u_2\|_V.
\end{aligned}$$

By considering in the previous inequality first $u_1 = u_N(\boldsymbol{\mu})$, $u_2 = u_h(\boldsymbol{\mu})$, and then $u_2 = u_N(\boldsymbol{\mu})$, $u_1 = u_h(\boldsymbol{\mu})$, (33) easily follows. \square

Although (32) cannot be used to estimate $\|u_h(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_V$, thanks to the approximation property of the space V_N , (33) can be regarded as an *a priori* convergence result. In fact, provided the RB approximation $u_N(\boldsymbol{\mu})$ is sufficiently close to $u_h(\boldsymbol{\mu})$ (which is the case for N sufficiently large), the stability factor $\beta_h(u_N(\boldsymbol{\mu}))$ related to the former can be properly approximated by the stability factor $\beta_h(u_h(\boldsymbol{\mu}))$ related to the latter, making thus possible to estimate the stability factor before assembling the reduced space.

We also remark that a result like (33) holds in case of a general nonlinear operator as long as its Fréchet derivative is Lipschitz continuous, i.e. if there exist $\eta(\boldsymbol{\mu}) > 0$, $L_h^N(\boldsymbol{\mu}) > 0$ such that

$$\|dA(u_N(\boldsymbol{\mu}); \boldsymbol{\mu})(\cdot, \cdot) - dA(v; \boldsymbol{\mu})(\cdot, \cdot)\|_{\mathcal{L}(V_h, V_h')} \leq L_h^N(\boldsymbol{\mu}) \|u_N(\boldsymbol{\mu}) - v\|_V,$$

holds for all $v \in B_\eta(u_N(\boldsymbol{\mu})) = \{w \in V_h : \|u_N(\boldsymbol{\mu}) - w\|_V \leq \eta(\boldsymbol{\mu})\}$. Then,

$$|\beta_h(u_h(\boldsymbol{\mu})) - \beta_h(u_N(\boldsymbol{\mu}))| \leq L_h^N(\boldsymbol{\mu}) \|u_h(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_V \quad \forall \boldsymbol{\mu} \in \mathcal{D}. \quad (35)$$

4 A linearized SCM for estimating the stability factor $\beta_h(u_h(\boldsymbol{\mu}))$

In this section we provide a linearized version of the Successive Constraint Method (SCM) [16] to compute an estimate of the stability factor $\beta_h(u_h(\boldsymbol{\mu}))$. Following [30, 16], we adopt a *natural norm* SCM procedure based on a set of local stability factors, properly computed for a (possibly small) set of J parameter values $\mathcal{S} = \{\boldsymbol{\mu}^{1*}, \dots, \boldsymbol{\mu}^{J*}\}$ selected through a greedy procedure. The key observation is provided by the following relation:

$$\begin{aligned}
\beta_h(u_h(\boldsymbol{\mu})) &= \inf_{v \in V_h} \sup_{w \in V_h} \frac{dA(u_h(\boldsymbol{\mu}); \boldsymbol{\mu})(v, w)}{\|T^{\boldsymbol{\mu}^*} w\|_V \|v\|_V} \frac{\|T^{\boldsymbol{\mu}^*} w\|_V}{\|w\|_V} \\
&\geq \inf_{v \in V_h} \sup_{w \in V_h} \frac{dA(u_h(\boldsymbol{\mu}); \boldsymbol{\mu})(v, w)}{\|T^{\boldsymbol{\mu}^*} w\|_V \|v\|_V} \inf_{w \in V_h} \frac{\|T^{\boldsymbol{\mu}^*} w\|_V}{\|w\|_V} \\
&= \beta_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) \beta_h(u_h(\boldsymbol{\mu}^*)) \geq \tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) \beta_h(u_h(\boldsymbol{\mu}^*)),
\end{aligned}$$

where

$$\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) := \inf_{v \in V_h} \frac{dA(u_h(\boldsymbol{\mu}); \boldsymbol{\mu})(v, T^{\boldsymbol{\mu}^*} v)}{\|T^{\boldsymbol{\mu}^*} v\|_V^2} = \inf_{v \in V_h} \frac{(T^{\boldsymbol{\mu}} v, T^{\boldsymbol{\mu}^*} v)_V}{\|T^{\boldsymbol{\mu}^*} v\|_V^2} \quad (36)$$

is a lower bound of $\beta_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ s.t.

$$\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) \leq \inf_{v \in V_h} \frac{\|T^{\boldsymbol{\mu}} v\|_V}{\|T^{\boldsymbol{\mu}^*} v\|_V} = \inf_{v \in V_h} \sup_{w \in V_h} \frac{dA(u_h(\boldsymbol{\mu}); \boldsymbol{\mu})(v, w)}{\|T^{\boldsymbol{\mu}^*} v\|_V \|w\|_V} =: \beta_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}), \quad (37)$$

thanks to Cauchy-Schwarz inequality and the definition of supremizer operator.

As in the linear case [30], we can show that, for $\boldsymbol{\mu}$ near $\boldsymbol{\mu}^* \in \mathcal{S}$, $\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ is a second-order accurate approximation to $\beta_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ (see Appendix A.2 for the proof):

Proposition 4 *Under the assumptions of Proposition 2, the following relations hold:*

$$\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) - 1 = O(|\boldsymbol{\mu} - \boldsymbol{\mu}^*|) \quad \text{as } \boldsymbol{\mu} \rightarrow \boldsymbol{\mu}^*, \quad (38)$$

$$\beta_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) - \tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) = O(|\boldsymbol{\mu} - \boldsymbol{\mu}^*|^2) \quad \text{as } \boldsymbol{\mu} \rightarrow \boldsymbol{\mu}^*. \quad (39)$$

In particular, (38) guarantees that, for $\boldsymbol{\mu}$ near $\boldsymbol{\mu}^*$, the bilinear form $\pi : V \times V \rightarrow \mathbb{R}$

$$\pi(u, v) = (T^{\boldsymbol{\mu}} u, T^{\boldsymbol{\mu}^*} v)_V$$

is coercive. Thus, we could compute a lower bound to $\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ by applying the SCM algorithm proposed in [18], since this surrogate problem is coercive thanks to (37). However, because of the parameter dependence through the solution $u_h(\boldsymbol{\mu})$, $\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ cannot be expressed as the solution of a *linear program*, which is the key ingredient of SCM in order to provide an efficient offline/online decomposition.

Therefore, we propose to approximate $\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ by the following surrogate:

$$\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) \approx \hat{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) := \inf_{v \in V_h} \frac{dA(u_h(\boldsymbol{\mu}^*); \boldsymbol{\mu})(v, T^{\boldsymbol{\mu}^*} v)}{\|T^{\boldsymbol{\mu}^*} v\|_V^2}. \quad (40)$$

In fact, by using the same argument of Proposition 2, it is possible to show that⁴

$$|\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) - \hat{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})| \leq \frac{\gamma_h^c(\boldsymbol{\mu})}{\beta_h(\boldsymbol{\mu}^*)} L_u^h |\boldsymbol{\mu} - \boldsymbol{\mu}^*|, \quad (41)$$

so that $\hat{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ turns out to be a reasonable approximation to $\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ for $\boldsymbol{\mu}$ sufficiently near to $\boldsymbol{\mu}^*$. The quality of this approximation depends on the ratio $\gamma_h^c(\boldsymbol{\mu})/\beta_h(\boldsymbol{\mu}^*)$, which for $\boldsymbol{\mu} = \boldsymbol{\mu}^*$ is nothing but the condition number of the problem.

Moreover, the approximation $\hat{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ can be obtained by solving a linear program amenable to a suitable offline/online decomposition. In fact, given $\boldsymbol{\mu}^* \in \mathcal{D}$,

$$\hat{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) = \inf_{\mathbf{y} \in \mathcal{Y}_s} \mathcal{J}(\mathbf{y}; \boldsymbol{\mu}), \quad \mathcal{J}(\mathbf{y}; \boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \Theta_q^a(\boldsymbol{\mu}) y_q + \sum_{q'=1}^{Q_c} \Theta_{q'}^c(\boldsymbol{\mu}) y_{Q_a+q'}, \quad (42)$$

⁴ Here L_u^h denotes the Lipschitz constant of the solution map $\boldsymbol{\mu} \rightarrow u_h(\boldsymbol{\mu})$. Thus L_u^h is the discrete counterpart of the Lipschitz constant L_u defined in (56).

with $\mathbf{y} = (y_1, \dots, y_{Q_a}, y_{Q_a+1}, \dots, y_{Q_a+Q_c})$. Here $\mathcal{Y}_* \subset \mathbb{R}^{Q_A}$ (with $Q_A = Q_a + Q_c$) is given by

$$\mathcal{Y}_* = \left\{ \mathbf{y} \in \mathbb{R}^{Q_A} : \exists w_h^{\mathbf{y}} \in V_h \left| \begin{aligned} y_q &= \frac{a_q(w_h^{\mathbf{y}}, T^{\mu^*} w_h^{\mathbf{y}})}{\|T^{\mu^*} w_h^{\mathbf{y}}\|_V^2}, \quad 1 \leq q \leq Q_a, \\ y_{Q_a+q'} &= \frac{d_{q'}(u_h(\mu^*))(w_h^{\mathbf{y}}, T^{\mu^*} w_h^{\mathbf{y}})}{\|T^{\mu^*} w_h^{\mathbf{y}}\|_V^2}, \quad 1 \leq q' \leq Q_c \end{aligned} \right. \right\}.$$

We can now use SCM to build a lower bound of $\hat{\beta}_{\mu^*}(\mu)$ through a sequence of suitable relaxed problems of (42), by seeking the minimum of \mathcal{J} on a descending sequence of larger sets, built by adding successively linear constraints. We also build an upper bound to $\hat{\beta}_{\mu^*}(\mu)$, which will serve to define a suitable *error indicator* in the greedy procedure for the construction of the local lower bound.

We underline that the original SCM would proceed by computing a local lower bound to $\hat{\beta}_{\mu^*}(\mu)$, thus providing a global lower bound to $\beta_h(u_h(\mu))$. Our linearized SCM computes instead a local lower bound to an approximation $\tilde{\beta}_{\mu^*}(\mu)$ of $\hat{\beta}_{\mu^*}(\mu)$, in order to enable the Offline/Online decomposition of the whole procedure. As a result, we obtain a global approximation – rather than a lower bound – to $\beta_h(u_h(\mu))$. We report the details of the procedure in the following subsections.

4.1 Construction of a local lower bound to $\hat{\beta}_{\mu^*}(\mu)$

We first remark that $\hat{\beta}_{\mu^*}(\mu) = (\lambda_{\min}^{\mu^*}(\mu))^{1/2}$, where $\lambda_{\min}^{\mu^*}(\mu)$ is the smallest eigenvalue $\lambda^{\mu^*}(\mu)$ such that $(\lambda^{\mu^*}(\mu), \mathbf{v}) \in \mathbb{R}_+ \times V_h$, $\mathbf{v} \neq 0$, satisfy

$$\frac{1}{2} \left[\mathbb{F}(\mu^*)^T \mathbb{X}^{-1} \hat{\mathbb{F}}(\mu; \mu^*) + \hat{\mathbb{F}}(\mu; \mu^*)^T \mathbb{X}^{-1} \mathbb{F}(\mu^*) \right] \mathbf{v} = \lambda^{\mu^*}(\mu) \mathbb{F}(\mu^*)^T \mathbb{X}^{-1} \mathbb{F}(\mu^*) \mathbf{v}, \quad (43)$$

being $\hat{\mathbb{F}}(\mu; \mu^*)$ the matrix resulting from the discretization of $dA(u_h(\mu^*); \mu)(\cdot; \cdot)$. By extending the procedure presented in [18], we report here the main steps required to construct a lower and an upper bound to $\hat{\beta}_{\mu^*}(\mu)$:

1. *Bounding box construction.* In order to guarantee that (42) is well-posed, we can construct a (continuity) bounding box $B_{\mu^*} \subset \mathbb{R}^{Q_A}$ given by [16]

$$B_{\mu^*} = \prod_{q=1}^{Q_a} \left[-\frac{\gamma_q^a}{\beta_h(u_h(\mu^*))}, \frac{\gamma_q^a}{\beta_h(u_h(\mu^*))} \right] \times \prod_{q'=1}^{Q_c} \left[-\frac{\gamma_{q'}^d(\mu^*)}{\beta_h(u_h(\mu^*))}, \frac{\gamma_{q'}^d(\mu^*)}{\beta_h(u_h(\mu^*))} \right], \quad (44)$$

where $\beta_h(u_h(\mu^*))$ is the solution of (28) computed for $\mu = \mu^*$. Alternatively, as recently proposed in [33], we can consider the following bounding box,

$$B_{\mu^*}^Y = \prod_{q=1}^{Q_a} \left[\inf_{v \in V_h} \frac{a_q(v, T^{\mu^*} v)}{\|T^{\mu^*} v\|_V^2}, \sup_{v \in V_h} \frac{a_q(v, T^{\mu^*} v)}{\|T^{\mu^*} v\|_V^2} \right] \times \prod_{q'=1}^{Q_c} \left[\inf_{v \in V_h} \frac{d_{q'}(u_h(\mu^*))(v, T^{\mu^*} v)}{\|T^{\mu^*} v\|_V^2}, \sup_{v \in V_h} \frac{d_{q'}(u_h(\mu^*))(v, T^{\mu^*} v)}{\|T^{\mu^*} v\|_V^2} \right], \quad (45)$$

which is proved to be tighter than (44), i.e. $B_{\mu^*}^Y \subset B_{\mu^*}$. Let us remark however that the computation of $B_{\mu^*}^Y$ requires additional operations, in particular: (i) for each μ^* the bounding box has to be fully recomputed, while for the former we can compute the γ_q^a 's once and for all, and only update the γ_q^d 's at each iteration; (ii) for each μ^* , $B_{\mu^*}^Y$ requires to compute not only the maximum but also the minimum eigenvalue of the involved bilinear forms. This is a demanding task, which can become unaffordable when Q_a and Q_c become too large. In Section 6 we will show a detailed comparison of these two options.

2. *Relaxed LP problem.* Given a SCM sample $\mathcal{C}_{\mu^*} = \{\mu_1^*, \dots, \mu_k^*\}$ associated to μ^* , compute the corresponding lower bounds $\hat{\beta}_{\mu^*}(\mu')$, by solving (43) $\forall \mu' \in \mathcal{C}_{\mu^*}$; then, define the relaxation set

$$\mathcal{Y}_*^{\text{LB}}(\mathcal{C}_{\mu^*}) = \left\{ \mathbf{y} \in B_{\mu^*} \mid J(\mathbf{y}; \mu') \geq \hat{\beta}_{\mu^*}(\mu'), \quad \forall \mu' \in \mathcal{C}_{\mu^*} \right\}$$

by selecting a set of additional linear constraints associated to \mathcal{C}_{μ^*} . Let us remark that the desired *local lower bound* $\hat{\beta}_{\mu^*}^{\text{LB}}(\mu)$ is provided by the solution of the following relaxed problem:

$$\hat{\beta}_{\mu^*}^{\text{LB}}(\mu) \equiv \hat{\beta}_{\mu^*}^{\text{LB}}(\mu; \mathcal{C}_{\mu^*}) = \inf_{\mathbf{y} \in \mathcal{Y}_*^{\text{LB}}(\mathcal{C}_{\mu^*})} \mathcal{J}(\mathbf{y}; \mu), \quad \forall \mu \in \mathcal{D}_{\mu^*}, \quad (46)$$

since $\hat{\beta}_{\mu^*}(\mu) \geq \hat{\beta}_{\mu^*}^{\text{LB}}(\mu)$. In fact, $\mathcal{Y}_* \subset \mathcal{Y}_*^{\text{LB}}(\mathcal{C}_{\mu^*})$ and thus the minimum is taken over a larger set. Note that (46) has to be solved $\forall \mu \in \Xi_{\text{train}}$ ($\Xi_{\text{train}} \subset \mathcal{D}$ being a very rich training sample), whereas the definition of $\mathcal{D}_{\mu^*} \subset \mathcal{D}$ will be made precise later on. We can also define an upper bound to $\hat{\beta}_{\mu^*}(\mu)$ as follows:

$$\hat{\beta}_{\mu^*}^{\text{UB}}(\mu) \equiv \hat{\beta}_{\mu^*}^{\text{UB}}(\mu; \mathcal{C}_k^*) = \inf_{\mathbf{y} \in \mathcal{Y}_*^{\text{UB}}(\mathcal{C}_{\mu^*})} \mathcal{J}(\mathbf{y}; \mu), \quad \forall \mu \in \mathcal{D}_{\mu^*}, \quad (47)$$

where

$$\mathcal{Y}_*^{\text{UB}}(\mathcal{C}_{\mu^*}) = \{ \tilde{\mathbf{y}} \in \mathbb{R}^{Q_A} : \tilde{\mathbf{y}} = \arg \min_{\mathbf{y} \in \mathcal{Y}_*} \mathcal{J}(\mathbf{y}; \mu'), \quad \forall \mu' \in \mathcal{C}_{\mu^*} \}.$$

Since $\mathcal{Y}_*^{\text{UB}}(\mathcal{C}_{\mu^*}) \subset \mathcal{Y}_*$ – see e.g. [18] for the proof – (47) is in fact an upper bound for $\hat{\beta}_{\mu^*}(\mu)$.

3. *Selection of the successive constraint.* The set \mathcal{C}_{μ^*} is built through a (local) greedy procedure. Starting from $\mathcal{C}_{\mu^*} = \{\mu^*\}$, we iteratively enrich the set \mathcal{C}_{μ^*} by adding the point $\hat{\mu}$ such that

$$\hat{\mu} = \arg \max_{\mu \in E_{\mu^*} \cap \Xi_{\text{train}}} \rho(\mu; \mathcal{C}_{\mu^*}), \quad \rho(\mu; \mathcal{C}_{\mu^*}) = \frac{\hat{\beta}_{\mu^*}^{\text{UB}}(\mu) - \hat{\beta}_{\mu^*}^{\text{LB}}(\mu)}{\hat{\beta}_{\mu^*}^{\text{UB}}(\mu)},$$

until the largest ratio is $\rho(\mu; \mathcal{C}_{\mu^*}) \leq \varepsilon_*$, i.e. under a chosen tolerance $\varepsilon_* \in (0, 1)$. Here we restrict the search for the maximum of $\rho(\cdot; \cdot)$ to a suitable neighborhood E_{μ^*} of μ^* , which shall represent an empirical approximation of the *coercivity region* (see Proposition 4) of $\hat{\beta}_{\mu^*}(\mu)$. The choice of E_{μ^*} is highly problem dependent and is usually made a priori, according to physical intuition, or a posteriori once the first iterations of the algorithm have been run. Further details can be found in Section 6.

Thus, we end up with $K = |\mathcal{C}_{\mu^*}|$ constraints and a local lower bound $\hat{\beta}_{\mu^*}^{\text{LB}}(\mu)$.

4.2 Computation of a global approximation

In order to turn the local lower bound $\hat{\beta}_{\mu^*}^{LB}(\mu)$, computed upon each selected value μ^* , into a global approximation for $\beta_h(\mu)$, we consider a greedy procedure such as the one addressed in [18,16] for the linear case. We remark that the output of the coverage procedure are the set $\mathcal{S} = \{\mu^{1*}, \dots, \mu^{J*}\}$, $J \leq J_{\max}$ and the associated samples $\mathcal{C}_{\mu^{j*}}$, for any $j = 1, \dots, J$, where $K(j) := |\mathcal{C}_{\mu^{j*}}| < K_{\max}$ is the number of constraints points related to each $\mu^{j*} \in \mathcal{S}$. Thus, a global approximation for $\beta_h(\mu)$ is

$$\beta_h^A(\mu) = \beta_h(\mu^{\sigma^*}) \hat{\beta}_{\mu^{\sigma^*}}^{LB}(\mu), \quad \text{being } \sigma \equiv \sigma(\mu) = \arg \max_{j \in \{1, \dots, J\}} \beta_h(\mu^{j*}) \hat{\beta}_{\mu^{j*}}^{LB}(\mu), \quad (48)$$

so that the subdomains $\mathcal{D}_{\mu^{*j}}$ are defined as

$$\mathcal{D}_{\mu^{*j}} = \{\mu \in \mathcal{D} : \beta_h(\mu^{j*}) \hat{\beta}_{\mu^{j*}}^{LB}(\mu) \geq \beta_h(\mu^{j'}) \hat{\beta}_{\mu^{j'}}^{LB}(\mu), \quad \forall j' = 1, \dots, J\}. \quad (49)$$

As in the original SCM, the global approximation $\beta_h^A(\mu)$ interpolates $\beta_h(\mu)$ at each $\mu^* \in \mathcal{S}$, being $\beta_h^A(\mu^*) = \beta_h(\mu^*)$. The set $\mathcal{S} = \{\mu^{1*}, \dots, \mu^{J*}\}$ is built through a (*global*) *greedy procedure*, which encapsulates the local ones used for building each sample. The whole procedure is summarized in the following algorithm.

Linearized SCM algorithm

Input: train sample Ξ_{train} , J_{\max} , K_{\max} , SCM tolerance ϵ_* , starting point μ^{1*} .

set $J = 1$, $\mathcal{C}_{\mu^{1*}} = \{\mu^{1*}\}$, $\mathcal{R}_J = \emptyset$

compute $\beta_h(\mu^{1*})$ by (28) and the bounding box $B_{\mu^{1*}}$

while $J < J_{\max}$, $\Xi_{\text{train}} \neq \emptyset$ and $\rho(\mu) > \epsilon_*$ **do**

 compute $\hat{\beta}_{\mu^{J*}}^{LB}(\mu)$, $\hat{\beta}_{\mu^{J*}}^{UB}(\mu)$

 construct $\mathcal{R}_J^* = \{\mu \in \Xi_{\text{train}} \mid \hat{\beta}_{\mu^{J*}}^{LB}(\mu) > 0 \text{ and } \rho(\mu; \mathcal{C}_{\mu^{J*}}) \leq \epsilon_*\}$

 compute $\beta_h^A(\mu)$ as in (48)

if $\mathcal{R}_J^* \setminus \mathcal{R}_J = \emptyset$ or $|\mathcal{C}_{\mu^{J*}}| = K_{\max}$ **do**

 update $\Xi_{\text{train}} = \Xi_{\text{train}} \setminus \mathcal{R}_J$

 set $J = J + 1$ and select a new μ^{J*}

 compute $\beta_h(\mu^{J*})$ by (28) and the bounding box $B_{\mu^{J*}}$

 set $\mathcal{C}_{\mu^{J*}} = \{\mu^{J*}\}$

 construct $\mathcal{R}_J = \{\mu \in \Xi_{\text{train}} \mid \hat{\beta}_{\mu^{J*}}^{LB}(\mu) > 0 \text{ and } \rho(\mu; \mathcal{C}_{\mu^{J*}}) \leq \epsilon_*\}$

else do

$\hat{\mu} = \arg \max_{\mu \in E_{\mu^{J*}}} \rho(\mu; \mathcal{C}_{\mu^{J*}})$

 set $\mathcal{C}_{\mu^{J*}} = \mathcal{C}_{\mu^{J*}} \cup \{\hat{\mu}\}$

 compute $\hat{\beta}_{\mu^*}(\hat{\mu})$ by solving (43)

 set $\mathcal{R}_J = \mathcal{R}_J^*$

end if

end while

Let us highlight which are the main computational costs of this problem. We denote by $n_{\text{train}} = |\Xi_{\text{train}}|$ and we define

$$n_{\beta} = \sum_{j=1}^J |\mathcal{C}_{\mu^{j*}}|, \quad n_c = \max_{j=1, \dots, J} |\mathcal{C}_{\mu^{j*}}|.$$

In the offline stage we have to: (i) solve J times problem (20) in order to compute $u_h(\boldsymbol{\mu})$ and assemble n_{β} times the Fréchet derivative; (ii) solve $n_{\text{eig}}^{(1)} = n_{\beta} + Q_a + JQ_c$ (respectively $n_{\text{eig}}^{(2)} = n_{\beta} + 2JQ_a + 2JQ_c$) eigenproblems when using the bounding box (44) (respectively (45)), (iii) solve $n_{\text{train}}n_{\beta}$ linear programs to compute the current global lower bounds (48) at each iteration of the algorithm.

In the online stage, each evaluation $\boldsymbol{\mu} \rightarrow \beta_h(u_h(\boldsymbol{\mu}))$ only requires to solve J linear programs in $Q_A = Q_a + Q_c$ variables with at most $n_c + 2Q_A$ constraints (independently of the employed bounding box).

Remark 2 As in the linear case, the computational complexity of the offline stage of the SCM depends inherently on $Q_A \mathcal{N}_h^{\alpha}$, where the dependence on the dimension \mathcal{N}_h is due to eigenvalues calculation (with $\alpha \in [1, 3]$). Thus, already for rather small problems, the size Q_A of the affine expansion may cause the Offline stage to become potentially very expensive. A *two-level* affine decomposition strategy was recently proposed in [20] to tackle the case of large affine operators (e.g. recovered through the empirical interpolation method).

5 A new heuristic strategy based on adaptive interpolation

Our direct experience indicates a rather slow convergence of the linearized SCM procedure when dealing with many ($p \geq 3$) parameters (see also the numerical results of Section 6). This prompts us to devise alternative strategies when dealing with nonlinear operators depending on many parameters.

A first, very simple approach would be to approximate the ($\boldsymbol{\mu}$ -dependent) stability factor $\beta_h(\boldsymbol{\mu})$ by the constant

$$\beta_{\text{LB}} = \min_{\boldsymbol{\mu} \in \mathcal{D}} \beta_h(\boldsymbol{\mu}). \quad (50)$$

Since $\beta_h(\boldsymbol{\mu})$ might be a non-convex function of $\boldsymbol{\mu}$, finding its global minimum on \mathcal{D} requires (i) to combine a local optimization solver with a suitable globalization strategy [14] and possibly (ii) to provide an explicit expression for the sensitivity of $\beta_h(\boldsymbol{\mu})$ with respect to the parameters. This approach is indeed effective when the stability factor changes mildly with respect to parameters. However, as soon as the dimension of \mathcal{D} increases, finding a global minimum becomes extremely expensive, not to mention that this strategy is over-conservative (thus inappropriate) when $\beta_h(\boldsymbol{\mu})$ varies significantly over \mathcal{D} .

For these reasons, we propose a heuristic strategy devised to meet an efficiency requisite (at both the offline and the online stages), which returns reliable and sufficiently tight approximations to parametrized stability factors.

5.1 Interpolant of the stability factor

Let us denote by $\Xi_{\text{fine}} \subset \mathcal{D}$ a sample set whose dimension $n_{\text{fine}} = |\Xi_{\text{fine}}|$ is sufficiently large. We (arbitrarily and a priori) select a (possibly *small*) set of interpolation points $\Xi_I = \{\boldsymbol{\mu}^j\}_{j=1}^{n_I} \subset \Xi_{\text{fine}}$ and compute the stability factor $\beta_h(\boldsymbol{\mu})$ for each $\boldsymbol{\mu} \in \Xi_I$. Then, we compute a suitable *interpolant* $\beta_I(\boldsymbol{\mu})$ such that

$$\beta_I(\boldsymbol{\mu}) = \beta_h(\boldsymbol{\mu}) \quad \forall \boldsymbol{\mu} \in \Xi_I \quad \text{and} \quad \beta_I(\boldsymbol{\mu}) > 0 \quad \forall \boldsymbol{\mu} \in \Xi_{\text{fine}}.$$

For any fixed $\boldsymbol{\mu} \in \mathcal{D}$, the computation of $\beta_h(\boldsymbol{\mu})$ requires to solve the following eigenvalue problem: find $(\lambda(\boldsymbol{\mu}), \mathbf{v}) \in \mathbb{R}_+ \times V_h$, $\mathbf{v} \neq 0$, such that

$$\mathbb{F}(\boldsymbol{\mu})^T \mathbb{X}^{-1} \mathbb{F}(\boldsymbol{\mu}) \mathbf{v} = \lambda(\boldsymbol{\mu}) \mathbb{X} \mathbf{v}, \quad (51)$$

where \mathbb{X} is the matrix defined by (26), while $\beta_h(\boldsymbol{\mu}) = \sqrt{\lambda_{\min}(\boldsymbol{\mu})}$.

Depending on the number of parameters and their range of variation, different interpolation methods might be employed. In the two-dimensional case considered in [24] we used a simple linear interpolant and an equally spaced grid of interpolation points. When the parameter space has higher dimension, using uniform grids would demand for $\beta_h(\boldsymbol{\mu})$ to be computed in a huge number of interpolation points. Following [22], we then replace Lagrange interpolation by *radial basis function* (RBF) interpolation. The latter is especially suited to interpolate scattered data in high-dimensional spaces (for a general introduction to RBF methods see, e.g., [4]). We define the RBF interpolant as

$$\beta_I(\boldsymbol{\mu}) = \omega_0 + \boldsymbol{\omega}^T \boldsymbol{\mu} + \sum_{j=1}^{n_I} \gamma_j \phi(|\boldsymbol{\mu} - \boldsymbol{\mu}^j|),$$

where ϕ is a radial basis function⁵, while the $1 + p + n_I$ interpolation weights $\{\omega_i\}_{i=0}^p$, $\{\gamma_j\}_{j=1}^{n_I}$ are determined by requiring the following conditions to hold:

$$\beta_I(\boldsymbol{\mu}^j) = \beta_h(\boldsymbol{\mu}^j) \quad j = 1, \dots, n_I, \quad (52a)$$

$$\sum_{j=1}^{n_I} \gamma_j = 0, \quad \sum_{j=1}^{n_I} \gamma_j \boldsymbol{\mu}_i^j = 0 \quad i = 1, \dots, p. \quad (52b)$$

Equations (52a)-(52b) lead to a symmetric linear system of dimension $1 + p + n_I$.

In order to avoid negative values of the interpolant $\beta_I(\boldsymbol{\mu})$, we first perform the interpolation on a starting grid Ξ_I of $n_I^{(0)}$ interpolation points and then evaluate the resulting interpolant on the fine grid. Note that the first step of the algorithm can be performed in parallel. Next, we enrich the interpolation grid Ξ_I by adding further interpolation points in those regions where the interpolant is negative. This yields a positive (provided that n_{max} is sufficiently large) interpolant $\beta_I(\boldsymbol{\mu})$. If n_{neg} denotes the number of points selected by the second step of the algorithm, we must in the offline stage (i) solve at most $n_I^{(0)} + n_{\text{neg}}$ eigenvalue problems; (ii) build $1 + n_{\text{neg}}$ times the RBF interpolant, involving $O(n_I^3 + n_I^2)$ operations (being n_I the dimension of the adaptively enriched set Ξ_I); (iii) evaluate $n_I^{(0)} + n_{\text{neg}}$ times the RBF interpolant, requiring $O(n_I n_{\text{fine}})$ operations.

⁵ In the numerical results of Section 6 we employ thin plate splines RBF, i.e. $\phi(r) = r^2 \log(r)$.

This strategy yields a good approximation of the stability factor, with a remarkably smaller computational effort with respect to the linearized SCM algorithm. However, as the dimension (and extent) of \mathcal{D} increases, the efficacy of this procedure highly depends on the full factorial grid Ξ_I adopted at the first step. Indeed, if Ξ_I is too coarse, most of the time is spent in the second step of the algorithm trying to ensure the positivity of the interpolant, and eventually resulting in a poor approximation. On the other hand, if Ξ_I is too fine, the additional computational effort can be often unnecessary, since too many points are added in regions where $\beta_h(\boldsymbol{\mu})$ changes mildly. We overcome these inconveniences by further improving this method in the following section.

5.2 RBF interpolant with adaptive sampling

In order to achieve a compromise between (i) adding new points in locations with a highly varying response, (ii) adding points in unsampled regions of the domain and (iii) ensuring the positivity of the interpolant, we propose an adaptive strategy based on a four-component criterion $C(\boldsymbol{\mu})$:

$$C(\boldsymbol{\mu}) = (\|\nabla\beta_I(\boldsymbol{\mu})\| + \epsilon)(|\Delta\beta_I(\boldsymbol{\mu})| + \epsilon) \left(\frac{h(\boldsymbol{\mu})}{\max h(\boldsymbol{\mu})} \right)^2 g(\beta_I(\boldsymbol{\mu})), \quad (53)$$

similarly to what proposed in [21]. Let us describe the role of each factor:

- the first two terms account for local changes of the interpolant⁶; an offset parameter ϵ ensures that $C(\boldsymbol{\mu}) > 0$ when $\|\nabla\beta_I(\boldsymbol{\mu})\| = 0$ or $|\Delta\beta_I(\boldsymbol{\mu})| = 0$;
- the third and the fourth terms promote the selection of space-filling points and penalize negative values of the interpolant, respectively, where

$$h(\boldsymbol{\mu}) = \min_{\boldsymbol{\mu}^j \in \Xi_I} \|\boldsymbol{\mu} - \boldsymbol{\mu}^j\|_2, \quad g(s) = \begin{cases} 1 & s > 0 \\ \alpha e^{-s} & s \leq 0 \end{cases}$$

and $\alpha > 0$ is a tuning parameter to be prescribed.

In this adaptive algorithm, the new sample locations are then selected as the ones which maximize C over Ξ_{fine} . This optimization problem is solved by enumeration, i.e. by evaluating C on the fine grid Ξ_{fine} and extracting the maximum, rather than using a global optimization algorithm. Indeed, evaluating C over Ξ_{fine} is a very fast operation (with complexity $O(n_I^2 n_{fine})$) which can be easily performed in parallel. The loop stops when either a predetermined number of interpolation points have been added, or a desired accuracy is reached. The complete algorithm reads as follows⁷:

⁶ Note that the derivatives of the interpolant are available analytically and, therefore, both $\nabla\beta_I(\boldsymbol{\mu})$ and $\Delta\beta_I(\boldsymbol{\mu})$ can be computed exactly, rather than approximated numerically.

⁷ The parameters domain \mathcal{D} is normalized to the unit hypercube $[0, 1]^p$ for the sake of interpolation; in this way, the results of the interpolation are not affected by possible different scales and range of variations of the parameters.

Adaptive RBF interpolant

Input: the evaluation grid Ξ_{fine} , a set of $n_I^{(0)}$ starting samples Ξ_I, n_{max}

(i) *Build initial coarse interpolation*

for $j = 1 : n_I^{(0)}$

 set $\boldsymbol{\mu}^j = \Xi_I(j)$ and assemble $\mathbb{F}(\boldsymbol{\mu}^j)$

 compute $\beta_h(\boldsymbol{\mu}^j)$ by solving the eigenvalue problem (51)

end for

build the RBF interpolant $\beta_I(\boldsymbol{\mu})$

(iii) *Enrich interpolation with adaptive sampling*

evaluate $\beta_I(\boldsymbol{\mu})$ on Ξ_{fine}

while $j < n_{\text{max}}$ and $E_j > \text{tol}$ **do**

 compute criterion $C(\boldsymbol{\mu})$ as defined in (53)

 set $\boldsymbol{\mu}^j = \arg \max_{\boldsymbol{\mu} \in \Xi_{\text{fine}}} C(\boldsymbol{\mu})$ and assemble $\mathbb{F}(\boldsymbol{\mu}^j)$

 compute $\beta_h(\boldsymbol{\mu}^j)$ by solving the eigenvalue problem (51)

 build the RBF interpolant $\beta_{I \cup \boldsymbol{\mu}^j}(\boldsymbol{\mu})$

 evaluate $E_j = \max_{\boldsymbol{\mu} \in \Xi_{\text{fine}}} |\beta_I(\boldsymbol{\mu}) - \beta_{I \cup \boldsymbol{\mu}^j}(\boldsymbol{\mu})| / |\beta_I(\boldsymbol{\mu})|$

 update the set $\Xi_I = \Xi_I \cup \{\boldsymbol{\mu}^j\}$

end while

Several techniques can be employed to assess the interpolation accuracy and possibly estimate the interpolation error, see [21] and references therein for further details. Here we simply require the $L^\infty(\Xi_{\text{fine}})$ -norm of two consecutive iterates to be under a prescribed tolerance tol . Let us remark that the offline costs are just slightly increased with respect to the interpolation technique of Sect. 5.1, since the number of operations required to evaluate $C(\boldsymbol{\mu})$ only depends on n_I and n_{fine} , and is independent of \mathcal{N}_h . Finally, we remark that the condition number of the RBF matrix arising from (52a)-(52b) rapidly increases as the number of interpolation points increase, although our greedy sampling helps in delaying this behavior. However, when a large number of interpolation points is needed, *ad hoc* preconditioning strategies [1] or suitable choices of RBF shape parameters can be put in place [12].

6 Numerical results: application to a backward facing step channel

In this section we illustrate the properties and the performances of the proposed techniques. As a test case, we consider a fluid flow over a backward facing step channel [2], described by the steady Navier-Stokes equations:

$$\begin{aligned}
 -\nu \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p &= \mathbf{0} && \text{in } \Omega_o(\boldsymbol{\mu}) \\
 \operatorname{div} \mathbf{u} &= 0 && \text{in } \Omega_o(\boldsymbol{\mu}) \\
 \mathbf{u} &= \mathbf{g} && \text{on } \Gamma_d^o \\
 \mathbf{u} &= \mathbf{0} && \text{on } \Gamma_w^o(\boldsymbol{\mu}) \\
 -p \mathbf{n} + \nu (\nabla \mathbf{u}) \mathbf{n} &= \mathbf{0} && \text{on } \Gamma_n^o(\boldsymbol{\mu}),
 \end{aligned} \tag{54}$$

where (\mathbf{v}, p) are the velocity and pressure defined over a parametrized domain $\Omega_o(\boldsymbol{\mu}) = \Omega_{o1} \cup \Omega_{o2}(\boldsymbol{\mu}) \cup \Omega_{o3}(\boldsymbol{\mu})$ (see Fig. 1). We denote by $\Gamma_D^o = \Gamma_d^o \cup \Gamma_w^o(\boldsymbol{\mu})$ the

Dirichlet portion of $\partial\Omega_o$, while $\Gamma_n^o(\boldsymbol{\mu})$ denotes the outflow boundary. We define the Reynolds number as $\text{Re} = D\mathbf{u}_b/\nu$, where ν is the kinematic viscosity, $D = 2h$ being $h = 1$ the height of the channel at the inflow, while $\mathbf{u}_b = 2/3 \max \mathbf{g} = 1$, being $\mathbf{g} = [6y(1-y), 0]^T$ the inflow profile. We consider $p = 3$ parameters: the Reynolds number $\mu_1 = \text{Re}$ (so that $\nu = 2/\mu_1$), the step height μ_2 and the channel length μ_3 (downstream of the step).

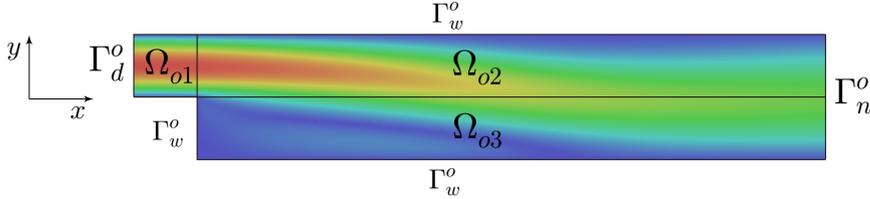


Fig. 1: Sketch of the channel geometry with boundaries and partition in affine subdomains. The first subdomain is $\boldsymbol{\mu}$ -independent, while $\Omega_{o2} = \Omega_{o2}(\mu_3)$ and $\Omega_{o3} = \Omega_{o3}(\mu_2, \mu_3)$. Coloring is given by the velocity field magnitude obtained for $\text{Re} = 250$.

Problem (54) can be rewritten in an affinely parametrized weak form. To do that, we first introduce a decomposition of $\Omega_o(\boldsymbol{\mu})$ into three subdomains (see Figure 1) and a suitable affine geometrical transformation, then we map the problem onto a fixed, reference domain Ω ; further details can be found, e.g., in [10, 22, 32].

Once the problem has been formulated as in (3), we introduce its FE discretization. We use a (inf-sup stable) \mathbf{P}_b^1 - P^1 approximation for the velocity and pressure variables, i.e. continuous linear FE enriched by bubble functions for the velocity and continuous linear FE for the pressure, see e.g. [26]. The total number of degrees of freedom is $\mathcal{N}_h = 40\,064$, obtained using a mesh of 11 485 triangular elements. For the solution of (54), we employ a few Picard iterations followed by some Newton iterations, to reach a relative tolerance of 10^{-8} on the norm of the increment. To facilitate the computation of extreme eigenvalues, we consider a weighted norm on V : for any $v = (\mathbf{v}, q) \in V$, $\|v\|_V^2 := \tilde{a}(\mathbf{v}, \mathbf{v}; \hat{\boldsymbol{\mu}}) + \lambda\|\mathbf{v}\|_{\mathbf{L}^2}^2 + \lambda\|q\|_{L^2}^2$, where $\hat{\boldsymbol{\mu}}$ is a reference parameter value (for instance the centroid of the parameter space), $\tilde{a}(\cdot, \cdot, \boldsymbol{\mu})$ corresponds to the diffusion term in the momentum equation, while

$$\lambda = \inf_{\mathbf{w} \in [H^1(\Omega)]^2} \frac{\tilde{a}(\mathbf{w}, \mathbf{w}; \hat{\boldsymbol{\mu}})}{\|\mathbf{w}\|_{\mathbf{L}^2}^2} > 0.$$

The research code we use in this work has been developed in the Matlab environment; all the linear systems are solved by the sparse direct solver provided by Matlab, whereas the eigenproblems are solved using Matlab `eigs` solver. We also take advantage of the existing SCM algorithm already developed (for linear problems) in the `rbMIT` library [17]. Parallelism is exploited to speed up the matrix assembly in the Navier-Stokes solver as well as to speed up some *embarrassingly parallel* portions of the algorithms we propose. The reported computational times will mainly serve to compare the different strategies.

6.1 Backward-facing step channel with a physical parameter

In this first test case we only consider the Reynolds number $\mu_1 \in [20, 250]$ as varying parameter, while the geometrical parameters are frozen to $\mu_2 = 1$ and $\mu_3 = 10$. The affine decomposition (12) is recovered for $Q_a = 3$, $Q_c = 1$ and $Q_f = 3$.

First, we numerically verify the inequality (33) proved in Proposition 3. We build the RB space following the procedure described in [23]: we select $N = 11$ basis functions to obtain a maximum error $\|u_h(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_V$ below 10^{-3} on the whole parameter space. In Figure 2 we report the graphs of the left- and right-hand sides of (33) with respect to N (computed on a test sample of 20 parameter values and then averaged).

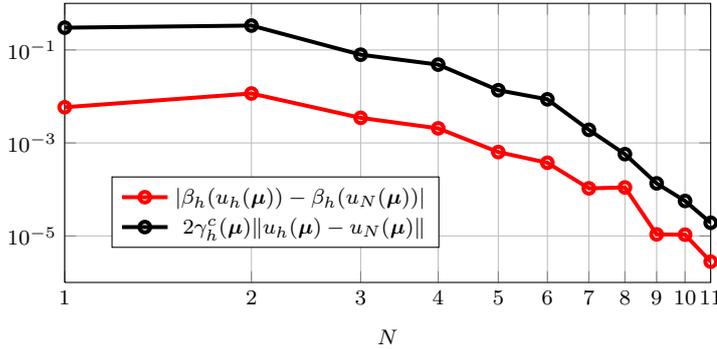


Fig. 2: Test 1. Numerical verification of Proposition 3.

Then, we numerically verify the coercivity property of the local lower bounds $\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$, i.e. that

$$\tilde{\beta}_{\boldsymbol{\mu}^*} = 1 + O(|\boldsymbol{\mu} - \boldsymbol{\mu}^*|) \quad \text{as } \boldsymbol{\mu} \rightarrow \boldsymbol{\mu}^*$$

as shown in Proposition 4. For the sake of verification, we select “by hand” $J = 9$ parameter points $\boldsymbol{\mu}^{j*}$, $1 \leq j \leq J$, and compute the corresponding $\tilde{\beta}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu})$, which are reported in Figure 3. As expected, for each $\boldsymbol{\mu}^{j*}$, $\tilde{\beta}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu})$ decreases linearly from 1. In Figure 3 we also report the approximation $\hat{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ to $\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$; as expected from estimate (40), the quality of the approximation deteriorates as the Reynolds number – and thus the condition number of the problem – increases. Note that for $\boldsymbol{\mu}$ sufficiently far from $\boldsymbol{\mu}^*$, $\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ and its approximation $\hat{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ become negative, and are therefore useless in order to build a positive global approximation to $\beta_h(\boldsymbol{\mu})$. For this reason, we restrict the search for successive constraints $\mathcal{C}_{\boldsymbol{\mu}^*}$ to the interval $E_{\boldsymbol{\mu}^*} = [\boldsymbol{\mu}^* - 20, \boldsymbol{\mu}^* + 20]$.

Let us now apply the linearized SCM algorithm with a tolerance $\epsilon_* = 0.7$, $n_{\text{train}} = 1000$ and using the original bounding box (44). The greedy procedure selects $J = 6$ anchor points $\boldsymbol{\mu}^*$ and $|\mathcal{C}_{\boldsymbol{\mu}^{j*}}|_{j=1}^6 = [3, 3, 6, 2, 2, 1]$ constraints, thus requiring to solve $n_{\text{eig}}^1 = 26$ eigenproblems. In Figure 4 we report the resulting global approximation $\beta_h^{\text{A}}(\boldsymbol{\mu})$ and the subdomains partition $\mathcal{D}_{\boldsymbol{\mu}^{*j}}$ induced by the algorithm.

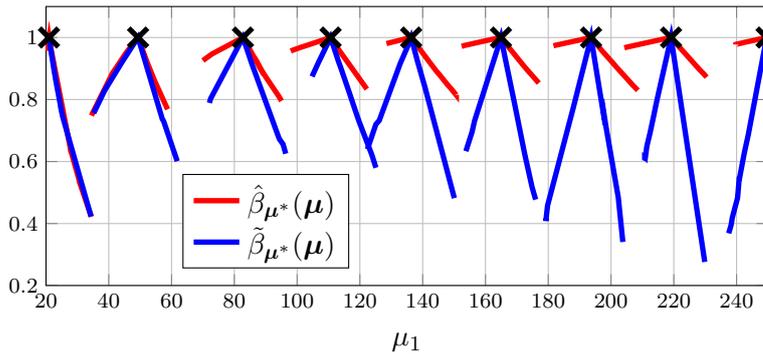


Fig. 3: Test 1. $\tilde{\beta}_{\mu^{j^*}}(\boldsymbol{\mu})$ and its approximation $\hat{\beta}_{\mu^{j^*}}(\boldsymbol{\mu})$, computed in the proximity of $J = 9$ (artificially imposed) parameter points $\boldsymbol{\mu}^{j^*}$, $1 \leq j \leq J$.

Then, in the same setting, we apply the linearized SCM using the tighter bounding box (44): we obtain $J = 4$, $|\mathcal{C}_{\mu^{j^*}}|_{j=1}^4 = [3, 3, 2, 1]$ and $n_{\text{eig}}^{(2)} = 41$. Regarding the computational performances, the two options require roughly the same time (about 20 minutes) to be performed. The (tighter) bounding box (45) lead to a sharper approximation (see Figure 4), yet selecting a smaller number of anchor points $\boldsymbol{\mu}^*$. However, it globally requires to solve a higher number of eigenvalue problems. In particular, its computational complexity depends on the number of terms Q_a and Q_c in the affine decomposition.

| | Bounding box (44) | Bounding box (45) |
|--|--------------------|-------------------|
| J (number of selected $\boldsymbol{\mu}^*$) | 6 | 4 |
| $ \mathcal{C}_{\mu^{j^*}} _{j=1, \dots, J}$ | [3; 3; 6; 2; 2; 1] | [3; 3; 2; 1] |
| Number of eigenproblems | 26 | 41 |
| SCM tolerance ϵ_* | 0.7 | 0.7 |
| Total time (s) | 1088 | 1191 |

Table 1: Test 1. Comparison of the computational cost of the linearized SCM algorithm when using the bounding box (44) and (45).

Let us now move to the other heuristic strategies. We first compute the *minimum stability factor* by performing a multi-start optimization with three different initial points: we find a minimum stability factor $\beta_{\text{LB}} = 0.1025$ (attained for $\mu_1 = 250$), which turns out to be the global minimum of $\beta_h(\boldsymbol{\mu})$ (see Figure 5). In this case, the algorithm requires to solve 45 eigenproblems. We highlight the importance of the multi-start strategy; indeed, if we start the optimization from $\mu_1 < 40$, the algorithm converges to the local minimum attained for $\mu_1 = 20$, thus largely overestimating the global one.

Then, we compute the adaptive RBF interpolant $\beta_I(\boldsymbol{\mu})$: starting from an initial coarse grid of 4 (uniformly distributed) interpolation points, the adaptive procedure selects 10 additional interpolation points so that $E_j < 10^{-3}$ (see Figure 6). The effectiveness of the adaptivity criterion is demonstrated by the evidence that most of the interpolation points are added in the region with the highest variation of $\beta_h(\boldsymbol{\mu})$, as it can be seen in Figure 5. In this case the construction of the in-

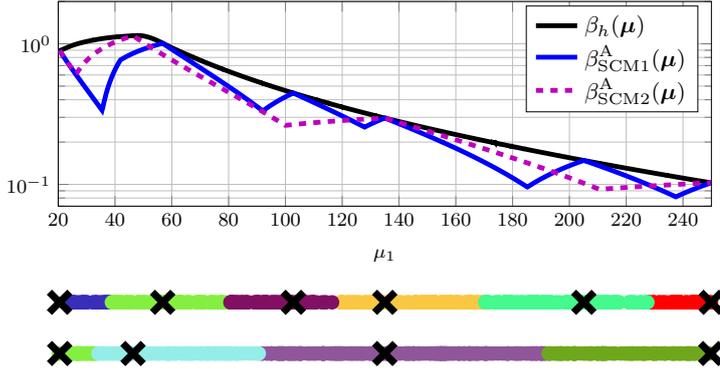


Fig. 4: Test 1. Comparison between the approximation of the stability factor obtained using the linearized SCM algorithm with different bounding box: β_{SCM1}^A refers to (44), while β_{SCM2}^A is obtained using (45). We also report the subdomains (49) induced by the algorithm, with different colors (top: β_{SCM1}^A , bottom: β_{SCM2}^A). The corresponding μ^{j*} , $1 \leq j \leq J$ are represented by black crosses.

terpolant only requires to solve 14 eigenproblems, taking about 8 minutes. Let us remark that, while the final interpolant is almost *coincident* with the exact stability factor, already the initial one (computed on the coarse grid) can be considered as a satisfactory approximation for our purposes.

In Figure 6 we also report a convergence analysis of the RBF interpolation comparing the adaptive versus the uniform refinement of the interpolation grid; in particular, we show the convergence of the $L^\infty(\Xi_{test})$ relative error between the stability factor and its interpolant, where $\Xi_{test} \subset \mathcal{D}$ is a uniform grid of 1000 points. We remark that the adaptive strategy allows to achieve the same accuracy with a considerably smaller number of interpolation points.

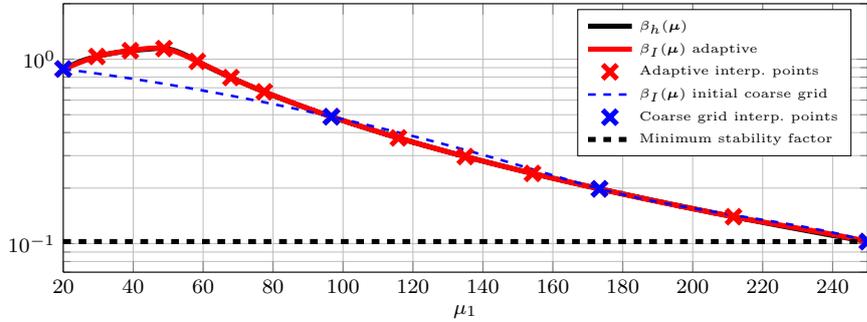


Fig. 5: Test 1. Comparison of the heuristic strategies. The value of the *minimum stability factor* and the RBF interpolant with respect to the true stability factor $\beta_h(\mu)$ (black line) are reported. For the latter, both the initial interpolation on a coarse grid of 4 points (blue dashed line) and the result of the adaptive strategy (red line) are shown.

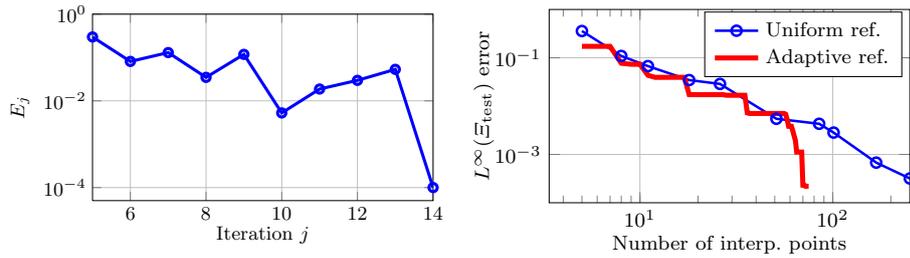


Fig. 6: Test 1. On the left: indicator E_j (used in the adaptive algorithm to monitor the accuracy of the interpolation) versus the number of iterations j . On the right: convergence of the $L^\infty(\Xi_{\text{test}})$ relative error between $\beta_h(\boldsymbol{\mu})$ and $\beta_I(\boldsymbol{\mu})$ with respect to the number of interpolation points (in the case of adaptive refinement, we stop the algorithm when a maximum budget of $n_{\text{max}} = 75$ points has been reached).

| | # eigenproblems | Time (s) |
|--------------------------|-----------------|----------|
| Linearized SCM with (44) | 26 | 1088 |
| Linearized SCM with (45) | 41 | 1191 |
| Minimum | 45 | 1320 |
| Adaptive RBF interpolant | 14 | 470 |

Table 2: Test 1. Comparison of the computational costs. Computations have been performed using 4 cores on a desktop computer.

6.2 Backward-facing step channel with both physical and geometrical parameters

In the second test case, we consider as parameters both the Reynolds number μ_1 and the height of the channel step μ_2 ; the parameter space is now given by $\mathcal{D} = [20, 200] \times [0.5, 1.5]$. We have slightly restricted the range of the parameter μ_1 to avoid numerical instabilities (due to the poor convergence of the nonlinear solver) occurring for high values of μ_1 and μ_2 . The affine decomposition (12) now holds with $Q_a = 5$ and $Q_c = 2$.

We first run the linearized SCM algorithm with a tolerance $\epsilon_* = 0.85$, $n_{\text{train}} = 10^4$ and using the original bounding box (44). The algorithm shows an extremely low convergence: $J = 195$ parameter values $\boldsymbol{\mu}^{1*}, \dots, \boldsymbol{\mu}^{J*}$ and about 750 constraints are selected, requiring to solve $O(10^3)$ eigenproblems. Its poor convergence rate is mainly due to the geometrical variation induced by the parameter μ_2 . Indeed, by running the SCM with the first parameter frozen to $\mu_1 = 100$, the computation of $\beta_h^A(\boldsymbol{\mu})$ (shown in Figure 8) required $J = 17$ parameter values $\boldsymbol{\mu}^{j*}$, a rather large number compared to the results of the previous section.

Then, we build the adaptive RBF interpolant, starting from a coarse grid of uniformly distributed 4×3 interpolation points. The algorithm stops after selecting 23 further samples, corresponding to a maximum budget of 35 interpolation points and $E_j \approx 2 \cdot 10^{-2}$. Note that this latter underestimates the interpolation error of one order of magnitude; in fact, as shown in Figure 9 we are approximating $\beta_h(\boldsymbol{\mu})$ with a relative $L^\infty(\Xi_{\text{fine}})$ error of about 10^{-1} . Nevertheless, the qualitative behavior of the stability factor is well captured. Moreover, the construction of the interpolant $\beta_I(\boldsymbol{\mu})$ takes less than 20 minutes, while the linearized SCM algo-

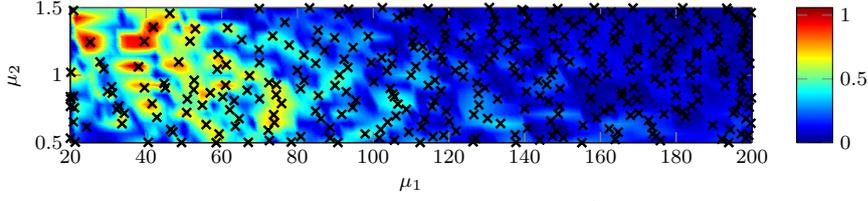


Fig. 7: Test 2. Approximation of the stability factor $\beta_h^A(\boldsymbol{\mu})$ as function of (μ_1, μ_2) ; black cross correspond to the parameter values $\boldsymbol{\mu}^{j*}$, $1 \leq j \leq J = 195$, selected by the linearized SCM algorithm.

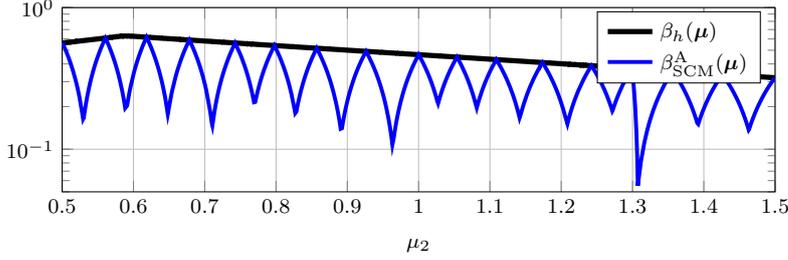


Fig. 8: Test 2. Approximation of the stability factor $\beta_h^A(\boldsymbol{\mu})$ as function of μ_2 , obtained running the linearized SCM algorithm with $\mu_1 = 100$ fixed. Despite the low variation of $\beta_h(\boldsymbol{\mu})$, SCM requires many iterations to converge (indeed $J = 17$).

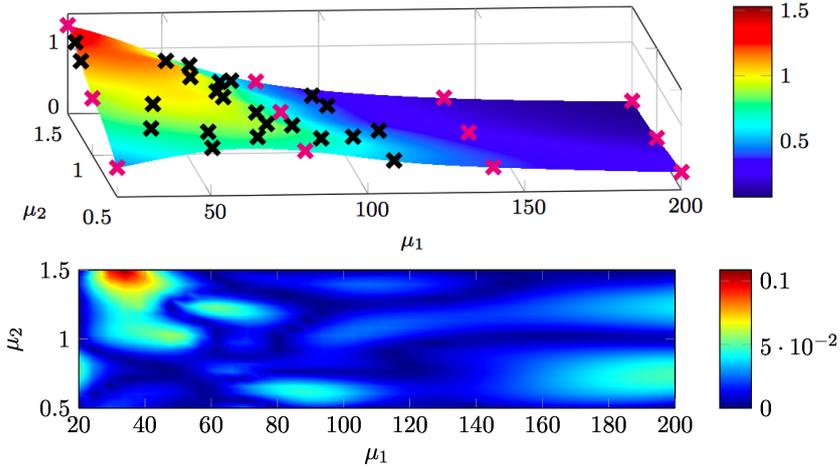


Fig. 9: Test 2. Top: RBF interpolant $\beta_I(\boldsymbol{\mu})$, initial coarse grid (magenta) and interpolation points (black) selected by the adaptive procedure. Bottom: relative error between the stability factor $\beta_h(\boldsymbol{\mu})$ and its RBF interpolant.

rithm requires many hours to build $\beta_h^A(\boldsymbol{\mu})$ (in both cases computations have been performed using 12 cores).

In Figure 10 we also report a convergence analysis of the RBF interpolation comparing the adaptive versus the uniform refinement of the interpolation grid; in this case $\mathcal{E}_{\text{test}} \subset \mathcal{D}$ is a factorial grid of 129×65 points.

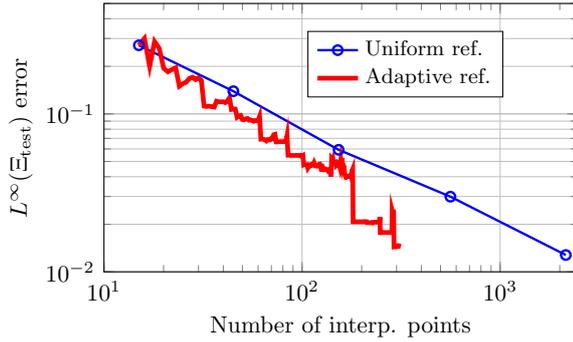


Fig. 10: Test 2. Convergence of the $L^\infty(\Xi_{\text{test}})$ relative error between $\beta_h(\boldsymbol{\mu})$ and $\beta_I(\boldsymbol{\mu})$ with respect to the number of interpolation points.

6.3 Backward-facing step channel with three parameters

In the third test case we let all the three parameters vary; in particular the parameter domain is now given by $\mathcal{D} = [20, 200] \times [0.5, 1.5] \times [9, 12]$, while $Q_a = 9$ and $Q_c = 4$.

It follows from our previous discussion that both the linearized SCM and the minimum stability factor strategies are no longer viable in this case. The adaptive RBF interpolant represents the only chance to obtain, with a reasonable (and somehow predictable) computational effort, a satisfactory approximation of the stability factor.

We start with a coarse grid of $3 \times 3 \times 3$ uniformly distributed interpolation points, and then we let the adaptive procedure select additional 18 samples. We report in Figure 11 the resulting approximation of the stability factor; once again, the adaptive criterion promotes the selection of interpolation points in the regions featuring the highest variations of $\beta_h(\boldsymbol{\mu})$. In Figure 12 we compare the stability factor and the interpolant $\beta_I(\boldsymbol{\mu})$ in the setting of the first test case, i.e. as functions of μ_1 , with $\mu_2 = 1$ and $\mu_3 = 10$ fixed; in the same figure we also report the interpolant obtained using 20 more points adaptively selected.

Once again, the adaptive procedure correctly selects the interpolation points in the most varying regions, so that a tight approximation can be easily obtained with a moderate computational effort. As a matter of fact, we experienced that the linearized SCM algorithm tends to select control points $\boldsymbol{\mu}^*$ from subregions of \mathcal{D} where the PDE solution (as well as the corresponding eigenpair) – rather than the stability factor – is more sensitive to changes in the parameters. On the other hand, the adaptive interpolation is only affected by the parametric response of the stability factor, enhancing the computational efficiency of this latter.

7 Approaching a singular point: the channel expansion case

In order to show the robustness of the adaptive interpolation strategy, we tackle a limit-case problem where multiple steady state solutions coexist as the result of a symmetry-breaking pitchfork bifurcation. In particular, we consider a two-dimensional laminar flow through a channel featuring a sudden expansion, whose

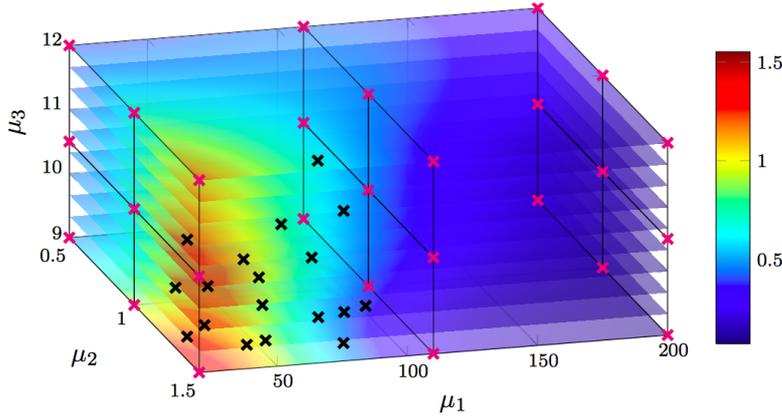


Fig. 11: Test 3. Slices of the adaptive RBF interpolant $\beta_I(\boldsymbol{\mu})$ for different values of μ_3 ; we report the initial full factorial grid (magenta) of $3 \times 3 \times 3$ points and the 18 interpolation points (black) selected by the adaptive procedure.

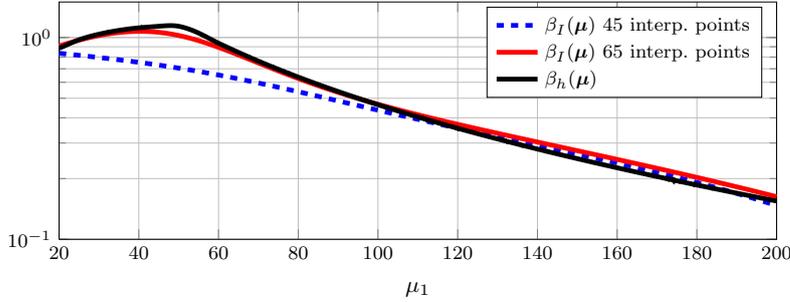


Fig. 12: Test 3. RBF interpolant $\beta_I(\boldsymbol{\mu})$ as a function of μ_1 , with $\mu_2 = 1$ and $\mu_3 = 10$ fixed (as in the setting of the first numerical test). We compare the adaptive RBF interpolants obtained using 45 (dashed blue line) and 65 (red line) interpolations points (in the whole parameter space) w.r.t the true stability factor $\beta_h(\boldsymbol{\mu})$. Note that none of the interpolation points lies along the $(\mu_1, 1, 10)$ line.

geometry and boundaries are reported in Figure 13. We define the Reynolds number as $\text{Re} = U h_1 / \nu$, where h_1 is the inlet height, while the characteristic velocity is $U = 2/3 \max \mathbf{g}$, being $\mathbf{g} = [4h_1^{-2}(h_1/2 - y)(h_1/2 + y), 0]^T$ the inflow profile. We consider as parameter the Reynolds number $\mu_1 = \text{Re}$, so that $\nu = 2h_1/(3\mu_1)$.

At very low Reynolds numbers the flow remains symmetric with separation regions of equal length on both channel walls. Increasing the Reynolds number the separation length increases too, and at a critical value $\mu_1 = \mu_1^*$ one recirculation region grows while the other shrinks. This symmetry breaking occurs as the result of a pitchfork bifurcation in the solution of the Navier-Stokes equations [9,11], i.e., for $\mu_1 > \mu_1^*$ two stable (asymmetric) solutions and one unstable (symmetric)

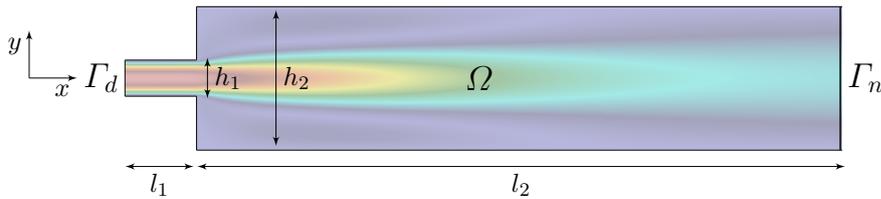


Fig. 13: Sketch of the expanding channel geometry: here we fix $l_1 = 1$, $l_2 = 30$, $h_1 = 0.5$, $h_2 = 2$ (yielding an expansion ratio $h_2/h_1 = 4$). Coloring is given by the velocity field magnitude obtained for $\text{Re} = 60$.

solution coexist (see Figure 14). In correspondence of $\mu_1 = \mu_1^*$ the problem becomes ill-posed, the tangent matrix is singular and therefore the stability factor vanishes⁸.

We initialize the interpolation procedure on a coarse grid of 4 uniformly distributed interpolation points, yielding a very rough approximation of the true stability factor (that is, unaware of the bifurcation, see Figure 14). Then, our adaptive algorithm, selects 16 additional interpolation points in order to reach a maximum budget of 20 points. Among them, the one corresponding to $\mu_1^* \approx 36.25$ is discarded since the solver does not reach convergence (thus indicating proximity to the bifurcation point). Moreover, since we use as initial guess for the Navier-Stokes solver the solution of the corresponding Stokes equations, we obtain convergence to the symmetric solution for the entire range of Reynolds numbers considered. Thus, for $\mu_1 > \mu_1^*$, the computed stability factor is the one corresponding to the unstable branch of solutions.

Once again, we highlight the efficacy of the adaptive criterion in selecting the interpolation points in the most varying region of the parameter space - that is, in the proximity of the bifurcation point. Thus, we obtain a reliable approximation of the stability factor even in this limit-case scenario, still entailing a moderate computational effort. Therefore, this heuristic strategy proves also to be viable when aiming at the detection of bifurcation points.

8 Further remarks and conclusions

In this paper we have developed some heuristic strategies for the rapid and reliable approximation of stability factors related to parametrized, inf-sup stable, (quadratically) nonlinear operators. These strategies feature a relevant novelty, being completely uncoupled to the generation of the reduced approximation space. After providing a suitable theoretical framework, we have developed a linearized version of the Successive Constraint Method. However, the rather expensive off-line stage of this procedure might compromise the efficiency of the whole reduction process. Not only, its convergence shows to be rather slow already when dealing with few ($p = 2, 3$) parameters, as shown by the numerical results presented in this work. For these reasons, we have proposed some alternative heuristic strategies to approximate parametrized stability factors in a more efficient way. They are based

⁸ From the computational standpoint, since the behavior of the unstable flow may be sensitive to the discretization error, we have employed a fine mesh made of 39222 triangular elements, obtained as the outcome of a suitable mesh convergence study.

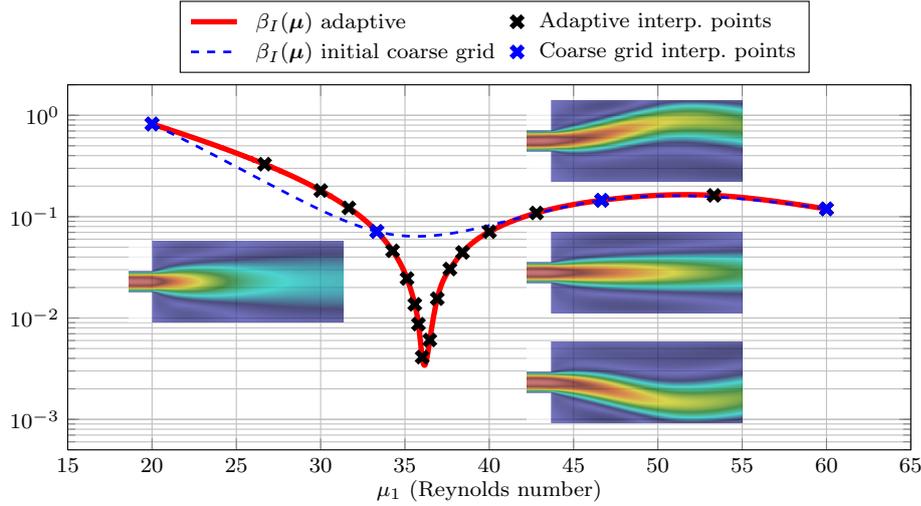


Fig. 14: Channel expansion case: initial interpolation on a coarse grid of 4 points (blue dashed line) and results obtained with the adaptive RBF strategy (red line).

on the construction of suitable radial basis interpolants, where the choice of the interpolation points is performed according to a greedy algorithm.

We have tested the proposed strategies on some numerical cases related to the approximation of 2D Navier-Stokes flows that depend on both physical and geometrical parameters. Our results are quite promising, in terms of both computational efficiency and reliability of the approximated lower bounds. Not only, we have assessed the behavior of the interpolation-based strategy also in the proximity of a bifurcation point. Since the proposed heuristic strategies are valid also in the case of more general nonlinear (i.e., other than quadratic) operators, we expect to exploit them in order to obtain more efficient error estimates for a wider range of problems in the near future.

A Appendix: Proofs of theoretical results

A.1 Proof of Proposition 2

From the definition (6) of $dA(\cdot; \boldsymbol{\mu})(\cdot, \cdot)$ and the affine decomposition (12), we have that

$$\begin{aligned} & dA(u(\boldsymbol{\mu}); \boldsymbol{\mu})(v, w) - dA(u(\boldsymbol{\mu}^*); \boldsymbol{\mu}^*)(v, w) \\ &= \underbrace{\sum_{q=1}^{Q_a} [\Theta_q^a(\boldsymbol{\mu}) - \Theta_q^a(\boldsymbol{\mu}^*)] a_q(v, w)}_{\text{(I)}} + \underbrace{\sum_{q=1}^{Q_c} \Theta_q^c(\boldsymbol{\mu}) d_q(u(\boldsymbol{\mu}), v, w) - \sum_{q=1}^{Q_c} \Theta_q^c(\boldsymbol{\mu}^*) d_q(u(\boldsymbol{\mu}^*), v, w)}_{\text{(II)}}. \end{aligned}$$

The first term can be easily bounded as

$$|\text{(I)}| \leq Q_a L_a \|\boldsymbol{\mu} - \boldsymbol{\mu}^*\| \tilde{\gamma}_a \|v\|_V \|w\|_V, \quad (55)$$

where $L_a = \max_{q=1, \dots, Q_a} L_a^q$, being the L_a^q 's the Lipschitz constants of the functions $\Theta_q^a(\cdot)$, while $\tilde{\gamma}_a = \max_{q=1, \dots, Q_a} \gamma_a^q$, being the γ_a^q 's the continuity constants of the bilinear forms

$a_q(\cdot, \cdot)$. Let us now rewrite the second term as

$$(II) = \sum_{q=1}^{Q_c} [\Theta_q^c(\boldsymbol{\mu}) - \Theta_q^c(\boldsymbol{\mu}^*)] d_q(u(\boldsymbol{\mu}), v, w) + \sum_{q=1}^{Q_c} \Theta_q^c(\boldsymbol{\mu}^*) d_q(u(\boldsymbol{\mu}) - u(\boldsymbol{\mu}^*), v, w),$$

which can be bounded as

$$|(II)| \leq Q_c L_c |\boldsymbol{\mu} - \boldsymbol{\mu}^*| \bar{\gamma}_d \|v\|_V \|w\|_V \|u(\boldsymbol{\mu})\|_V + Q_c M_{\Theta}^c \bar{\gamma}_d \|v\|_V \|w\|_V \|u(\boldsymbol{\mu}) - u(\boldsymbol{\mu}^*)\|_V.$$

Here $L_c = \max_{q=1, \dots, Q_c} L_c^q$, being the L_c^q 's the Lipschitz constants of the functions $\Theta_q^c(\cdot)$, $\bar{\gamma}_d$ is the larger among the continuity constants of the trilinear forms $d_q(\cdot, \cdot, \cdot)$, and

$$M_{\Theta}^c = \max_{\boldsymbol{\mu} \in \mathcal{D}} \max_{q=1, \dots, Q_c} \Theta_q^c(\boldsymbol{\mu}).$$

Since the solution $u(\boldsymbol{\mu})$ of problem (3) is bounded for every $\boldsymbol{\mu} \in \mathcal{D}$ – thanks to (16) – and Lipschitz continuous with respect to $\boldsymbol{\mu}$ (see Proposition 1), there exist positive constants K_u and L_u such that

$$\|u(\boldsymbol{\mu})\|_V \leq K_u, \quad \|u(\boldsymbol{\mu}) - u(\boldsymbol{\mu}^*)\|_V \leq L_u |\boldsymbol{\mu} - \boldsymbol{\mu}^*|, \quad (56)$$

uniformly in \mathcal{D} . Therefore,

$$|(II)| \leq (L_c K_u + M_{\Theta}^c L_u) Q_c \bar{\gamma}_d |\boldsymbol{\mu} - \boldsymbol{\mu}^*| \|v\|_V \|w\|_V. \quad (57)$$

Combining (55) and (57), in the end we obtain (17) with constant

$$C = Q_a L_a \bar{\gamma}_a + Q_c \bar{\gamma}_d (L_c K_u + M_{\Theta}^c L_u).$$

Furthermore, we have

$$\begin{aligned} \|T^{\boldsymbol{\mu}} w - T^{\boldsymbol{\mu}^*} w\|_V^2 &= (T^{\boldsymbol{\mu}} w - T^{\boldsymbol{\mu}^*} w, T^{\boldsymbol{\mu}} w - T^{\boldsymbol{\mu}^*} w)_V \\ &= dA(u(\boldsymbol{\mu}); \boldsymbol{\mu})(w, T^{\boldsymbol{\mu}} w - T^{\boldsymbol{\mu}^*} w) - dA(u(\boldsymbol{\mu}^*); \boldsymbol{\mu}^*)(w, T^{\boldsymbol{\mu}} w - T^{\boldsymbol{\mu}^*} w) \\ &\leq C |\boldsymbol{\mu} - \boldsymbol{\mu}^*| \|w\|_V \|T^{\boldsymbol{\mu}} w - T^{\boldsymbol{\mu}^*} w\|_V \leq \frac{C}{\beta(\boldsymbol{\mu}^*)} |\boldsymbol{\mu} - \boldsymbol{\mu}^*| \|T^{\boldsymbol{\mu}^*} w\|_V \|T^{\boldsymbol{\mu}} w - T^{\boldsymbol{\mu}^*} w\|_V \end{aligned}$$

by exploiting (17) and (11), from which we obtain,

$$\|T^{\boldsymbol{\mu}} w - T^{\boldsymbol{\mu}^*} w\|_V \leq \frac{C}{\beta(\boldsymbol{\mu}^*)} |\boldsymbol{\mu} - \boldsymbol{\mu}^*| \|T^{\boldsymbol{\mu}^*} w\|_V \quad \forall w \in V.$$

A.2 Proof of Proposition 4

In order to show (38), we start by observing that

$$\begin{aligned} (T^{\boldsymbol{\mu}} v, T^{\boldsymbol{\mu}^*} v)_V &= (T^{\boldsymbol{\mu}} v - T^{\boldsymbol{\mu}^*} v + T^{\boldsymbol{\mu}^*} v, T^{\boldsymbol{\mu}^*} v)_V = \|T^{\boldsymbol{\mu}^*} v\|_V^2 + (T^{\boldsymbol{\mu}} v - T^{\boldsymbol{\mu}^*} v, T^{\boldsymbol{\mu}^*} v)_V \\ &= \|T^{\boldsymbol{\mu}^*} v\|_V^2 + dA(u_h(\boldsymbol{\mu}); \boldsymbol{\mu})(v, T^{\boldsymbol{\mu}^*} v) - dA(u_h(\boldsymbol{\mu}^*); \boldsymbol{\mu}^*)(v, T^{\boldsymbol{\mu}^*} v) \end{aligned}$$

so that

$$\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) \leq 1 + \inf_{v \in V_h} \frac{|dA(u_h(\boldsymbol{\mu}); \boldsymbol{\mu})(v, T^{\boldsymbol{\mu}^*} v) - dA(u_h(\boldsymbol{\mu}^*); \boldsymbol{\mu}^*)(v, T^{\boldsymbol{\mu}^*} v)|}{\|T^{\boldsymbol{\mu}^*} v\|_V^2}.$$

In order to bound this quantity, we exploit the result (17) of Proposition 2, which is valid for any $v, w \in V_h$, too (see e.g. [5, Remark 13.2]). Thus, for any $v \in V_h$,

$$\begin{aligned} |dA(v(\boldsymbol{\mu}); \boldsymbol{\mu})(v, T^{\boldsymbol{\mu}^*} v) - dA(v(\boldsymbol{\mu}^*); \boldsymbol{\mu}^*)(v, T^{\boldsymbol{\mu}^*} v)| \\ \leq C |\boldsymbol{\mu} - \boldsymbol{\mu}^*| \|v\|_V \|T^{\boldsymbol{\mu}^*} v\|_V \leq \frac{C}{\beta(\boldsymbol{\mu}^*)} |\boldsymbol{\mu} - \boldsymbol{\mu}^*| \|T^{\boldsymbol{\mu}^*} v\|_V^2, \end{aligned}$$

so that

$$\tilde{\beta}_{\mu^*}(\mu) \leq 1 + \frac{C}{\beta(\mu^*)} |\mu - \mu^*|$$

or, equivalently, $\tilde{\beta}_{\mu^*}(\mu) - 1 = O(|\mu - \mu^*|)$ as $\mu \rightarrow \mu^*$. In order to show (39), we first expand

$$\begin{aligned} \beta_{\mu^*}^2(\mu) &= \inf_{v \in V_h} \frac{\|T^\mu v\|_V^2}{\|T^{\mu^*} v\|_V^2} = \inf_{v \in V_h} \frac{(T^{\mu^*} v + (T^\mu v - T^{\mu^*} v), T^{\mu^*} v + (T^\mu v - T^{\mu^*} v))_V}{\|T^{\mu^*} v\|_V^2} \\ &= 1 + \inf_{v \in V_h} \left(2 \frac{(T^\mu v - T^{\mu^*} v, T^{\mu^*} v)_V}{\|T^{\mu^*} v\|_V^2} + \frac{\|T^\mu v - T^{\mu^*} v\|_V^2}{\|T^{\mu^*} v\|_V^2} \right). \end{aligned}$$

Thanks to (18), we have

$$\frac{\|T^\mu v - T^{\mu^*} v\|_V^2}{\|T^{\mu^*} v\|_V^2} \leq \frac{C^2}{\beta^2(\mu^*)} |\mu - \mu^*|^2, \quad \forall v \in V_h$$

and by recognizing that

$$\inf_{v \in V_h} \frac{(T^\mu v - T^{\mu^*} v, T^{\mu^*} v)_V}{\|T^{\mu^*} v\|_V^2} = \tilde{\beta}_{\mu^*}(\mu) - 1,$$

we end up with $\beta_{\mu^*}^2(\mu) = 1 + 2(\tilde{\beta}_{\mu^*}(\mu) - 1) + O(|\mu - \mu^*|^2)$ as $\mu \rightarrow \mu^*$. By taking the square root, using a Taylor series expansion and the fact that $O(\tilde{\beta}_{\mu^*}(\mu) - 1) = O(|\mu - \mu^*|)$ thanks to (38), we obtain:

$$\begin{aligned} &\sqrt{-1 + 2\tilde{\beta}_{\mu^*}(\mu) + O(|\mu - \mu^*|^2)} = \\ &= 1 + \frac{1}{2}(2\tilde{\beta}_{\mu^*}(\mu) - 2 + O(|\mu - \mu^*|^2)) - \frac{1}{8}(2\tilde{\beta}_{\mu^*}(\mu) - 2 + O(|\mu - \mu^*|^2))^2 + O(|\mu - \mu^*|^3) \\ &= 1 + (\tilde{\beta}_{\mu^*}(\mu) - 1) - \frac{1}{2}(\tilde{\beta}_{\mu^*}(\mu) - 1)^2 + O(|\mu - \mu^*|^2), \end{aligned}$$

so that

$$\beta_{\mu^*}(\mu) = \tilde{\beta}_{\mu^*}(\mu) + O(\tilde{\beta}_{\mu^*}(\mu) - 1)^2 + O(|\mu - \mu^*|^2), \quad \text{as } \mu \rightarrow \mu^*.$$

Finally, by exploiting again (38), we end up with (39). \square

Acknowledgements

We thank Dr. Masayuki Yano (MIT) for several useful discussions and Prof. Anthony T. Patera (MIT) for his initial encouragement in deepening the analysis of the methodologies presented in this paper. We are also grateful to Prof. Alfio Quarteroni (EPFL) for his valuable comments and suggestions. This work has been supported in part by the Swiss National Science Foundation (Project 141034), by the SHARM 2012-2014 SISSA post-doctoral research grant on the project ‘‘Reduced Basis Methods for shape optimization in computational fluid dynamics’’ and by the ‘‘National Group of Computing Science’’ (GNCS-INDAM).

References

1. Beatson, R., Cherrie, J., Mouat, C.: Fast fitting of radial basis functions: Methods based on preconditioned GMRES iteration. *Adv. Comput. Math.* **11**(2-3), 253–270 (1999)
2. Biswas, G., Breuer, M., Durst, F.: Backward-Facing Step Flows for Various Expansion Ratios at Low and Moderate Reynolds Numbers. *J. Fluids Eng.* **126**(3), 362–374 (2004)
3. Brezzi, F., Rappaz, J., Raviart, P.: Finite dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions. *Numer. Math.* **36**, 1–25 (1980)

4. Buhmann, M.: Radial Basis Functions: Theory and Implementations, *Cambridge monographs on applied and computational mathematics*, vol. 12. Cambridge University Press, UK (2003)
5. Caloz, G., Rappaz, J.: Numerical analysis for nonlinear and bifurcation problems. In: P. Ciarlet, J. Lions (eds.) *Handbook of Numerical Analysis*, Vol. V, Techniques of Scientific Computing (Part 2), pp. 487–637. Elsevier Science B.V. (1997)
6. Canuto, C., Tonn, T., Urban, K.: A posteriori error analysis of the reduced basis method for non-affine parameterized nonlinear PDEs. *SIAM J. Numer. Anal.* **47**(3), 2001–2022 (2009)
7. Chen, Y., Hesthaven, J., Maday, Y., Rodriguez, J.: A monotonic evaluation of lower bounds for inf-sup stability constants in the frame of reduced basis approximations. *C. R. Acad. Sci. Paris, Ser. I* **346**, 1295–1300 (2008)
8. Chen, Y., Hesthaven, J., Maday, Y., Rodriguez, J.: Improved successive constraint method based a posteriori error estimate for reduced basis approximation of 2D Maxwell’s problem. *ESAIM Math. Modelling Numer. Anal.* **43**, 1099–1116 (2009)
9. Cliffe, K., Spence, A., Tavener, S.: The numerical analysis of bifurcation problems with application to fluid mechanics. *Acta Numerica* **9**(00), 39–131 (2000)
10. Deparis, S.: Reduced basis error bound computation of parameter-dependent Navier-Stokes equations by the natural norm approach. *SIAM J. Num. Anal.* **46**(4), 2039–2067 (2008)
11. Drikakis, D.: Bifurcation phenomena in incompressible sudden expansion flows. *Physics of Fluids* **9**(1), 76–87 (1997)
12. Fornberg, B., Zuev, J.: The runge phenomenon and spatially variable shape parameters in RBF interpolation. *Comput. Math. Appl.* **54**(3), 379–398 (2007)
13. Girault, V., Raviart, P.A.: *Finite element methods for Navier-Stokes equations: Theory and algorithms*. Springer-Verlag, Berlin and New York (1986)
14. Hendrix, E., G-Tóth, B.: *Introduction to nonlinear and global optimization*. Springer, New York (2010)
15. Hildebrandt, T.H., Graves, L.M.: Implicit functions and their differentials in general analysis. *Trans. Amer. Math. Soc.* **29**(1), 127–153 (1927)
16. Huynh, D., Knezevic, D., Chen, Y., Hesthaven, J., Patera, A.: A natural-norm successive constraint method for inf-sup lower bounds. *Comput. Meth. Appl. Mech. Engrg.* **199**(29–32), 1963–1975 (2010)
17. Huynh, D., Nguyen, N., Patera, A., Rozza, G.: Rapid reliable solution of the parametrized partial differential equations of continuum mechanics and transport (2008-2012). Website: <http://augustine.mit.edu>
18. Huynh, D., Rozza, G., Sen, S., Patera, A.: A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants. *C. R. Acad. Sci. Paris. Sér. I Math.* **345**, 473–478 (2007)
19. Ito, K., Kunisch, K.: *Lagrange multiplier approach to variational problems and applications*. Society for Industrial and Applied Mathematics, Philadelphia, PA (2008)
20. Lassila, T., Manzoni, A., Rozza, G.: On the approximation of stability factors for general parametrized partial differential equations with a two-level affine decomposition. *ESAIM Math. Modelling Numer. Anal.* **46**, 1555–1576 (2012)
21. Mackman, T.J., Allen, C.B.: Investigation of an adaptive sampling method for data interpolation using radial basis functions. *Int. J. Numer. Meth. Eng.* **83**(7), 915–938 (2010)
22. Manzoni, A.: *Reduced models for optimal control, shape optimization and inverse problems in haemodynamics*. Ph.D. thesis, École Polytechnique Fédérale de Lausanne (2012)
23. Manzoni, A.: An efficient computational framework for reduced basis approximation and a posteriori error estimation of parametrized Navier-Stokes flows. *ESAIM Math. Modelling Numer. Anal.* **48**, 1199–1226 (2014)
24. Negri, F., Rozza, G., Manzoni, A., Quarteroni, A.: Reduced basis method for parametrized elliptic optimal control problems. *SIAM J. Sci. Comp.* **35**(5), A2316–A2340 (2013)
25. Quarteroni, A., Rozza, G., Manzoni, A.: Certified reduced basis approximation for parametrized partial differential equations in industrial applications. *J. Math. Ind.* **1**(3) (2011)
26. Quarteroni, A., Valli, A.: *Numerical Approximation of Partial Differential Equations (1st Ed.)*. Springer-Verlag, Berlin-Heidelberg (1994)
27. Rasty, M., Grepl, M.: Efficient reduced basis solution of quadratically nonlinear diffusion equations. In: *Proceedings of 7th Vienna Conference on Mathematical Modelling – MATHMOD 2012* (2012). DOI 10.3182/20120215-3-AT-3016.00125

28. Rozza, G., Huynh, D., Manzoni, A.: Reduced basis approximation and a posteriori error estimation for Stokes flows in parametrized geometries: roles of the inf-sup stability constants. *Numer. Math.* **125**(1), 115–152 (2013)
29. Rozza, G., Huynh, D., Patera, A.: Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Arch. Comput. Methods Engrg.* **15**, 229–275 (2008)
30. Sen, S., Veroy, K., Huynh, D., Deparis, S., Nguyen, N., Patera, A.: “Natural norm” a posteriori error estimators for reduced basis approximations. *J. Comp. Phys.* **217**(1), 37–62 (2006)
31. Temam, R.: *Navier-Stokes Equations*. AMS Chelsea, Providence, Rhode Island (2001)
32. Veroy, K., Patera, A.: Certified real-time solution of the parametrized steady incompressible Navier-Stokes equations: rigorous reduced-basis a posteriori error bounds. *Int. J. Numer. Meth. Fluids* **47**, 773–788 (2005)
33. Yano, M.: A space-time Petrov-Galerkin certified reduced basis method: application to the Boussinesq equations. *SIAM J. Sci. Comput.* **36**(1), A232–A266 (2014)
34. Yano, M., Patera, A.T.: A space-time variational approach to hydrodynamic stability theory. *Proc. R. Soc. A* **469**(2155), 20130036 (2013)
35. Zeidler, E.: *Nonlinear Functional Analysis and its Applications, vol. I: Fixed-Point Theorems*. Springer-Verlag (1985)

Recent publications:

MATHEMATICS INSTITUTE OF COMPUTATIONAL SCIENCE AND ENGINEERING
Section of Mathematics
Ecole Polytechnique Fédérale
CH-1015 Lausanne

- 40.2013** D. KRESSNER, M. STEINLECHNER, A. USCHMAJEW:
Low-rank tensor methods with subspace correction for symmetric eigenvalue problems
- 41.2013** A. ABDULLE, O. BUDÁČ:
An adaptive finite element heterogeneous multiscale method for Stokes flow in porous media
- 42.2013** C. JÄGGLI, L. IAPICHINO, G. ROZZA:
An improvement on geometrical parametrizations by transfinite maps
- 43.2013** P. PACCIARINI, G. ROZZA:
Stabilized reduced basis method for parametrized advection-diffusion PDEs
- 44.2013** A. KOSHAKJI, A. QUARTERONI, G. ROZZA:
Free form deformation techniques applied to 3D shape optimization problems
- 45.2013** J. E. CATRILLON-CANDAS, F. NOBILE, R. F. TEMPONE:
Analytic regularity and collocation approximation for PDEs with random domain deformations
- 01.2014** GIOVANNI MIGLIORATI:
Multivariate Markov-type and Nicolskii-type inequalities for polynomials associated with downward closed multi-index sets
- 02.2014** FEDERICO NEGRI, ANDREA MANZONI, GIANLUIGI ROZZA:
Certified reduced basis method for parametrized optimal control problems governed by the Stokes equations
- 03.2014** CEDRIC EFFENBERGER, DANIEL KRESSNER:
On the residual inverse iteration for nonlinear eigenvalue problems admitting a Rayleigh functional
- 04.2014** TAKAHITO KASHIWABARA, CLAUDIA M. COLCIAGO, LUCA DEDÈ, ALFIO QUARTERONI:
Numerical Well-posedness, regularity, and convergence analysis of the finite element approximation of a generalized Robin boundary value problem
- 05.2014** BJÖRN ADLERBORN, BO KAGSTRÖM, DANIEL KRESSNER:
A parallel QZ algorithm for distributed memory HPC systems
- 06.2014** MICHELE BENZI, SIMONE DEPARIS, GWENOL GRANDPERRIN, ALFIO QUARTERONI:
Parameter estimates for the relaxed dimensional factorization preconditioner and application to hemodynamics
- 07.2014** ASSYR ABDULLE, YUN BAI:
Reduced order modelling numerical homogenization
- 08.2014 NEW :** ANDREA MANZONI, FEDERICO NEGRI:
Heuristic strategies for the approximation of stability factors in quadratically nonlinear parametrized PDEs