

A non-intrusive reduced basis method using artificial neural networks

J.S. Hesthaven, S. Ubbiali

École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

Abstract

In this work, we develop a non-intrusive reduced basis (RB) method for parametrized time-independent partial differential equations (PDEs). The proposed method extracts a reduced basis from a collection of high-fidelity solutions via proper orthogonal decomposition (POD) and employs artificial neural networks (ANNs), particularly multi-layer perceptrons (MLPs), to accurately approximate the coefficients of the reduced model. The search for the optimal number of inner neurons and the minimum amount of training samples which avoids overfitting is carried out in the offline phase through an automatic routine, relying upon a joint use of the Latin Hypercube Sampling (LHS) and the Levenberg-Marquardt training algorithm. This guarantees a complete offline-online decoupling, leading to an efficient RB method - referred to as POD-NN - suitable also for nonlinear problems featuring a non-affine parametric dependence. Numerical studies are presented for the linear and nonlinear Poisson equation and for driven cavity viscous flows, modeled through the steady incompressible Navier-Stokes equations. Both physical and geometrical parametrizations are considered. Several results confirm the accuracy of the POD-NN method and show the substantial speed-up enabled at the online stage with respect to a traditional RB strategy based on the Galerkin projection process.

Keywords: non-intrusive reduced basis method, proper orthogonal decomposition, multi-layer perceptron, Levenberg-Marquardt algorithm, Poisson equation, driven cavity flow

1. Introduction

Several applications arising in engineering and applied sciences involve mathematical models expressed as parametrized partial differential equations (PDEs), in which boundary conditions, material properties, source terms, loads or geometrical factors of the underlying physical problem are addressed to a parameter μ [18, 29, 33]. A list of notable examples includes parameter estimation [6], topology optimization [5], optimal control [39] and uncertainty quantification [38]. In these contexts, one is typically interested in a real-time evaluation of an *output of interest* (defined as a functional of the state variable [16]) for many parameter entries, i.e., for many configurations of the problem.

The continuously growing availability of computational power and the simultaneous algorithmic improvements make possible nowadays the *high-fidelity* numerical resolution of complex problems via standard discretization procedures, such as finite difference (FD), finite volume (FV), finite element (FE), or spectral methods [52]. However, these schemes remain prohibitively expensive in many-query and real-time contexts, both in terms of CPU time and memory demand. This follows from the large amount of degrees of freedom (DOFs) they imply, resulting from the (fine) spatial discretization needed to accurately solve the underpinning PDE [1]. In light of this, *reduced order modeling* (ROM) methods have received a significant attention in the last decades. The objective of these methods is to replace the full-order system by one of significant smaller dimension, thus to decrease the computational burden while leading to a reasonable loss of accuracy [12].

Reduced basis (RB) methods constitute a well-known and widely-used instance of reduced order modeling techniques. They are generally implemented pursuing an offline-online paradigm [40]. Based upon an ensemble of *snapshots* (i.e., high-fidelity solutions to the parametrized differential problem), the goal of the *offline* step is to construct a solution-dependent basis, yielding a reduced space of globally approximating functions capable of representing the main dynamics of the full-order model [2, 12]. For this, two major approaches have been proposed in the literature: proper orthogonal

Email addresses: jan.hesthaven@epfl.ch (J.S. Hesthaven), stefano.ubbioli@epfl.ch (S. Ubbiali)

decomposition (POD) [58] and greedy algorithms [30]. The former relies on a deterministic or random sample to generate the snapshots and then employs a singular value decomposition (SVD) to recover the reduced basis. Whereas, the latter adopt a different strategy, as the basis vectors coincide with the snapshots themselves, carefully selected according to some optimality criterion. As a result, a greedy strategy is typically more effective and efficient than POD, as it enables the exploration of a wider region of the parameter space while entailing the computation of many fewer high-fidelity solutions [29]. However, there may exist problems for which a greedy approach is not feasible, simply because a natural criterion for the choice of the snapshots is not available [2].

When a reduced-order environment has been properly set up, given a new parameter input, an approximation to the *truth* solution is sought *online* as a linear combination of the RB functions. The expansion coefficients are determined via projection of the full-order system onto the reduced space [7]. To this end, a Galerkin procedure is the most popular choice.

Despite their established effectiveness, for complex nonlinear problems with a non-affine dependence on the parameters, projection-based RB methods do not provide any computational gain with respect to a direct (expensive) approach, as the cost to compute the projection coefficients depends on the dimension of the full-order model. In fact, a full decoupling between the online stage and the high-fidelity scheme is the ultimate secret for the success of any RB procedure [52]. For this purpose, one may recover an affine expansion of the differential operator through the empirical interpolation method (EIM) [3] or its discrete variants [10, 48].

A valuable alternative to address this concern is represented by *non-intrusive* RB methods, which refer to the high-fidelity model solely to generate the snapshots, without involving it in the projection process [12]. The projection coefficients are then obtained via interpolation over the parameter domain of a database of reduced-order informations [9]. However, it should be noted that reduced bases generally belong to nonlinear, matrix manifolds. Hence, unless using a large amount of samples, standard interpolation techniques may fail, as they cannot enforce the constraints characterizing those manifolds [1, 4].

In this work, we develop a non-intrusive RB method employing POD for the generation of the reduced basis and resorting to (artificial) neural networks, in particular multi-layer perceptrons, in the interpolation step. Hence, in the following we refer to the proposed RB procedure as the POD-NN method. Being of non-intrusive nature, POD-NN is suitable for a fast and reliable resolution of complex nonlinear PDEs featuring a non-affine parametric dependence. To test its accuracy and efficiency, the POD-NN method is applied to the linear and nonlinear Poisson equation and to the steady incompressible Navier-Stokes equations. Both physical and geometrical parametrizations are considered.

As their biological counterpart, *artificial neural networks* (ANNs) [25] are computing systems consisting of information-processing units, called *neurons*, interconnected through *weighted synapses*. The attractive feature of ANNs lies in their capability of *learning* from experience [37]. As a result, ANNs have found most used in those applications, such as cluster detection, image identification or speech recognition, which are difficult to address via computer algorithms using rule-based programming [42]. However, ANNs have been successfully applied also in more traditional contexts, e.g., continuous function approximation, as a valid alternative to the programming approach to investigate complexity, nonlinearity and uncertainties of a high order [47].

The learning of an ANN is accomplished through a *training* process, during which the network is exposed to a collection of examples and its weights are properly adjusted so that it can provide reasonable responses in similar (yet not identical) situations. As shown later in the paper, the training and the subsequent evaluation of a neural network perfectly fit the offline-online framework, offering the POD-NN method a significant online speed-up with respect to the standard projection-based POD-Galerkin (POD-G) RB procedure.

TODO: paper outline

2. Parametrized partial differential equations

Assume $\mathcal{P}_{ph} \subset \mathbb{R}^{P_{ph}}$ and $\mathcal{P}_g \subset \mathbb{R}^{P_g}$ be compact sets, and let $\mu_{ph} \in \mathcal{P}_{ph}$ and $\mu_g \in \mathcal{P}_g$ be respectively the *physical* and *geometrical* parameters characterizing the differential problem, so that $\mu = (\mu_{ph}, \mu_g) \in \mathcal{P} = \mathcal{P}_{ph} \times \mathcal{P}_g \subset \mathbb{R}^P$, $P = P_{ph} + P_g$ represents the overall *input vector parameter*. While μ_{ph} addresses material properties, source terms and boundary conditions, μ_g defines the shape of the computational domain $\tilde{\Omega} = \tilde{\Omega}(\mu_g) \subset \mathbb{R}^d$, $d = 1, 2$, with Lipschitz boundary $\tilde{\Gamma} = \partial\tilde{\Omega}$; we denote by $\tilde{\Gamma}_D$ and $\tilde{\Gamma}_N$ the portions of $\tilde{\Gamma}$ where Dirichlet and Neumann boundary conditions are enforced, with $\tilde{\Gamma}_D \cup \tilde{\Gamma}_N = \tilde{\Gamma}$ and $\tilde{\Gamma}_D \cap \tilde{\Gamma}_N = \emptyset$.

Consider then a Hilbert space $\tilde{V} = \tilde{V}(\mu_g) = \tilde{V}(\tilde{\Omega}(\mu_g))$ defined over the domain $\tilde{\Omega}(\mu_g)$, equipped with the scalar product $(\cdot, \cdot)_{\tilde{V}}$ and the induced norm $\|\cdot\|_{\tilde{V}} = \sqrt{(\cdot, \cdot)_{\tilde{V}}}$. Furthermore, let $\tilde{V}' = \tilde{V}'(\mu_g)$ be the dual of \tilde{V} , i.e., the space of linear and

continuous functionals over \tilde{V} . Denoting by $\tilde{G} : \tilde{V} \times \mathcal{P}_{ph} \rightarrow \tilde{V}'$ the map representing a parametrized nonlinear second-order PDE, the differential (strong) form of the problem of interest reads: given $\boldsymbol{\mu} = (\boldsymbol{\mu}_{ph}, \boldsymbol{\mu}_g) \in \mathcal{P}$, find $\tilde{u}(\boldsymbol{\mu}) \in \tilde{V}(\boldsymbol{\mu}_g)$ such that

$$\tilde{G}(\tilde{u}(\boldsymbol{\mu}); \boldsymbol{\mu}_{ph}) = 0 \quad \text{in } \tilde{V}'(\boldsymbol{\mu}_g), \quad (1)$$

namely

$$\langle \tilde{G}(\tilde{u}(\boldsymbol{\mu}); \boldsymbol{\mu}_{ph}), v \rangle_{\tilde{V}', \tilde{V}} = 0 \quad \forall v \in \tilde{V}(\boldsymbol{\mu}_g).$$

Here, $\langle \cdot, \cdot \rangle_{\tilde{V}', \tilde{V}} : \tilde{V}' \times \tilde{V} \rightarrow \mathbb{R}$ represents the duality pairing between \tilde{V}' and \tilde{V} , encoding the action of any functional of \tilde{V}' onto elements of \tilde{V} .

The finite element method requires problem (1) to be stated in a weak (or variational) form [51]. To this end, let us introduce the form $\tilde{g} : \tilde{V} \times \tilde{V} \times \mathcal{P} \rightarrow \mathbb{R}$, with $g(\cdot, \cdot; \boldsymbol{\mu})$ defined as:

$$\tilde{g}(w, v; \boldsymbol{\mu}) = \langle \tilde{G}(w; \boldsymbol{\mu}_{ph}); v \rangle_{\tilde{V}', \tilde{V}} \quad \forall w, v \in \tilde{V}.$$

The variational formulation of (1) then reads: given $\boldsymbol{\mu} = (\boldsymbol{\mu}_{ph}, \boldsymbol{\mu}_g) \in \mathcal{P}$, find $\tilde{u}(\boldsymbol{\mu}) \in \tilde{V}(\boldsymbol{\mu}_g)$ such that

$$\tilde{g}(\tilde{u}(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = 0 \quad \forall v \in \tilde{V}(\boldsymbol{\mu}). \quad (2)$$

Note that the definition of $\tilde{g}(\cdot, \cdot; \boldsymbol{\mu})$ relies on the duality pairing $\langle \cdot, \cdot \rangle_{\tilde{V}', \tilde{V}}$ between \tilde{V} and \tilde{V}' . Hence, from the nonlinearity of \tilde{G} follows the nonlinearity of \tilde{g} with respect to its first argument.

2.1. From physical to reference domain

As anticipated in the Introduction, any reduced basis method seeks an approximated solution to a differential problem as a combination of (few) well-chosen basis vectors, resulting in a finite-dimensional model which features a remarkably decreased dimension with respect to canonical discretization techniques (e.g., finite element method). As further detailed in Section ??, the method relies on the combination of a collection of high-fidelity approximations $\{\tilde{u}_h(\boldsymbol{\mu}^{(1)}), \dots, \tilde{u}_h(\boldsymbol{\mu}^{(N)})\}$, called *snapshots*, corresponding to the parameter values $\{\boldsymbol{\mu}^{(1)}, \dots, \boldsymbol{\mu}^{(N)}\}$. However, we may be concerned with boundary value problems for stationary PDEs defined on variable shape geometries, so that two snapshots $\tilde{u}_h(\boldsymbol{\mu}^{(j)})$ and $\tilde{u}_h(\boldsymbol{\mu}^{(k)})$, with $1 \leq j, k \leq N$ and $j \neq k$, are likely to have been computed on two different domains $\tilde{\Omega}(\boldsymbol{\mu}_g^{(j)})$ and $\tilde{\Omega}(\boldsymbol{\mu}_g^{(k)})$. Hence, the underlying high-fidelity solver should be carefully designed to guarantee the *compatibility* among snapshots. In particular, let $\mathbf{u}_h(\boldsymbol{\mu}^{(j)})$ and $\mathbf{u}_h(\boldsymbol{\mu}^{(k)})$ be the vectors collecting the degrees of freedom for $\tilde{u}_h(\boldsymbol{\mu}^{(j)})$ and $\tilde{u}_h(\boldsymbol{\mu}^{(k)})$, respectively. Then, we should ensure that:

- (a) $\dim(\mathbf{u}_h(\boldsymbol{\mu}^{(j)})) = \dim(\mathbf{u}_h(\boldsymbol{\mu}^{(k)}))$, i.e. the number of degrees of freedom must be the same;
- (b) corresponding entries of the two vectors must be correlated.

For a mesh-based numerical method, the conditions (a) and (b) can be satisfied by preserving the connectivity of the underlying meshes across different domains, i.e., different values of $\boldsymbol{\mu}_g$. To this end, we formulate and solve the differential problem over a fixed, *parameter-independent* domain Ω . This can be accomplished upon introducing a parametrized map $\Phi : \Omega \times \mathcal{P}_g \rightarrow \tilde{\Omega}$ such that

$$\tilde{\Omega}(\boldsymbol{\mu}_g) = \Phi(\Omega; \boldsymbol{\mu}_g).$$

The transformation Φ allows to restate the general problem (1). Let V be a suitable Hilbert space over Ω and V' be its dual. Suppose V is equipped with the scalar product $(\cdot, \cdot)_V : V \times V \rightarrow \mathbb{R}$ and the induced norm $\|\cdot\|_V = \sqrt{(\cdot, \cdot)_V} : V \rightarrow [0, \infty)$. Given the parametrized map $G : V \times \mathcal{P} \rightarrow V'$ representing the (nonlinear) PDE over the reference domain Ω , we focus on differential problems of the form: given $\boldsymbol{\mu} \in \mathcal{P}$, find $u(\boldsymbol{\mu}) \in V$ such that

$$G(u(\boldsymbol{\mu}); \boldsymbol{\mu}) = 0 \quad \text{in } V'. \quad (3)$$

The weak formulation of problem (3) reads: given $\boldsymbol{\mu} \in \mathcal{P}$, seek $u(\boldsymbol{\mu}) \in V$ such that

$$g(u(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = 0 \quad \forall v \in V, \quad (4)$$

where $g : V \times V \times \mathcal{P} \rightarrow \mathbb{R}$ is defined as

$$g(w, v; \boldsymbol{\mu}) = \langle G(w); \boldsymbol{\mu} \rangle_{V', V} \quad \forall w, v \in V,$$

with $\langle \cdot, \cdot \rangle_{V', V} : V' \times V \rightarrow \mathbb{R}$ the dual pairing between V and V' . Observe that the explicit expression of $g(\cdot, \cdot; \boldsymbol{\mu})$ involves the map Φ , thus keeping track of the original domain $\tilde{\Omega}(\boldsymbol{\mu}_g)$. Then, the solution $\tilde{u}(\boldsymbol{\mu})$ over the original domain $\tilde{\Omega}(\boldsymbol{\mu}_g)$ can be recovered as

$$\tilde{u}(\boldsymbol{\mu}) = u(\boldsymbol{\mu}) \circ \Phi(\boldsymbol{\mu}).$$

For any $\boldsymbol{\mu} \in \mathcal{P}$ we seek a discrete solution $u_h(\boldsymbol{\mu})$ to the problem (3) on a *parameter-independent* cover Ω_h of the domain Ω . Provided a convenient choice for Ω , this makes the mesh generation process easier. In addition, we note that discretizing the problem (3) over Ω_h is equivalent to approximating the original problem (1) over the mesh $\tilde{\Omega}_h(\boldsymbol{\mu}_g)$, given by

$$\tilde{\Omega}_h(\boldsymbol{\mu}_g) = \Phi(\Omega_h; \boldsymbol{\mu}_g).$$

Therefore, the requirements (a) and (b) are automatically fulfilled provided the map $\Phi(\cdot; \boldsymbol{\mu}_g)$ is *conformal* for any $\boldsymbol{\mu}_g \in \mathcal{P}_g$. To ensure conformality, in our numerical tests we resort to a particular choice for Φ - the boundary displacement-dependent transfinite map (BDD TM) proposed in [33].

2.2. Problems of interest

Within the wide range of PDEs which suit the functional framework portrayed so far, in this work we focus on two relevant examples - the nonlinear Poisson equation and the stationary Navier-Stokes equations - which will serve as test cases in Section ??.

2.2.1. Nonlinear Poisson equation

Despite a rather simple form, the Poisson equation has proved itself to be effective to model steady phenomena occurring in, e.g., electromagnetism, heat transfer, and underground flows [43]. We consider the following version of the parametrized Poisson equation for a state variable $\tilde{u} = \tilde{u}(\boldsymbol{\mu})$:

$$\begin{cases} -\tilde{\nabla} \cdot (\tilde{k}(\tilde{\mathbf{x}}, \tilde{u}(\boldsymbol{\mu}); \boldsymbol{\mu}_{ph}) \tilde{\nabla} \tilde{u}(\boldsymbol{\mu})) = \tilde{s}(\tilde{\mathbf{x}}; \boldsymbol{\mu}_{ph}) & \text{in } \tilde{\Omega}(\boldsymbol{\mu}_g), \\ \tilde{u}(\boldsymbol{\mu}) = \tilde{g}(\tilde{\boldsymbol{\sigma}}; \boldsymbol{\mu}_{ph}) & \text{on } \tilde{\Gamma}_D, \\ \tilde{k}(\tilde{\boldsymbol{\sigma}}, \tilde{u}(\boldsymbol{\mu}); \boldsymbol{\mu}_{ph}) \tilde{\nabla} \tilde{u}(\boldsymbol{\mu}) \cdot \tilde{\mathbf{n}} = 0 & \text{on } \tilde{\Gamma}_N. \end{cases} \quad \begin{matrix} (5a) \\ (5b) \\ (5c) \end{matrix}$$

Here, for any $\boldsymbol{\mu}_g \in \mathcal{P}_g$:

- $\tilde{\mathbf{x}}$ and $\tilde{\boldsymbol{\sigma}}$ denote a generic point in $\tilde{\Omega}$ and on $\tilde{\Gamma}$, respectively;
- $\tilde{\nabla}$ is the nabla operator with respect to $\tilde{\mathbf{x}}$;
- $\tilde{\mathbf{n}} = \tilde{\mathbf{n}}(\tilde{\boldsymbol{\sigma}})$ denotes the outward normal to $\tilde{\Gamma}$ in $\tilde{\boldsymbol{\sigma}}$;
- $\tilde{k} : \tilde{\Omega} \times \mathbb{R} \times \mathcal{P}_{ph} \rightarrow (0, \infty)$ is the diffusion coefficient, $\tilde{s} : \tilde{\Omega} \times \mathcal{P}_{ph} \rightarrow \mathbb{R}$ is the source term, and $\tilde{g} : \tilde{\Gamma}_D \times \mathcal{P}_{ph} \rightarrow \mathbb{R}$ encodes the Dirichlet boundary conditions;
- to ease the subsequent discussion, we limit the attention to homogeneous Neumann boundary constraints.

Let us fix $\boldsymbol{\mu} \in \mathcal{P}$ and set

$$\tilde{V} = H_{\tilde{\Gamma}_D}^1(\tilde{\Omega}) = \{v \in H^1(\tilde{\Omega}) : v|_{\tilde{\Gamma}_D} = 0\},$$

Multiplying (5a) by a *test* function $v \in \tilde{V}$, integrating over $\tilde{\Omega}$, and exploiting integration by parts on the left-hand side, yields:

$$\int_{\tilde{\Omega}(\boldsymbol{\mu}_g)} \tilde{k}(\tilde{u}(\boldsymbol{\mu}); \boldsymbol{\mu}_{ph}) \tilde{\nabla} \tilde{u}(\boldsymbol{\mu}) \cdot \tilde{\nabla} v \, d\tilde{\Omega}(\boldsymbol{\mu}_g) = \int_{\tilde{\Omega}(\boldsymbol{\mu}_g)} \tilde{s}(\boldsymbol{\mu}_{ph}) v \, d\tilde{\Omega}(\boldsymbol{\mu}_g), \quad (6)$$

where we have omitted the dependence on the space variable $\tilde{\mathbf{x}}$ for ease of notation. For the integrals in Eq. (6) to be well-defined, we require, for any $\boldsymbol{\mu}_g \in \mathcal{P}_g$,

$$|\tilde{k}(\tilde{\mathbf{x}}, r; \boldsymbol{\mu}_g)| < \infty \text{ for almost any (a.a.) } \tilde{\mathbf{x}} \in \tilde{\Omega}(\boldsymbol{\mu}_g), r \in \mathbb{R} \quad \text{and} \quad \tilde{s}(\boldsymbol{\mu}_{ph}) \in L^2(\tilde{\Omega}(\boldsymbol{\mu}_g)).$$

Let then $\tilde{l} = \tilde{l}(\boldsymbol{\mu}) \in H^1(\tilde{\Omega}(\boldsymbol{\mu}_g))$ be a *lifting* function such that $\tilde{l}(\boldsymbol{\mu})|_{\tilde{\Gamma}_D} = \tilde{g}(\boldsymbol{\mu}_{ph})$, with $\tilde{g}(\boldsymbol{\mu}_{ph}) \in H^{1/2}(\tilde{\Gamma}_D)$. We assume that such a function can be constructed, e.g., by interpolation of the boundary condition. Hence, upon defining

$$\begin{aligned} \tilde{a}(w, v; \boldsymbol{\mu}) &:= \int_{\tilde{\Omega}(\boldsymbol{\mu}_g)} \tilde{k}(w + \tilde{l}(\boldsymbol{\mu}); \boldsymbol{\mu}_{ph}) \tilde{\nabla} w \cdot \tilde{\nabla} v \, d\tilde{\Omega}(\boldsymbol{\mu}_g) \\ &\quad + \int_{\tilde{\Omega}(\boldsymbol{\mu}_g)} \tilde{k}(w + \tilde{l}(\boldsymbol{\mu}); \boldsymbol{\mu}_{ph}) \tilde{\nabla} \tilde{l}(\boldsymbol{\mu}) \cdot \tilde{\nabla} v \, d\tilde{\Omega}(\boldsymbol{\mu}_g) \quad \forall w, v \in \tilde{V}(\boldsymbol{\mu}_g), \end{aligned} \quad (7a)$$

$$\tilde{f}(v; \boldsymbol{\mu}) := \int_{\tilde{\Omega}(\boldsymbol{\mu}_g)} \tilde{s}(\boldsymbol{\mu}_{ph}) v \, d\tilde{\Omega}(\boldsymbol{\mu}_g) \quad \forall v \in \tilde{V}(\boldsymbol{\mu}_g), \quad (7b)$$

the weak formulation of problem (5) reads: given $\boldsymbol{\mu} \in \mathcal{P}$, find $\tilde{u}(\boldsymbol{\mu}) \in \tilde{V}(\boldsymbol{\mu}_g)$ such that

$$\tilde{a}(\tilde{u}(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = \tilde{f}(v; \boldsymbol{\mu}) \quad \forall v \in \tilde{V}(\boldsymbol{\mu}_g), \quad (8)$$

Then, the weak solution of problem (5) is given by $\tilde{u}(\boldsymbol{\mu}) + \tilde{l}(\boldsymbol{\mu})$. Note that using a lifting function makes the formulation (8) *symmetric*, i.e., both the solution and the test functions are picked up from the same functional space [51].

Let us now re-state the variational problem (8) onto the reference domain Ω . For this purpose, let Γ_D and Γ_N be the portions of the boundary $\Gamma = \partial\Omega$ on which we impose Dirichlet and Neumann boundary conditions, respectively. Moreover, we denote by $\mathbb{J}_\Phi(\boldsymbol{\mu})$ the Jacobian of the parametrized map $\Phi(\cdot; \boldsymbol{\mu})$, with determinant $|\mathbb{J}_\Phi(\boldsymbol{\mu})|$, and we set

$$k(\mathbf{x}, \cdot; \boldsymbol{\mu}) = \tilde{k}(\Phi(\mathbf{x}; \boldsymbol{\mu}), \cdot; \boldsymbol{\mu}), \quad s(\mathbf{x}; \boldsymbol{\mu}) = \tilde{s}(\Phi(\mathbf{x}; \boldsymbol{\mu}); \boldsymbol{\mu}) \quad \text{and} \quad g(\mathbf{x}; \boldsymbol{\mu}) = \tilde{g}(\Phi(\mathbf{x}; \boldsymbol{\mu}); \boldsymbol{\mu}).$$

Letting

$$V = H_{\Gamma_D}^1(\Omega)$$

and exploiting standard change of variables formulas, the variational formulation of the Poisson problem (5) over Ω reads: given $\boldsymbol{\mu} \in \mathcal{P}$, find $u(\boldsymbol{\mu}) \in V$ such that

$$a(u(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = f(v; \boldsymbol{\mu}) \quad \forall v \in V, \quad (9)$$

with

$$\begin{aligned} a(w, v; \boldsymbol{\mu}) &= \int_{\Omega} k(w + l(\boldsymbol{\mu}); \boldsymbol{\mu}) \mathbb{J}_\Phi^{-T}(\boldsymbol{\mu}) \nabla w \cdot \mathbb{J}_\Phi^{-T}(\boldsymbol{\mu}) \nabla v \, |\mathbb{J}_\Phi(\boldsymbol{\mu})| \, d\Omega \\ &\quad + \int_{\Omega} k(w + l(\boldsymbol{\mu}); \boldsymbol{\mu}) \mathbb{J}_\Phi^{-T}(\boldsymbol{\mu}) \nabla l(\boldsymbol{\mu}) \cdot \mathbb{J}_\Phi^{-T}(\boldsymbol{\mu}) \nabla v \, |\mathbb{J}_\Phi(\boldsymbol{\mu})| \, d\Omega, \end{aligned} \quad (10a)$$

$$f(v; \boldsymbol{\mu}) = \int_{\Omega} s(\boldsymbol{\mu}) v \, |\mathbb{J}_\Phi(\boldsymbol{\mu})| \, d\Omega, \quad (10b)$$

for any $w, v \in V$ and $\boldsymbol{\mu} \in \mathcal{P}$. Note that, as done in (7), we resort to a lifting function $l(\boldsymbol{\mu}) \in H^1(\Omega)$ with $l(\boldsymbol{\mu})|_{\Gamma_D} = g(\boldsymbol{\mu})$, such that the weak solution to problem (5) re-stated over Ω is obtained as $u(\boldsymbol{\mu}) + l(\boldsymbol{\mu})$.

Let us also remark that, when on a parameter-independent configuration, one can avoid to distinguish between physical and geometrical parameters, since even the latter now affect the integrands in (10a) and (10b), and not the domain of integration.

2.2.2. Steady incompressible Navier-Stokes equations

The system of the Navier-Stokes equations model the conservation of mass and momentum for an incompressible Newtonian viscous fluid confined in a region $\tilde{\Omega}(\boldsymbol{\mu}_g) \subset \mathbb{R}^d$, $d = 2, 3$ [53]. Letting $\tilde{\mathbf{v}} = \tilde{\mathbf{v}}(\tilde{\mathbf{x}}; \boldsymbol{\mu})$ and $\tilde{p} = \tilde{p}(\tilde{\mathbf{x}}; \boldsymbol{\mu})$ be the velocity and pressure of the fluid, respectively, the parametrized steady version of the Navier-Stokes equations considered in this

work reads:

$$\left\{ \begin{array}{ll} \tilde{\nabla} \cdot \tilde{\mathbf{v}}(\boldsymbol{\mu}) = 0 & \text{in } \tilde{\Omega}(\boldsymbol{\mu}_g), \\ -\nu(\boldsymbol{\mu}) \tilde{\Delta} \tilde{\mathbf{v}}(\boldsymbol{\mu}) + (\tilde{\mathbf{v}}(\boldsymbol{\mu}) \cdot \tilde{\nabla}) \tilde{\mathbf{v}}(\boldsymbol{\mu}) + \frac{1}{\rho(\boldsymbol{\mu})} \tilde{\nabla} \tilde{p}(\boldsymbol{\mu}) = \mathbf{0} & \text{in } \tilde{\Omega}(\boldsymbol{\mu}_g), \\ \tilde{\mathbf{v}}(\boldsymbol{\mu}) = \tilde{\mathbf{g}}(\boldsymbol{\mu}_{ph}) & \text{on } \tilde{\Gamma}_D(\boldsymbol{\mu}_g), \\ \tilde{p}(\boldsymbol{\mu}) \tilde{\mathbf{n}} - \nu(\boldsymbol{\mu}) \tilde{\nabla} \tilde{\mathbf{v}}(\boldsymbol{\mu}) \cdot \tilde{\mathbf{n}} = \mathbf{0} & \text{on } \tilde{\Gamma}_N(\boldsymbol{\mu}_g). \end{array} \right. \quad \begin{array}{l} (11a) \\ (11b) \\ (11c) \\ (11d) \end{array}$$

Here, $\tilde{\mathbf{g}}(\boldsymbol{\mu}_{ph})$ denotes the velocity field prescribed on $\tilde{\Gamma}_D$, whereas homogeneous Neumann conditions are applied on $\tilde{\Gamma}_N$. Furthermore, $\rho(\boldsymbol{\mu})$ and $\nu(\boldsymbol{\mu})$ represent the uniform density and kinematic viscosity of the fluid, respectively. Note that, despite these quantities encode physical properties, we let them depend on the geometrical parameters as well. Indeed, fluid dynamics can be characterized (and controlled) by means of a wealth of dimensionless quantities, e.g., the Reynolds number, which combine physical properties of the fluid with geometrical features of the domain. Therefore, a numerical study of the sensitivity of the system (11) with respect to $\boldsymbol{\mu}_g$ may be carried out by adapting either $\rho(\boldsymbol{\mu})$ or $\nu(\boldsymbol{\mu})$ as $\boldsymbol{\mu}_g$ varies, so to preserve a dimensionless quantity of interest; we refer the reader to Section ?? for a practical example. For our purpose, it is worth noticing that the nonlinearity of problem (11) lies in the convective term

$$(\tilde{\mathbf{v}}(\boldsymbol{\mu}) \cdot \tilde{\nabla}) \tilde{\mathbf{v}}(\boldsymbol{\mu}),$$

which gives rise to a *quadratic* nonlinearity.

To write the differential system (11) in weak form over a $\boldsymbol{\mu}_g$ -independent configuration Ω , let us introduce the velocity and pressure spaces

$$\tilde{X}(\boldsymbol{\mu}_g) = [H_{\tilde{\Gamma}_D}^1(\tilde{\Omega}(\boldsymbol{\mu}_g))]^d \quad \text{and} \quad \tilde{Q}(\boldsymbol{\mu}_g) = L^2(\tilde{\Omega}(\boldsymbol{\mu}_g)),$$

respectively, with

$$X = [H_{\Gamma_D}^1(\Omega)]^d \quad \text{and} \quad Q = L^2(\Omega)$$

be their respective counterparts over Ω . By multiplying (11) for test functions $(\tilde{\chi}, \tilde{\xi}) \in \tilde{X}(\boldsymbol{\mu}_g) \times \tilde{Q}(\boldsymbol{\mu}_g)$, integrating by parts and then tracing everything back onto Ω by means of the parametrized map $\Phi(\cdot; \boldsymbol{\mu})$, we end up with the following parametrized weak variational problem: given $\boldsymbol{\mu} \in \mathcal{P}$, find $u(\boldsymbol{\mu}) = (\mathbf{v}(\boldsymbol{\mu}), p(\boldsymbol{\mu})) \in V = X \times Q$ so that

$$a(\mathbf{v}(\boldsymbol{\mu}), \boldsymbol{\chi}; \boldsymbol{\mu}) + c(\mathbf{v}(\boldsymbol{\mu}), \mathbf{v}(\boldsymbol{\mu}), \boldsymbol{\chi}; \boldsymbol{\mu}) + d(\mathbf{v}(\boldsymbol{\mu}), \boldsymbol{\chi}; \boldsymbol{\mu}) + b(p(\boldsymbol{\mu}), \nabla \cdot \boldsymbol{\chi}; \boldsymbol{\mu}) = f_1(\boldsymbol{\chi}; \boldsymbol{\mu}), \quad (12a)$$

$$b(\nabla \cdot \mathbf{v}(\boldsymbol{\mu}), \xi; \boldsymbol{\mu}) = f_2(\xi; \boldsymbol{\mu}), \quad (12b)$$

for all $(\boldsymbol{\chi}, \xi) \in V$, with, for any $\boldsymbol{\mu} \in \mathcal{P}$,

$$c(\boldsymbol{\psi}, \boldsymbol{\chi}, \boldsymbol{\eta}; \boldsymbol{\mu}) = \int_{\Omega} (\boldsymbol{\psi} \cdot \mathbb{J}_{\Phi}^{-T}(\boldsymbol{\mu}) \nabla) \boldsymbol{\chi} \cdot \boldsymbol{\eta} |\mathbb{J}_{\Phi}(\boldsymbol{\mu})| d\Omega \quad \forall \boldsymbol{\psi}, \boldsymbol{\chi}, \boldsymbol{\eta} \in [H^1(\Omega)]^d, \quad (13a)$$

$$a(\boldsymbol{\psi}, \boldsymbol{\chi}; \boldsymbol{\mu}) = \nu(\boldsymbol{\mu}) \int_{\Omega} \mathbb{J}_{\Phi}^{-T}(\boldsymbol{\mu}) \nabla \boldsymbol{\psi} : \mathbb{J}_{\Phi}^{-T}(\boldsymbol{\mu}) \nabla \boldsymbol{\chi} |\mathbb{J}_{\Phi}(\boldsymbol{\mu})| d\Omega \quad \forall \boldsymbol{\psi}, \boldsymbol{\chi} \in X, \quad (13b)$$

$$b(\boldsymbol{\psi}, \xi; \boldsymbol{\mu}) = -\frac{1}{\rho(\boldsymbol{\mu})} \int_{\Omega} (\mathbb{J}_{\Phi}^{-T}(\boldsymbol{\mu}) \nabla \cdot \boldsymbol{\psi}) \xi |\mathbb{J}_{\Phi}(\boldsymbol{\mu})| d\Omega \quad \forall \boldsymbol{\psi} \in X, \forall \xi \in Q, \quad (13c)$$

$$d(\boldsymbol{\psi}, \boldsymbol{\chi}; \boldsymbol{\mu}) = c(\mathbf{l}(\boldsymbol{\mu}), \boldsymbol{\psi}, \boldsymbol{\chi}; \boldsymbol{\mu}) + c(\boldsymbol{\psi}, \mathbf{l}(\boldsymbol{\mu}), \boldsymbol{\chi}; \boldsymbol{\mu}); \quad \forall \boldsymbol{\psi}, \boldsymbol{\chi} \in X, \quad (13d)$$

$$f_1(\boldsymbol{\psi}; \boldsymbol{\mu}) = -a(\mathbf{l}(\boldsymbol{\mu}), \boldsymbol{\psi}; \boldsymbol{\mu}) - c(\mathbf{l}(\boldsymbol{\mu}), \mathbf{l}(\boldsymbol{\mu}), \boldsymbol{\psi}; \boldsymbol{\mu}) \quad \forall \boldsymbol{\psi} \in X, \quad (13e)$$

$$f_2(\xi; \boldsymbol{\mu}) = -b(\mathbf{l}(\boldsymbol{\mu}), \xi; \boldsymbol{\mu}) \quad \forall \xi \in Q. \quad (13f)$$

Here, $\mathbf{l}(\boldsymbol{\mu}) \in [H^1(\Omega)]^d$ denotes the lifting vector field, with $\mathbf{l}(\boldsymbol{\mu})|_{\Gamma_D} = \mathbf{g}(\boldsymbol{\mu})$, $\mathbf{g}(\mathbf{x}; \boldsymbol{\mu}) = \tilde{\mathbf{g}}(\Phi(\mathbf{x}; \boldsymbol{\mu}); \boldsymbol{\mu})$ being the velocity field prescribed on Γ_D . Hence, the weak solution to (11) defined over the fixed domain Ω is given by $(\mathbf{v}(\boldsymbol{\mu}) + \mathbf{l}(\boldsymbol{\mu}), p(\boldsymbol{\mu}))$.

3. Discrete full-order model

Let $V_h \subset V$ be a finite-dimensional subspace of V of dimension M . The FE approximation of the weak problem (4) can be cast in the form: given $\boldsymbol{\mu} \in \mathcal{P}$, find $u_h(\boldsymbol{\mu}) \in V_h$ such that

$$g(u_h(\boldsymbol{\mu}), v_h; \boldsymbol{\mu}) = 0 \quad \forall v_h \in V_h. \quad (14)$$

The discretization (14) is known as *Galerkin approximation*, and therefore $u_h(\boldsymbol{\mu})$ is referred to as the *Galerkin solution* to the problem (3).

Due to the nonlinearity of $g(\cdot, \cdot; \boldsymbol{\mu})$ in its first argument, one has to resort to some iterative method, e.g., Newton's method, to solve the Galerkin problem (14). In this regard, let

$$dg[z](\cdot, \cdot; \boldsymbol{\mu}) : V \times V \rightarrow \mathbb{R}$$

be the partial Frechét derivative of $g(\cdot, \cdot; \boldsymbol{\mu})$ with respect to its first argument, evaluated at $z \in V$. Starting from an initial guess $u_h^0(\boldsymbol{\mu})$, we construct a collection of approximations $\{u_h^k(\boldsymbol{\mu})\}_{k \geq 0}$ to the Galerkin solution $u_h(\boldsymbol{\mu})$ by iteratively solving the linearized problems

$$dg[u_h^k(\boldsymbol{\mu})](\delta u_h^k(\boldsymbol{\mu}), v_h; \boldsymbol{\mu}) = -g(u_h^k(\boldsymbol{\mu}), v_h; \boldsymbol{\mu}) \quad \forall v_h \in V_h$$

in the unknown $\delta u_h^k(\boldsymbol{\mu}) \in V_h$, and then setting $u_h^{k+1}(\boldsymbol{\mu}) = u_h^k(\boldsymbol{\mu}) + \delta u_h^k(\boldsymbol{\mu})$.

To derive the algebraic counterpart of the Galerkin-Newton method, let $\{\phi_1, \dots, \phi_M\}$ be a basis for the M -dimensional space V_h , so that the solution $u_h(\boldsymbol{\mu})$ can be expressed as a linear combination of the basis functions, i.e.,

$$u_h(\mathbf{x}; \boldsymbol{\mu}) = \sum_{j=1}^M u_h^{(j)}(\boldsymbol{\mu}) \phi_j(\mathbf{x}). \quad (15)$$

Hence, denoting by $\mathbf{u}_h(\boldsymbol{\mu}) \in \mathbb{R}^M$ the vector collecting the *degrees of freedom* $\{u_h^{(j)}\}_{j=1}^M$ and exploiting the linearity of $g(\cdot, \cdot; \boldsymbol{\mu})$ in the second argument, the problem (14) is equivalent to: given $\boldsymbol{\mu} \in \mathcal{P}$, find $\mathbf{u}_h(\boldsymbol{\mu}) \in \mathbb{R}^M$ such that

$$g\left(\sum_{j=1}^M u_h^{(j)}(\boldsymbol{\mu}) \phi_j, \phi_i; \boldsymbol{\mu}\right) = 0 \quad \forall i = 1, \dots, M.$$

We observe now that the above problem can be written in compact form as

$$\mathbf{G}_h(\mathbf{u}_h(\boldsymbol{\mu}); \boldsymbol{\mu}) = \mathbf{0} \in \mathbb{R}^M, \quad (16)$$

where the i -th component of the *residual vector* $\mathbf{G}(\cdot; \boldsymbol{\mu}) \in \mathbb{R}^M$ is given by

$$(\mathbf{G}_h(\mathbf{u}_h(\boldsymbol{\mu}); \boldsymbol{\mu}))_i = g\left(\sum_{j=1}^M u_h^{(j)}(\boldsymbol{\mu}) \phi_j, \phi_i; \boldsymbol{\mu}\right), \quad i = 1, \dots, M. \quad (17)$$

Then, for $k \geq 0$, the k -th iteration of Newton's method applied to the system (16) entails the resolution of the *linear* system

$$\mathbb{J}_h(\mathbf{u}_h^k(\boldsymbol{\mu}); \boldsymbol{\mu}) \delta \mathbf{u}_h^k(\boldsymbol{\mu}) = -\mathbf{G}_h(\mathbf{u}_h^k(\boldsymbol{\mu}); \boldsymbol{\mu}), \quad \delta \mathbf{u}_h^k(\boldsymbol{\mu}) \in \mathbb{R}^M, \quad (18)$$

so that $\mathbf{u}_h^{k+1}(\boldsymbol{\mu}) = \mathbf{u}_h^k(\boldsymbol{\mu}) + \delta \mathbf{u}_h^k(\boldsymbol{\mu})$. Here, $\mathbb{J}_h(\cdot; \boldsymbol{\mu}) \in \mathbb{R}^{M \times M}$ denotes the Jacobian of the residual vector $\mathbf{G}_h(\cdot; \boldsymbol{\mu})$; exploiting the bilinearity of $dg[z](\cdot, \cdot; \boldsymbol{\mu})$, $\mathbb{J}_h(\cdot; \boldsymbol{\mu})$ is defined as

$$(\mathbb{J}_h(\mathbf{u}_h^k(\boldsymbol{\mu}); \boldsymbol{\mu}))_{i,j} = dg[u_h^k(\boldsymbol{\mu})](\phi_j, \phi_i; \boldsymbol{\mu}), \quad i, j = 1, \dots, M.$$

4. Projection-based reduced basis method

As detailed in the previous section, the finite element discretization of the $\boldsymbol{\mu}$ -dependent nonlinear differential problem (4), combined with Newton's method, entails the assembly and resolution of (possibly) many linear systems of the form (18), whose dimension is directly related to (i) the size of the underlying grid and (ii) the order of the polynomial FE spaces adopted. Since the accuracy of the resulting discretization heavily relies on these two factors, a direct numerical approximation of the full-order model implies severe computational costs. Therefore, even resorting to high-performance parallel workstations, this approach is hardly affordable in *many-query* and *real-time* contexts. This motivates the broad use of *reduced-order* models, across several inter-disciplinary areas, e.g., parameter estimation, optimal control, shape optimization and uncertainty quantification [29, 52].

Reduced basis methods seek an approximated solution to the problem (4) as a linear combination of parameter-independent functions

$$\{\psi_1, \dots, \psi_L\} \subset V_h,$$

called *reduced basis functions*, built from a collection of high-fidelity solutions

$$\{u_h(\boldsymbol{\mu}^{(1)}), \dots, u_h(\boldsymbol{\mu}^{(N)})\} \subset \mathcal{M}_h,$$

called *snapshots*, where the discrete and finite set

$$\Xi_N = \{\boldsymbol{\mu}^{(1)}, \dots, \boldsymbol{\mu}^{(N)}\} \subset \mathcal{P}$$

may consist of either a uniform lattice or randomly generated points over the parameter domain \mathcal{P} [29]. The basis functions $\{\psi_l\}_{l=1}^L$ generally follow from a principal component analysis (PCA) of the set of snapshots (in that case, $N > L$), or they might coincide with the snapshots themselves (in that case, $N = L$). In the latter approach, typical of any *greedy* method, the parameters $\{\boldsymbol{\mu}^{(n)}\}_{n=1}^N$ must be carefully chosen according to some optimality criterium (see, e.g., [12]). Here, we pursue the first approach, employing the well-known Proper Orthogonal Decomposition (POD) method [58], detailed in the following subsection.

Assume now that a reduced basis is available and let $V_{\text{rb}} \subset V_h$ be the associated *reduced basis space*, i.e.,

$$V_{\text{rb}} = \text{span}\{\psi_1, \dots, \psi_L\}.$$

A *reduced basis solution* $u_L(\boldsymbol{\mu})$ is sought in the form

$$u_L(\mathbf{x}; \boldsymbol{\mu}) = \sum_{l=1}^L u_{\text{rb}}^{(l)}(\boldsymbol{\mu}) \psi_l(\mathbf{x}) \in V_{\text{rb}}, \quad (19)$$

with

$$\mathbf{u}_{\text{rb}}(\boldsymbol{\mu}) = [u_{\text{rb}}^{(1)}(\boldsymbol{\mu}), \dots, u_{\text{rb}}^{(L)}(\boldsymbol{\mu})]^T \in \mathbb{R}^L$$

be the *coefficients* (also called *generalized coordinates*) for the expansion of the RB solution in the RB basis functions.

As one can foresee, RB methods (potentially) enable a relevant reduction in the computational effort when the dimension of the associated test and trial space V_{rb} is significantly smaller than the dimension of the original finite element space V_h ; in other terms, $L \ll M$ must hold. However, this assumes that the solution manifold $\mathcal{M} \subset V$ (or, equivalently, $\mathcal{M}_h \subset V_h$) is actually of low-dimension, and can then be accurately approximated by a subspace of reduced dimension L [29]. To further investigate this necessary hypothesis, it is worth introducing the notion of Kolmogorov L -width. The definition is [40]:

Definition 4.1. Let X be a linear space equipped with the norm $\|\cdot\|_X$, A be a subset of X and X_L be a generic L -dimensional subspace of X . The deviation of A from X_L is defined as

$$E(A; X_L) = \sup_{x \in A} \inf_{y \in X_L} \|x - y\|_X.$$

Then, the Kolmogorov L -width of A in X is given by

$$\begin{aligned} d_L(A, X) &= \inf \{E(A; X_L) : X_L \text{ is an } L\text{-dimensional subspace of } X\} \\ &= \inf_{X_L} \sup_{x \in A} \inf_{y \in X_L} \|x - y\|_X. \end{aligned} \quad (20)$$

Therefore, $d_L(A, X)$ measures the extent to which the subset A of the vector space X can be well-approximated by an L -dimensional subspace [40]. Indeed, there exist many situations in which the Kolmogorov L -width shows a graceful behaviour with L , e.g., an exponential decay. In our case, $X = V$ and $A = \mathcal{M}$, and we can refer to regularity of the solutions $u(\boldsymbol{\mu})$ with respect to the parameter $\boldsymbol{\mu}$, or even to analyticity in the parameter dependence [7].

Nevertheless, we still have to ensure that the RB space V_{rb} , i.e., the chosen X_L in (20), attains the infimum $d_L(\mathcal{M}, V)$, or at least a value close to the infimum. At this regard, let us consider the following bound for the error committed when approximating the continuous solution $u(\boldsymbol{\mu})$ with $u_L(\boldsymbol{\mu})$:

$$\|u(\boldsymbol{\mu}) - u_L(\boldsymbol{\mu})\|_V \leq \|u(\boldsymbol{\mu}) - u_h(\boldsymbol{\mu})\|_V + \|u_h(\boldsymbol{\mu}) - u_L(\boldsymbol{\mu})\|_V \quad \forall \boldsymbol{\mu} \in \mathcal{P}.$$

The first term on the right-hand side measures the discrepancy between the continuous solution and its high-fidelity approximation. For what previously said, this error can be lower to any desired level of accuracy. Therefore, the reliability of the solution provided by any reduced basis technique relies on a sound control of the second term of the right-hand side $\|u_h(\boldsymbol{\mu}) - u_L(\boldsymbol{\mu})\|_V$, i.e., the error between the truth and the reduced solution. The last decade has witnessed the development of different *a priori* and *a posteriori* estimates for such error (see, e.g., [7, 29, 40]), thus *certifying* the RB procedure, that is, enabling the user to trust the output of the RB method. However, as already mentioned, the range of application of these estimates is usually limited to linear problems with an affine dependence on the parameters. Although recent and relevant improvements have been achieved also for the Navier-Stokes equations (see, e.g., [16, 52]), they rely on non-trivial results from functional analysis and involve rather long calculations. Hence, it is not the intent of this work to further investigate and employ these estimates. Yet, in our numerical simulations we will *empirically* study the effectiveness of the POD-Galerkin method by evaluating the error $\|u_h(\boldsymbol{\mu}) - u_L(\boldsymbol{\mu})\|_V$ on a finite and discrete *test* dataset $\Xi_{te} \in \mathcal{P}$. We refer the reader to Section ?? for a deeper discussion.

Let us now resume the derivation of the POD-Galerkin RB method. To unearth $u_L(\boldsymbol{\mu})$, whose general form is given in (19), we proceed to project the variational problem (4) onto the RB space V_{rb} by pursuing a standard Galerkin approach, leading to the following *reduced basis problem*: given $\boldsymbol{\mu} \in \mathcal{P}$, find $u_L(\boldsymbol{\mu}) \in V_{\text{rb}}$ so that

$$g(u_L(\boldsymbol{\mu}), v_L; \boldsymbol{\mu}) = 0 \quad \forall v_L \in V_{\text{rb}}. \quad (21)$$

Then, Newton's method applied to (21) entails, at each iteration $k \geq 0$, the solution of the linearized problem: given $\boldsymbol{\mu} \in \mathcal{P}$, seek $\delta u_L^k(\boldsymbol{\mu})$ such that

$$dg[u_L^k(\boldsymbol{\mu})](\delta u_L^k(\boldsymbol{\mu}), v_L; \boldsymbol{\mu}) = -g(u_L^k(\boldsymbol{\mu}), v_L; \boldsymbol{\mu}) \quad \forall v_L \in V_{\text{rb}},$$

with $u_L^{k+1}(\boldsymbol{\mu}) = u_L^k(\boldsymbol{\mu}) + \delta u_L^k(\boldsymbol{\mu})$.

Let us point out that the RB functions $\{\psi_l\}_{l=1}^L$ belong to V_h , i.e., they are actual FE functions. Hence, we denote by $\boldsymbol{\psi}_l \in \mathbb{R}^M$ the vector collecting the nodal values of ψ_l , for $l = 1, \dots, L$, and introduce the matrix $\mathbb{V} = [\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_L] \in \mathbb{R}^{M \times L}$. For any $v_L \in V_{\text{rb}}$, the matrix \mathbb{V} encodes the change of variables from the RB basis to the standard (Lagrangian) FE basis, i.e.,

$$\mathbf{v}_L = \mathbb{V} \mathbf{v}_{\text{rb}}. \quad (22)$$

Therefore, each element v_L of the reduced space admits two (algebraic) representations:

- $\mathbf{v}_{\text{rb}} \in \mathbb{R}^L$, collecting the coefficients for the expansion of v_L in terms of the RB basis $\{\psi_1, \dots, \psi_L\}$;
- $\mathbf{v}_L \in \mathbb{R}^M$, collecting the coefficients for the expansion of v_L in terms of the FE basis $\{\phi_1, \dots, \phi_M\}$.

Note that the latter is also available for any $v_h \in V_h$, while the former characterizes the element in the subspace V_{rb} .

Upon choosing $v_L = \psi_l$, $l = 1, \dots, L$, in the RB problem (21), for any $\boldsymbol{\mu} \in \mathcal{P}$ we get the set of equations

$$g(u_L(\boldsymbol{\mu}), \psi_l; \boldsymbol{\mu}) = 0 \quad 1 \leq l \leq L. \quad (23)$$

Inserting into (23) the expansion of ψ_l , $l = 1, \dots, L$, in terms of the canonical FE basis $\{\phi_m\}_{m=1}^M$, i.e.,

$$\psi_l(\mathbf{x}) = \sum_{m=1}^M \psi_l^{(m)} \phi_m(\mathbf{x}) = \sum_{m=1}^M \mathbb{V}_{m,l} \phi_m(\mathbf{x}),$$

and exploiting the linearity of $g(\cdot, \cdot; \boldsymbol{\mu})$ in the second argument, yields:

$$0 = \sum_{m=1}^M \mathbb{V}_{m,l} g(u_L(\boldsymbol{\mu}), \phi_m; \boldsymbol{\mu}) = (\mathbb{V}^T \mathbf{G}_h(\mathbf{u}_L(\boldsymbol{\mu}); \boldsymbol{\mu}))_l \quad 1 \leq l \leq L,$$

where the last equality follows from the definition (17) of the residual vector $\mathbf{G}_h(\cdot; \boldsymbol{\mu})$ and the notation introduced in (22). Then, the algebraic formulation of the reduced basis problem (21) can be written in compact form as:

$$\mathbf{G}_{\text{rb}}(\mathbf{u}_{\text{rb}}(\boldsymbol{\mu}); \boldsymbol{\mu}) = \mathbb{V}^T \mathbf{G}_h(\mathbf{u}_L(\boldsymbol{\mu}); \boldsymbol{\mu}) = \mathbb{V}^T \mathbf{G}_h(\mathbb{V} \mathbf{u}_{\text{rb}}(\boldsymbol{\mu}); \boldsymbol{\mu}) = \mathbf{0} \in \mathbb{R}^L. \quad (24)$$

This *reduced nonlinear system* imposes the orthogonality (in the Euclidean scalar product) of the residual vector $\mathbf{G}_h(\cdot; \boldsymbol{\mu})$, evaluated in $\mathbb{V} \mathbf{u}_{\text{rb}}(\boldsymbol{\mu})$, to the columns of \mathbb{V} , thus encoding the Galerkin approach pursued at the variational level.

Finally, exploiting the chain rule and the Jacobian $\mathbb{J}_h(\cdot; \boldsymbol{\mu})$ of $\mathbf{G}_h(\cdot; \boldsymbol{\mu})$, the Jacobian $\mathbb{J}_{\text{rb}}(\cdot; \boldsymbol{\mu})$ of $\mathbf{G}_{\text{rb}}(\cdot; \boldsymbol{\mu})$ is given by

$$\mathbb{J}_{\text{rb}}(\mathbf{w}; \boldsymbol{\mu}) = \mathbb{V}^T \mathbb{J}_h(\mathbb{V} \mathbf{w}; \boldsymbol{\mu}) \mathbb{V} \in \mathbb{R}^{L \times L} \quad \forall \mathbf{w} \in \mathbb{R}^L, \forall \boldsymbol{\mu} \in \mathcal{P}. \quad (25)$$

Hence, starting from an initial guess $\mathbf{u}_{\text{rb}}^0 \in \mathbb{R}^L$, each iteration k , $k \geq 0$, of Newton's method applied to the reduced nonlinear system (24) entails the resolution of the linear system

$$\mathbb{J}_{\text{rb}}(\mathbf{u}_{\text{rb}}^k(\boldsymbol{\mu}); \boldsymbol{\mu}) \delta \mathbf{u}_{\text{rb}}^k(\boldsymbol{\mu}) = -\mathbf{G}_{\text{rb}}(\mathbf{u}_{\text{rb}}^k(\boldsymbol{\mu}); \boldsymbol{\mu}), \quad (26)$$

with $\mathbf{u}_{\text{rb}}^{k+1}(\boldsymbol{\mu}) = \mathbf{u}_{\text{rb}}^k(\boldsymbol{\mu}) + \delta \mathbf{u}_{\text{rb}}^k(\boldsymbol{\mu})$. Therefore, applying the POD-Galerkin RB method enables a dramatic reduction of the size of the linear systems to solve whenever the dimension L of the reduced space V_{rb} is much lower than the dimension M of the underlying finite element space V_h .

4.1. Proper Orthogonal Decomposition

Consider a collection of N snapshots $\{u_h(\boldsymbol{\mu}^{(1)}), \dots, u_h(\boldsymbol{\mu}^{(N)})\} \subset \mathcal{M}_h$ corresponding to the finite and discrete parameter set $\Xi_N = \{\boldsymbol{\mu}^{(1)}, \dots, \boldsymbol{\mu}^{(N)}\} \subset \mathcal{P}$, and let \mathcal{M}_{Ξ_N} be the subspace spanned by the snapshots, i.e.,

$$\mathcal{M}_{\Xi_N} = \text{span}\{u_h(\boldsymbol{\mu}^{(1)}), \dots, u_h(\boldsymbol{\mu}^{(N)})\}.$$

Clearly, $\mathcal{M}_{\Xi_N} \subset \mathcal{M}_h$ and we can assume that \mathcal{M}_{Ξ_N} provides a good approximation of \mathcal{M}_h , as long as the number of snapshots is sufficiently large (but typically much smaller than the dimension M of the FE space). Then, we aim at finding a parameter-independent *reduced basis* for \mathcal{M}_{Ξ_N} , i.e., a collection of FE functions $\{\psi_1, \dots, \psi_L\} \subset \mathcal{M}_{\Xi_N}$, with $L \ll M, N$, and L *independent* of both M and N , so that the associated linear space

$$V_{\text{rb}} = \text{span}\{\psi_1, \dots, \psi_L\}$$

constitutes a low-rank approximation of \mathcal{M}_{Ξ_N} , optimal in some later defined sense.

To this end, we work at the algebraic level. Let $\mathbf{u}_h(\boldsymbol{\mu}^{(n)}) \in \mathbb{R}^M$ be the vector collecting the degrees of freedom (with respect to the FE basis) for the n -th snapshots $u_h(\boldsymbol{\mu}^{(n)})$, $n = 1, \dots, N$, and consider the *snapshot matrix* $\mathbb{S} \in \mathbb{R}^{M \times N}$ storing such vectors in a column-wise sense, i.e.,

$$\mathbb{S} = [\mathbf{u}_h(\boldsymbol{\mu}^{(1)}) \mid \dots \mid \mathbf{u}_h(\boldsymbol{\mu}^{(N)})].$$

Denoting by R the rank of \mathbb{S} , with $R \leq \min\{M, N\}$, the singular value decomposition (SVD) of \mathbb{S} ensures the existence of two orthogonal matrices $\mathbb{W} = [\mathbf{w}_1 \mid \dots \mid \mathbf{w}_M] \in \mathbb{R}^{M \times M}$ and $\mathbb{Z} = [\mathbf{z}_1 \mid \dots \mid \mathbf{z}_N] \in \mathbb{R}^{N \times N}$, and a diagonal matrix $\mathbb{D} = \text{diag}(\sigma_1, \dots, \sigma_R) \in \mathbb{R}^{R \times R}$, with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_R > 0$, such that

$$\mathbb{S} = \mathbb{W} \begin{bmatrix} \mathbb{D} & 0 \\ 0 & 0 \end{bmatrix} \mathbb{Z}^T = \mathbb{W} \Sigma \mathbb{Z}^T, \quad (27)$$

where the zeros denote null matrices of appropriate dimensions. The real values $\{\sigma_i\}_{i=1}^R$ are called *singular values* of \mathbb{S} , the columns $\mathbf{w}_m \in \mathbb{R}^M$, $m = 1, \dots, M$, of \mathbb{W} are called *left singular vectors* of \mathbb{S} , and the columns $\mathbf{z}_n \in \mathbb{R}^N$, $n = 1, \dots, N$, of \mathbb{Z} are called *right singular vectors* of \mathbb{S} , and they are related by the following relations:

$$\mathbb{S}\mathbb{S}^T\mathbf{w}_m = \begin{cases} \sigma_m^2 \mathbf{w}_m & \text{for } 1 \leq m \leq R, \\ \mathbf{0} & \text{for } R+1 \leq m \leq M, \end{cases} \quad (28)$$

$$\mathbb{S}^T\mathbb{S}\mathbf{z}_n = \begin{cases} \sigma_n^2 \mathbf{z}_n & \text{for } 1 \leq n \leq R, \\ \mathbf{0} & \text{for } R+1 \leq n \leq N, \end{cases} \quad (29)$$

$$\mathbb{S}\mathbf{z}_i = \sigma_i \mathbf{w}_i \quad \text{for } 1 \leq i \leq R, \quad (30)$$

$$\mathbb{S}^T\mathbf{w}_i = \sigma_i \mathbf{z}_i \quad \text{for } 1 \leq i \leq R. \quad (31)$$

Due to the sparsity pattern of Σ in (27), the SVD of \mathbb{S} can be cast in the compact form:

$$\mathbb{S} = \mathbb{W}_R \mathbb{D} \mathbb{Z}_R^T,$$

with $\mathbb{W}_R \in \mathbb{R}^{M \times R}$ and $\mathbb{Z}_R \in \mathbb{R}^{N \times R}$ retaining only the first R columns of \mathbb{W} and \mathbb{Z} , respectively, i.e.,

$$\mathbb{W}_R = [\mathbf{w}_1 \mid \dots \mid \mathbf{w}_R] \quad \text{and} \quad \mathbb{Z}_R = [\mathbf{z}_1 \mid \dots \mid \mathbf{z}_R]. \quad (32)$$

Letting $\mathbb{B}_R = \mathbb{D} \mathbb{Z}_R^T \in \mathbb{R}^{R \times N}$ and exploiting the orthonormality of the columns of \mathbb{W}_R , the generic column $\mathbf{s}_n = \mathbf{u}_h(\boldsymbol{\mu}^{(n)})$ of \mathbb{S} , $n = 1, \dots, N$, can be expressed as [58]:

$$\begin{aligned} \mathbf{s}_n &= \sum_{r=1}^R (\mathbb{B}_R)_{r,n} \mathbf{w}_r = \sum_{r=1}^R (\mathbb{D} \mathbb{Z}_R^T)_{r,n} \mathbf{w}_r = \sum_{r=1}^R \underbrace{(\mathbb{W}_R^T \mathbb{W}_R \mathbb{D} \mathbb{Z}_R^T)_{r,n}}_{\mathbb{I}_R \in \mathbb{R}^{R \times R}} \mathbf{w}_r \\ &= \sum_{r=1}^R ((\mathbb{W}_R)^T \mathbb{S})_{r,n} \mathbf{w}_r = \sum_{r=1}^R (\mathbf{w}_r^T \mathbf{s}_n) \mathbf{w}_r = \sum_{r=1}^R (\mathbf{s}_n, \mathbf{w}_r)_{\mathbb{R}^M} \mathbf{w}_r, \end{aligned}$$

where $(\cdot, \cdot)_{\mathbb{R}^M}$ denotes the Euclidean scalar product in \mathbb{R}^M . Therefore, the columns $\{\mathbf{w}_1, \dots, \mathbf{w}_R\}$ of \mathbb{W}_R constitute an orthonormal basis for the column space $\text{Col}(\mathbb{S})$ of \mathbb{S} . Moreover, from the above calculations follow that $(\mathbb{B}_R)_{r,n} = (\mathbf{s}_n, \mathbf{w}_r)_{\mathbb{R}^M}$ for $n = 1, \dots, N$ and $r = 1, \dots, R$.

Assume we seek to approximate the columns of \mathbb{S} by means of L orthonormal vectors $\{\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_L\}$, with $L < R$. It is an easy matter to show that for each \mathbf{s}_n , $n = 1, \dots, N$, the element of $\text{span}\{\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_L\}$ closest to \mathbf{s}_n in the Euclidean norm $\|\cdot\|_{\mathbb{R}^M}$ is given by

$$\sum_{l=1}^L (\mathbf{s}_n, \tilde{\mathbf{w}}_l)_{\mathbb{R}^M} \tilde{\mathbf{w}}_l.$$

Hence, we could measure the error committed by approximating the columns of \mathbb{S} via the vectors $\{\tilde{\mathbf{w}}_l\}_{l=1}^L$ through the quantity

$$\varepsilon(\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_L) = \sum_{n=1}^N \left\| \mathbf{s}_n - \sum_{l=1}^L (\mathbf{s}_n, \tilde{\mathbf{w}}_l)_{\mathbb{R}^M} \tilde{\mathbf{w}}_l \right\|_{\mathbb{R}^M}^2. \quad (33)$$

The following theorem states that the basis $\{\mathbf{w}_1, \dots, \mathbf{w}_L\}$ consisting of the first L left singular values of \mathbb{S} minimizes (33) among all the orthonormal bases of \mathbb{R}^L .

Theorem 4.1 (Schmidt-Eckart-Young). *Consider a rectangular matrix $\mathbb{S} = [\mathbf{s}_1 \mid \dots \mid \mathbf{s}_N] \in \mathbb{R}^{M \times N}$ with rank $R \leq \min\{M, N\}$, and let $\mathbb{S} = \mathbb{W} \Sigma \mathbb{Z}^T$ be the singular value decomposition (SVD) of \mathbb{S} , with $\mathbb{W} = [\mathbf{w}_1 \mid \dots \mid \mathbf{w}_M] \in \mathbb{R}^{M \times M}$ and $\mathbb{Z} = [\mathbf{z}_1 \mid \dots \mid \mathbf{z}_N] \in \mathbb{R}^{N \times N}$ orthogonal matrices, and $\Sigma \in \mathbb{R}^{M \times N}$ defined as in (27). Further, let $\mathbb{S} = \mathbb{W}_R \mathbb{D} \mathbb{Z}_R^T = \mathbb{W}_R \mathbb{B}_R$ be the compact form of the SVD of \mathbb{S} , with $\mathbb{W}_R \in \mathbb{R}^{M \times R}$ and $\mathbb{Z}_R \in \mathbb{R}^{N \times R}$ defined as in (32), $\mathbb{D} = \text{diag}(\sigma_1, \dots, \sigma_R) \in \mathbb{R}^{R \times R}$, and $\mathbb{B}_R = \mathbb{D} \mathbb{Z}_R^T \in \mathbb{R}^{R \times N}$. Suppose that $\hat{\mathbb{W}}_R \in \mathbb{R}^{M \times R}$ denotes a matrix with pairwise orthonormal columns $\hat{\mathbf{w}}_r$, $r = 1, \dots, R$, and that the expansion of*

the columns of \mathbb{S} in the basis $\{\widehat{\mathbf{w}}_n\}_{n=1}^N$ is given by

$$\mathbb{S} = \widehat{\mathbb{W}}_R \mathbb{C}_R,$$

with $\mathbb{C}_R \in \mathbb{R}^{R \times N}$, defined as

$$(\mathbb{C}_R)_{r,n} = (\widehat{\mathbf{w}}_r, \mathbf{s}_n)_{\mathbb{R}^M} \quad \text{for } 1 \leq r \leq R, 1 \leq n \leq N.$$

Then for every $L \in \{1, \dots, R\}$ we have

$$\|\mathbb{S} - \mathbb{W}_L \mathbb{B}_L\|_F \leq \|\mathbb{S} - \widehat{\mathbb{W}}_L \mathbb{C}_L\|_F. \quad (34)$$

Here, $\|\cdot\|_F$ denotes the Frobenius norm given by

$$\|\mathbb{A}\|_F = \sqrt{\sum_{m=1}^M \sum_{n=1}^N |A_{m,n}|^2} = \sqrt{\text{tr}(\mathbb{A}^T \mathbb{A})} \quad \text{for any } \mathbb{A} \in \mathbb{R}^{M \times N},$$

the matrix \mathbb{W}_L (respectively, $\widehat{\mathbb{W}}_L$) denotes the first L columns of \mathbb{W} (resp., $\widehat{\mathbb{W}}$), and \mathbb{B}_L (resp., \mathbb{C}_L) denotes the first L rows of \mathbb{B} (resp., \mathbb{C}) [58].

The orthonormal basis $\{\mathbf{w}_1, \dots, \mathbf{w}_L\}$ is known as the *POD basis* of rank L . Returning to the POD-Galerkin RB method, we set $\boldsymbol{\psi}_l = \mathbf{w}_l$, for all $l = 1, \dots, L$, so that

$$\mathbb{V} = [\mathbf{w}_1 \mid \dots \mid \mathbf{w}_L].$$

Hence, in the reduced basis problem (21) the test and trial functions are picked from the subspace V_{rb} of V_h spanned by the functions $\{\psi_l\}_{l=1}^L$, given by

$$\psi_l(\mathbf{x}) = \sum_{m=1}^M \psi_l^{(m)} \phi_m(\mathbf{x}) = \sum_{m=1}^M (\mathbf{w}_l)_m \phi_m(\mathbf{x}) \quad \text{for } 1 \leq l \leq L. \quad (35)$$

From a computational viewpoint, the first L left singular vectors $\{\mathbf{w}_l\}_{l=1}^L$ of \mathbb{S} can be efficiently computed through the so-called *method of snapshots*. We should distinguish two cases:

- (a) if $M \leq N$: directly solve the eigenvalue problems

$$\mathbb{S} \mathbb{S}^T \mathbf{w}_l = \lambda_l \mathbf{w}_l \quad \text{for } 1 \leq l \leq L;$$

- (b) if $M > N$: compute the *correlation* matrix $\mathbb{M} = \mathbb{S}^T \mathbb{S}$ and solve the eigenvalue problems

$$\mathbb{M} \mathbf{z}_l = \lambda_l \mathbf{z}_l \quad \text{for } 1 \leq l \leq L.$$

Then, by (30) we have

$$\mathbf{w}_l = \frac{1}{\sqrt{\lambda_l}} \mathbb{S} \mathbf{z}_l \quad \text{for } 1 \leq l \leq L.$$

References

- [1] Amsallem, D. (2010). *Interpolation on manifolds of CFD-based fluid and finite element-based structural reduced-order models for on-line aeroelastic predictions*. Doctoral dissertation, Department of Aeronautics and Astronautics, Stanford University.
- [2] Ballarin, F., Manzoni, A., Quarteroni, A., Rozza, G. (2014). *Supremizer stabilization of POD-Galerkin approximation of parametrized Navier-Stokes equations*. MATHICSE Technical Report, École Polytechnique Fédérale de Lausanne.
- [3] Barrault, M., Maday, Y., Nguyen, N. C., Patera, A. T. (2004). *An 'empirical interpolation' method: Application to efficient reduced-basis discretization of partial differential equations*. *Comptes Rendus Mathématique*, 339(9):667-672.
- [4] Barthelmann, V., Novak, E., Ritter, K. (2000). *High-dimensional polynomial interpolation on sparse grids*. *Advances in Computational Mathematics*, 12(4):273-288.
- [5] Bendsøe, M. P., Sigmund, O. (2004). *Topology optimization: Theory, methods and applications*. Heidelberg, DE: Springer Science & Business Media.
- [6] Brown, P. F., Pietra, V. J. D., Pietra, S. A. D., Mercer, R. L. (1993). *The mathematics of statistical machine translation: Parameter estimation*. *Computational linguistics*, 19(2):263-311.
- [7] Buffa, A., Maday, Y., Patera, A. T., Prud'Homme, C., Turinici, G. (2012). *A priori convergence of the greedy algorithm for the parametrized reduced basis method*. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46:595-603.
- [8] Burkardt, J., Gunzburger, M., Lee, H. C. (2006). *POD and CVT-based reduced-order modeling of Navier-Stokes flows*. *Computer Methods in Applied Mechanics and Engineering*, 196:337-355.
- [9] Casenave, F., Ern, A., Lelièvre, T. (2015). *A nonintrusive reduced basis method applied to aeroacoustic simulations*. *Advances in Computational Mathematics*, 41:961-986.
- [10] Chaturantabut, S., Sorensen, D. C. (2010). *Nonlinear model reduction via discrete empirical interpolation*. *SIAM Journal on Scientific Computing*, 32(5):2737-2764.
- [11] Caloz, G., Rappaz, J. (1997). *Numerical analysis and bifurcation problems*. *Handbook of numerical analysis*, 5(2):487-637.
- [12] Chen, W., Hesthaven, J. S., Junqiang, B., Yang, Z., Tihao, Y. (2017). *A greedy non-intrusive reduced order model for fluid dynamics*. Submitted to American Institute of Aeronautics and Astronautics.
- [13] Cybenko, G. (1988). *Continuous valued neural networks with two hidden layers are sufficient*. Technical Report, Department of Computer Science, Tufts University.
- [14] Cybenko, G. (1989). *Approximation by superpositions of a sigmoidal function*. *Mathematics of Control, Signals, and Systems*, 2(4):303-314.
- [15] De Boor, C. (1978). *A practical guide to splines*. New York, NY: Springer-Verlag.
- [16] Deparis, S. (2008). *Reduced basis error bound computation of parameter-dependent Navier-Stokes equations by the natural norm approach*. *SIAM Journal of Numerical Analysis*, 46(4):2039-2067.
- [17] Dhondt, G. (2014). *CalculiX CrunchiX user's manual*. Available at http://web.mit.edu/calculix_v2.7/CalculiX/ccx_2.7/doc/ccx/node1.html.
- [18] Eftang, J. L. (2008). *Reduced basis methods for partial differential equations*. Master thesis, Department of Mathematical Sciences, Norwegian University of Science and Technology.
- [19] Elman, H. C., Silvester, D. J., Wathen, A. (2004). *Finite elements and fast iterative solvers with applications in incompressible fluid dynamics*. New York, NY: Oxford University Press.

- [20] Fahlman, S. E. (1988). *An empirical study of learning speed in back-propagation networks*. Technical Report CMU-CS-88-162, CMU.
- [21] Gill, P. E., Murray, W., Wright, M. H. (1981). *Practical optimization*. Academic Press.
- [22] Hagan, M. T., Menhaj, M. B. (1994). *Training feedforward networks with the Marquardt algorithm*. IEEE Transactions on Neural Networks, 5(6):989-993.
- [23] Hagan, M. T., Demuth, H. B., Beale, M. H., De Jesús, O. (2014). *Neural Network Design, 2nd Edition*. Retrieved from <http://hagan.okstate.edu/NNDesign.pdf>.
- [24] Hassdonk, B. (2013). *Model reduction for parametric and nonlinear problems via reduced basis and kernel methods*. CEES Computational Geoscience Seminar, Stanford University.
- [25] Haykin, S. (2004). *Neural Networks: A comprehensive foundation*. Upper Saddle River, NJ: Prentice Hall.
- [26] Hebb, D. O. (1949). *The organization of behaviour: A neuropsychological theory*. New York, NY: John Wiley & Sons.
- [27] Liang, Y. C., Lee, H. P., Lim, S. P., Lin, W. Z., Lee, K. H., Wu, C. G. (2002) *Proper Orthogonal Decomposition and its applications - Part I: Theory*. Journal of Sound and Vibration, 252(3):527-544.
- [28] Hassibi, B., Stork, D. G. (1993). *Second order derivatives for network pruning: Optimal Brain Surgeon*. Advances in neural information processing systems, 164-171.
- [29] Hesthaven, J. S., Starn, B., Rozza, G. (2016). *Certified reduced basis methods for parametrized partial differential equations*. New York, NY: Springer.
- [30] Hesthaven, J. S., Starn, B., Zhang, S. (2014). *Efficient greedy algorithms for high-dimensional parameter spaces with applications to empirical interpolation and reduced basis methods*. ESAIM: Mathematical Modelling and Numerical Analysis, 48(1):259-283.
- [31] Hopfield, J. J. (1982). *Neural networks and physical systems with emergent collective computational abilities*. Proceedings of the National Academy of Science, 79:2554-2558.
- [32] Imam, R. L. (2008). *Latin hypercube sampling*. Encyclopedia of Quantitative Risk Analysis and Assessment.
- [33] Jaggi, C., Iapichino, L., Rozza, G. (2014). *An improvement on geometrical parametrizations by transfinite maps*. Comptes Rendus de l'Académie des Sciences Paris, Series I, 352:263-268.
- [34] Kaelbling, L. P., Littman, M. L., Moore, A. W. (1996). *Reinforcement Learning: A Survey*. Journal of Artificial Intelligence Research, 4:237-285.
- [35] Kohavi, R. (1995). *A study of cross-validation and bootstrap for accuracy estimation and model selection*. Proceedings of the 40th International Joint Conference on Artificial Intelligence, 2(12):1137-1143.
- [36] Kohonen, T. (1998). *The self-organizing map*. Neurocomputing, 21(1-3):1-6.
- [37] Kriesel, D. (2007). *A Brief Introduction to Neural Networks*. Retrieved from http://www.dkriesel.com/en/science/neural_networks.
- [38] Le Maître, O., Knio, O. M. (2010). *Spectral methods for uncertainty quantification with applications to computational fluid dynamics*. Berlin, DE: Springer Science & Business Media.
- [39] Lee, E. B., Markus, L. (1967). *Foundations of optimal control theory*. New York, NY: John Wiley & Sons.
- [40] Maday, Y. (2006) *Reduced basis method for the rapid and reliable solution of partial differential equations*. Proceedings of the International Congress of Mathematicians, Madrid, Spain, 1255-1269.
- [41] Marquardt, D. W. (1963). *An algorithm for least-squares estimation of nonlinear parameters*. Journal of the Society for Industrial and Applied Mathematics, 11(2):431-441.

- [42] The MathWorks, Inc. (2016). *Machine learning challenges: Choosing the best model and avoiding overfitting*. Retrieved from <https://it.mathworks.com/campaigns/products/offer/common-machine-learning-challenges.html>.
- [43] Mitchell, W., McClain, M. A. (2010). *A collection of 2D elliptic problems for testing adaptive algorithms*. NISTIR 7668.
- [44] Manzoni, A., Negri, F. (2016). *Automatic reduction of PDEs defined on domains with variable shape*. MATHICSE technical report, École Polytechnique Fédérale de Lausanne.
- [45] Møller, M. D. (1993). *A scaled conjugate gradient algorithm for fast supervised learning*. Neural Networks, 6:525-533.
- [46] McClelland, J. L., Rumelhart, D. E. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, UK: MIT Press.
- [47] Nielsen, M. A. (2015). *Neural Networks and Deep Learning*. Determination Press.
- [48] Negri, F., Manzoni, A., Amsallem, D. (2015). *Efficient model reduction of parametrized systems by matrix discrete empirical interpolation*. Journal of Computational Physics, 303:431-454.
- [49] Persson, P. O. (2002). *Implementation of finite-element based Navier-Stokes solver*. Massachussets Institue of Technology.
- [50] Prud'homme, C., Rovas, D. V., Veroy, K., Machiels, L., Maday, Y., Patera, A. T., Turinici, G. (2002). *Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods*. Journal of Fluids Engineering, 124(1):70-80.
- [51] Quarteroni, A. (2010). *Numerical models for differential problems* (Vol. 2). New York, NY: Springer Science & Business Media.
- [52] Quarteroni, A., Manzoni, A., Negri, F. (2015). *Reduced basis methods for partial differential equations: An introduction* (Vol. 92). New York, NY: Springer, 2015.
- [53] Rannacher, R. (1999). *Finite element methods for the incompressible Navier-Stokes equations*. Lecture notes, Institute of Applied Mathematics, University of Heidelberg.
- [54] Riedmiller, M., Braun, H. (1993). *A direct adaptive method for faster backpropagation learning: The rprop algorithm*. Neural Networks, IEEE International Conference on, 596-591.
- [55] Rosenblatt, F. (1958). *The perceptron: A probabilistic model for information storage and organization in the brain*. Psychological Review, 65:386-408.
- [56] Rudin, W. (1964). *Principles of mathematical analysis* (Vol. 3). New York, NY: McGraw-Hill.
- [57] Stergiou, C., Siganos, D. (2013). *Neural Networks*. Retrieved from https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html#Introductiontoneuralnetworks.
- [58] Volkwein, S. (2008). *Model reduction using proper orthogonal decomposition*. Lecture notes, Department of Mathematics, University of Konstanz.
- [59] Widrow, B., Hoff, M. E. (1960). *Adaptive switching circuits*. Proceedings WESCON, 96-104.