THÈSE DE DOCTORAT DE

École Centrale de Nantes
COMUE UNIVERSITÉ BRETAGNE LOIRE

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies
de l'Information et de la Communication
Spécialité : Mathématiques et leurs Interactions

Par

Oleg BALABANOV

Randomized linear algebra for model order reduction

Thèse présentée et soutenue à l'École Centrale de Nantes, le 11 Octobre, 2019
Unité de recherche : Laboratoire de Mathématiques Jean Leray, UMR CNRS 6629
Thèse N° :

Rapporteurs avant soutenance :

Bernard HAASDONK	Professeur, University of Stuttgart
Tony LELIÈVRE	Professeur, École des Ponts ParisTech

Composition du Jury :

Président :	Albert COHEN	Professeur, Sorbonne Université
Examinateurs :	Christophe PRUD'HOMME	Professeur, Université de Strasbourg
	Laura GRIGORI	Directeur de recherche, Inria Paris, Sorbonne Université
	Marie BILLAUD-FRIESS	Maître de Conférences, École Centrale de Nantes
Dir. de thèse :	Anthony NOUY	Professeur, École Centrale de Nantes
Dir. de thèse :	Núria PARÉS MARINÉ	LaCàN, Universitat Politècnica de Catalunya, Spain

ÉCOLE CENTRALE DE NANTES

Nantes, France

UNIVERSITAT POLITÉCNICA DE CATALUNYA

Barcelona, Spain

Randomized linear algebra for model order reduction

by

Oleg Balabanov

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the

Doctoral school of Mathematics and STIC
École Centrale de Nantes

Department of Civil and Environmental Engineering
Universitat Politécnica de Catalunya

September 2019

President :	Albert Cohen	Sorbonne Université
Jury members :	Bernard Haasdonk	University of Stuttgart
	Tony Lelièvre	École des Ponts ParisTech
	Christophe Prud'homme	Université de Strasbourg
	Laura Grigori	Inria Paris, Sorbonne Université
	Marie Billaud-Friess	École Centrale de Nantes
Advisor :	Anthony Nouy	École Centrale de Nantes
Co-advisor:	Núria Parés Mariné	Universitat Politècnica de Catalunya

To beloved mom and dad.

Abstract

Solutions to high-dimensional parameter-dependent problems are in great demand in the contemporary applied science and engineering. The standard approximation methods for parametric equations can require computational resources that are exponential in the dimension of the parameter space, which is typically referred to as the curse of dimensionality. To break the curse of dimensionality one has to appeal to nonlinear methods that exploit the structure of the solution map, such as projection-based model order reduction methods.

This thesis proposes novel methods based on randomized linear algebra to enhance the efficiency and stability of projection-based model order reduction methods for solving parameter-dependent equations. Our methodology relies on random projections (or random sketching). Instead of operating with high-dimensional vectors we first efficiently project them into a low-dimensional space. The reduced model is then efficiently and numerically stably constructed from the projections of the reduced approximation space and the spaces of associated residuals.

Our approach allows drastic computational savings in basically any modern computational architecture. For instance, it can reduce the number of flops and memory consumption and improve the efficiency of the data flow (characterized by scalability or communication costs). It can be employed for improving the efficiency and numerical stability of classical Galerkin and minimal residual methods. It can also be used for the efficient estimation of the error, and post-processing of the solution of the reduced order model. Furthermore, random sketching makes computationally feasible a dictionary-based approximation method, where for each parameter value the solution is approximated in a subspace with a basis selected from a dictionary of vectors. We also address the efficient construction (using random sketching) of parameter-dependent preconditioners that can be used to improve the quality of Galerkin projections or for effective error certification for problems with ill-conditioned operators. For all proposed methods we provide precise conditions on the random sketch to guarantee accurate and stable estimations with a user-specified probability of success. A priori estimates to determine the sizes of the random matrices are provided as well as a more effective adaptive procedure based on a posteriori estimates.

Key words— model order reduction, parameter-dependent equations, random sketching, subspace embedding, reduced basis, dictionary-based approximation, preconditioner

Acknowledgements

This work is a result of a fruitful four years cooperation with Anthony Nouy – my advisor, teacher and simply, a brilliant mind. It was a blessing to have a chance to work with him. I would like to express a sincere gratitude for his special guidance, patience and, especially, our vast discussions, every single one of which was very instructive and fascinating. Anthony's exceptional view and ideas have led to very interesting discoveries partially presented in this manuscript.

A significant contribution to this work was also done by my co-supervisor Núria Parés Mariné. Our few, yet productive meetings have led to a better presentation and polishing of the methodology. The discussions with Núria have as well resulted in many interesting ideas and open questions, which lie out of the scope of the present manuscript and are to be investigated in the future. In addition to Anthony and Núria, I also would like to thank Pedro Díez for all his guidelines and help since the beginning of my scientific career.

I additionally would like to thank professors Bernard Haasdonk, Tony Lelièvre, Albert Cohen, Christophe Prud'homme, Laura Grigori and Marie Billaud-Friess for their kind agreement to be on my defence committee.

Another word of gratitude goes to the colleagues at GeM and LMJL at the Centrale Nantes, and LaCàn at the Polytechnic University of Catalonia for enriching my life in the lab with cheerful debates and discussions.

Special thanks are to Daria Koliesnikova.

Finally, I am infinitely grateful to my family for unconditional support during this difficult path. My loving, kind and mindful parents Balabanov Ihor and Balabanova Tetyana were always there for me. My brothers Valeriy and Oleksandr never stopped boosting my brain by clever riddles and new interesting ideas. Also thanks to all my friends for their support.

Contents

Chapter 1

Introduction and overview of the literature

Contents

The past decades have seen a tremendous advance in the computational technology that had a monumental impact on modern science and engineering. Nowadays, the scientists, analytics and engineers in their developments rely on machine computations, which accelerate everyday throughout improvement of hardware. Along with enhancement of the hardware, a crucial role for the contemporary industry is played by the development of new sophisticated algorithms providing ability of finding rapid, reliable and non-intrusive solutions to difficult tasks.

Numerical modeling has become an inherent step for practically any industrial project. Many problems in numerical modeling, however, still remain intractable on industrial level even with the recent vast growth of computational capabilities and progress in the development of numerical methods. Finding solutions to complex parametric equations is an excellent example of an industrially demanded problem, suffering from the aforementioned limitations. This manuscript makes a step forward to enhance the effectiveness of the existing methods [24, 25, 89, 128, 129, 133] for parametric problems by exploiting randomized linear algebra. For the introduction of randomized linear algebra, see Section 1.2.

Parametric models arise throughout many fields of applied science and engineering. Examples include modeling of heat transfer, diffusion, wave scattering phenomenon, (fluid-)structure interactions, fluid flows, problems in quantum mechanics and solid state physics, population dynamics, problems in finance and others. The models can be formulated in terms of algebraic equations, or governed by partial differential equations (PDEs) or integral equations. If the equations are not given in algebraic form, they should be discretized using approximation methods (finite elements, finite volumes, discontinuous Galerkin methods, etc.) [52, 80, 92, 152], in their turn yielding large-scale systems of parametrized algebraic equations. This work focuses on improving methods related to parameter-dependent systems of algebraic equations, possibly obtained after a discretization of the original problem. We then assume that the discrete solution may serve as a “truth” solution for the original undiscretized problem from any practical perspective. Furthermore, we shall restrict ourselves to linear steady models, noting that similar considerations apply also to a wide range of nonlinear or non-stationary problems.

The discrete problem of interest can be formulated as follows: find $s(\mu) := l(\mathbf{u}(\mu);\mu)$ with $\mathbf{u}(\mu)$ satisfying

$$\mathbf{A}(\mu)\mathbf{u}(\mu) = \mathbf{b}(\mu),\ \mu \in \mathcal{P}, \tag{1.1}$$

where $\mathcal{P}$ is the parameter set, $l(\cdot;\mu) : \mathbb{K}^n \to \mathbb{K}$ (with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) is a parameter-dependent functional for extraction of the quantity of a interest, $\mathbf{A}(\mu) \in \mathbb{K}^{n\times n}$ is a large-scale parameter-dependent matrix and $\mathbf{b}(\mu) \in \mathbb{K}^n$ is a parameter-dependent vector.

For the typical applications we have $\mathcal{P} \subseteq \mathbb{R}^e$, but the parameters may as well be chosen as elements of infinite dimensional spaces such as function spaces. The parameters can describe material properties, domain geometry (for PDEs) and many more. For applications where the parameters describing the model are unknown and one deals with a set of linear systems, $\mathcal{P}$ can be taken as the system's index set. The parametric equations can be considered in many contexts such as optimization, control, uncertainty quantification (UQ) and inverse problems. In the contexts of optimization and control one seeks a rapid (possibly in real-time) prediction of an objective function (derived from the quantity of interest $s(\mu)$). The predictions are then used for driving the model to a configuration, which satisfies the desired conditions or fulfills the problem's constraints. The objective of UQ is to provide statistical analysis (typically using Monte-Carlo methods) of a problem with random input variables. The objective for inverse problems is to determine the value of the parameter (or its probability distribution) from partial information on the solution. All the mentioned contexts require an (approximate) solution to a large-scale parameter-dependent system of equations.

The standard approach for solving (1.1) proceeds with an approximation of the solution map $\mathbf{u} : \mathcal{P} \rightarrow \mathbb{K}^n$ on some subspace of functions. This can be done by interpolation (also called stochastic collocation in the UQ community) [9, 18, 126] or (stochastic) Galerkin projection [10, 76]. Over the past years a variety of approximation tools have been considered including polynomials, piecewise polynomials, wavelets, etc. [10, 44, 76].

Approximation of the solution map on classical approximation spaces may lead to solutions of prohibitively large problems with exponential complexity with respect to the dimension of the parameter space. This is typically referred to as the curse of dimensionality [22, 23]. A remedy can be to reduce the complexity of the problem by exploiting the structure of the solution map. For instance, the solution $\mathbf{u}(\mu)$ may have certain properties such as symmetries or anisotropies, be almost constant along some directions while having strong variations along others. Such properties of $\mathbf{u}(\mu)$ can be exploited by sparse approximation methods [55, 63, 116]. These nonlinear methods consist in selecting an approximation space using a dictionary of functions. For some problems the dictionary-based approximation can break the curse of dimensionality [49, 116]. An alternative to the sparse approximation methods are the so-called projection-based model order reduction methods, which is the focus of the present thesis and is described below.

1.1 Projection-based model reduction

Model order reduction (MOR) methods [24, 25, 89, 119, 128] are developed for the efficient output of an approximate solution for each parameter value. Unlike methods that provide an approximation of the solution map in an explicit form, MOR methods proceed with a reduction of the complexity of the model and its subsequent efficient solution for each parameter value, therefore yielding an efficient implicit approximation of the solution map. The construction of the reduced order model usually involves expensive computations but is performed only once in the offline stage. Then for each parameter value the reduced model is used for rapid approximation of the solution (typically) with a computational cost independent of the dimension of the original system of equations.

1.1.1 Semi-discrete framework

Throughout the manuscript a general setting is considered with notations that are standard in the context of variational methods for PDEs. We use the notions of solution space $U := \mathbb{K}^n$ and dual space $U' := \mathbb{K}^n$ to specify the origin of vectors and matrices involved in the discrete problem. In the framework of numerical methods for PDEs, if a vector represents an element from the solution space for the PDE, then it is said to belong to U. On the other hand, the vectors identified with functions from the dual space are considered to lie in U'. The canonical ℓ_2-inner product between a vector from U and a vector from U' represents the duality pairing between the solution space and the dual space.

The spaces U and U' are equipped with inner products $\langle \cdot, \cdot \rangle_U := \langle \mathbf{R}_U \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle_{U'} := \langle \cdot, \mathbf{R}_U^{-1} \cdot \rangle$, where $\langle \cdot, \cdot \rangle$ is the canonical ℓ_2-inner product and $\mathbf{R}_U \in \mathbb{K}^{n \times n}$ is some self-adjoint (symmetric if $\mathbb{K} = \mathbb{R}$ and Hermitian if $\mathbb{K} = \mathbb{C}$) positive definite matrix. In addition, we let $\| \cdot \|_U$ and $\| \cdot \|_{U'}$ be the norms associated with $\langle \cdot, \cdot \rangle_U$ and $\langle \cdot, \cdot \rangle_{U'}$. Note that in this case $\| \cdot \|_{U'}$ corresponds to the canonical (dual) norm

$$\| \cdot \|_{U'} = \max_{\mathbf{w} \in U} \frac{\langle \cdot, \mathbf{w} \rangle}{\| \mathbf{w} \|_U}.$$

The inner product $\langle \cdot, \cdot \rangle_U$ between two vectors from U is chosen as the inner product of the corresponding elements from the solution space for the PDE. The matrix $\mathbf{R}_U$ is seen as a map from U to U'. In the framework of numerical methods for PDEs, the entries of $\mathbf{R}_U$ can be obtained by evaluating inner products of corresponding basis functions. For example, for a PDE defined on the Sobolev space H^1_0, $\mathbf{R}_U$ may be chosen as the discrete Laplacian. This (semi-)discrete algebraic formulation is essentially equivalent to the variational formulations typically used for numerical methods for PDEs and integral equations, but is more convenient for the introduction of randomized linear algebra techniques. The usage of the discrete formulation rather than the variational formulation considerably simplifies the complexity of the theory and in particular, the proofs. Yet, our methodology in principle should also be applicable (or could be extended) to equations defined on infinite dimensional function spaces.

If a model is simply described by algebraic equations, the notions of solution spaces, dual spaces, etc., can be disregarded and $\mathbf{R}_U$ can be taken as identity.

Further, we shall consider projection-based MOR methods, where for each parameter value the solution $\mathbf{u}(\mu)$ is approximated by a projection $\mathbf{u}_r(\mu)$ onto (possibly parameter-dependent) subspace $U_r(\mu) \subset U$, called reduced approximation space, of small or moderately large dimension r.

1.1.2 State-of-the-art model reduction

Model order reduction based on projections on low-dimensional spaces (for PDEs) can be traced back to the late 1970s with broader development in 1980s and 1990s [6, 14, 17, 26, 117, 118]. It received a lot of attention at the beginning of the 21th century [85, 101, 102, 125, 127, 132]. Among projection-based MOR methods the most established are reduced basis (RB) and Proper Orthogonal Decomposition (POD) methods. Balanced Truncation [25, 79, 113] is another popular approach for reduction of parametric models, although it lies out of the scope of the present manuscript. Furthermore, in some cases the recycling Krylov methods [4, 98, 124, 154] also can be a good alternative to the RB and POD methods.

An excellent presentation of classical RB and POD methods can be found in [82]. These methods both consist of approximating the solution manifold $\mathcal{M} := \{\mathbf{u}(\mu) : \mu \in \mathcal{P}\}$ with a fixed low-dimensional space obtained from solution samples, called snapshots, at some parameter values. The difference between the two approaches is the way the approximation space is constructed.

Proper Orthogonal Decomposition

In the POD method the approximation space U_r is defined as the one which minimizes the projection error of the solution manifold in the mean-square sense. This method proceeds with computing a training set of snapshots $\{\mathbf{u}(\mu^i) : 1 \leq i \leq m\}$, and approximating the range of $\mathbf{U}_m = [\mathbf{u}(\mu^1), \mathbf{u}(\mu^2), \ldots, \mathbf{u}(\mu^m)]$ with a Singular Value Decomposition (SVD) of $\mathbf{R}_U^{1/2}\mathbf{U}_m$. In practice, the SVD of $\mathbf{R}_U^{1/2}\mathbf{U}_m$ can be performed by the so-called method of snapshots, which consists in solving the eigenvalue problem

$$\mathbf{G}\mathbf{t} = \lambda \mathbf{t},$$

where $[\mathbf{G}]_{i,j} = \langle \mathbf{u}(\mu^i), \mathbf{u}(\mu^j) \rangle_U, 1 \leq i,j \leq m$. If $\{(\lambda_i, \mathbf{t}_i) : 1 \leq i \leq l\}$, with $l = \mathrm{rank}(\mathbf{U}_m)$, is the solution to the eigenvalue problem ordered such that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_l$, then the approximation space can be taken as

$$U_r := \mathrm{span}\{\mathbf{U}_m \mathbf{t}_i : 1 \leq i \leq r\}.$$

It can be shown that such U_r minimizes the mean-square error over all r-dimensional subspaces of U. Moreover, the mean-square error associated with U_r is given by

$$\frac{1}{m} \sum_{i=1}^{m} \min_{\mathbf{w} \in U_r} \|\mathbf{u}(\mu^i) - \mathbf{w}\|_U^2 = \frac{1}{m} \sum_{i=r+1}^{l} \lambda_i.$$

Greedy algorithm

The objective of RB methods is usually to minimize the maximal error of approximation over the parameter set. The RB methods typically proceed with an iterative greedy construction of the reduced approximation space at iteration i enriching the basis for the reduced space by a snapshot associated with the maximum of an error indicator $\widetilde{\Delta}_i(\mu)$ (an estimation of the projection error $\min_{\mathbf{w}\in U_i}\|\mathbf{u}(\mu)-\mathbf{w}\|_U$ on U_i) at iteration $i-1$. The greedy algorithm is summarized in Algorithm 1.

Algorithm 1 Classical greedy algorithm

```
Given: $\mathcal{P}_{\text{train}}$, $\mathbf{A}(\mu)$, $\mathbf{b}(\mu)$, $\tau$.
Output: $U_r$
1. Set $i := 0$, $U_0 = \{\mathbf{0}\}$ and pick $\mu^1 \in \mathcal{P}_{\text{train}}$.
while $\max_{\mu\in\mathcal{P}_{\text{train}}} \widetilde{\Delta}_i(\mu) \geq \tau$ do
    2. Set $i := i+1$.
    3. Evaluate $\mathbf{u}(\mu^i)$ and set $U_i := U_{i-1} + \text{span}(\mathbf{u}(\mu^i))$.
    4. Update provisional online solver.
    5. Find $\mu^{i+1} := \underset{\mu\in\mathcal{P}_{\text{train}}}{\text{argmax}}\ \widetilde{\Delta}_i(\mu)$.
end while
```

In Algorithm 1, $\mathcal{P}_{\text{train}} \subseteq \mathcal{P}$ is the training set and τ is the desired tolerance of the algorithm. The error indicator $\widetilde{\Delta}_i(\mu)$ should be picked according to the particular situation. The typical choice is $\widetilde{\Delta}_i(\mu) = \Delta_i(\mathbf{u}_i(\mu);\mu)$, which estimates the error of an approximation $\mathbf{u}_i(\mu) \in U_i$ of $\mathbf{u}(\mu)$. The quasi-optimality of the greedy selection with such an error indicator can then be characterized by the quasi-optimality of $\mathbf{u}_i(\mu)$ compared to the best approximation in U_i, and the effectivity of the error estimator, which is explained below in details. Let us consider the i-th iteration of Algorithm 1. Define

$$\kappa_i(\mu) := \frac{\|\mathbf{u}(\mu)-\mathbf{u}_i(\mu)\|_U}{\min_{\mathbf{w}\in U_r}\|\mathbf{u}(\mu)-\mathbf{w}\|_U} \quad \text{and} \quad \sigma_i(\mu) := \frac{\Delta_i(\mathbf{u}_i(\mu);\mu)}{\|\mathbf{u}(\mu)-\mathbf{u}_i(\mu)\|_U},$$

and assume that $\min_{\mu\in\mathcal{P}_{\text{train}}}\sigma_i(\mu) \geq \sigma_0$ where σ_0 is a positive constant. Then

$$\min_{\mathbf{w}\in U_i}\|\mathbf{u}(\mu^{i+1})-\mathbf{w}\|_U \geq \frac{1}{\gamma_i}\max_{\mu\in\mathcal{P}_{\text{train}}}\min_{\mathbf{w}\in U_i}\|\mathbf{u}(\mu)-\mathbf{w}\|_U, \tag{1.2}$$

where $\gamma_i = \frac{1}{\sigma_0}\sup_{\mu\in\mathcal{P}_{\text{train}}}\kappa_i(\mu)\sigma_i(\mu)$.

The ideal error indicator is $\widetilde{\Delta}_i(\mu) = \|\mathbf{u}(\mu)-\mathbf{u}_i(\mu)\|_U$, where $\mathbf{u}_i(\mu)$ is the orthogonal projection of $\mathbf{u}(\mu)$ onto U_i. Such a choice for $\widetilde{\Delta}_i(\mu)$, however, requires computation and maintenance of $\mathbf{u}(\mu)$ on the whole training set of parameter values, which can be intractable for large training set $\mathcal{P}_{\text{train}}$. This problem can be circumvented by considering $\mathbf{u}_i(\mu)$ as the Galerkin projection and using a residual-based error estimator (introduced below). In this case the efficiency of Algorithm 1 is attained thanks to a local offline/online splitting of computations. More specifically, at the i-th iteration of the greedy algorithm a provisional online solver (obtained from a reduced model) associated with reduced subspace U_i is constructed allowing efficient evaluation of $\operatorname{argmax}_{\mu\in\mathcal{P}_{\text{train}}} \widetilde{\Delta}_i(\mu)$.

In [28, 35] the convergence of the greedy algorithms has been analyzed. In these works the authors basically proved that a greedy algorithm will generate an approximation space U_r with approximation error close to the optimal one given by the Kolmogorov r-width

$$d_r(\mathcal{M}) := \inf_{\dim(W_r)=r} \sup_{\mathbf{u}\in\mathcal{M}} \min_{\mathbf{w}\in W_r} \|\mathbf{u}-\mathbf{w}\|_U, \tag{1.3}$$

with the infinitum taken over all r-dimensional spaces. More rigorously, if

$$d_r(\mathcal{M}) \leq cr^{\alpha},$$

for some constants c and α, then the approximation error of U_r generated with a greedy algorithm will be at most Cr^{α}, with C being a constant independent of r, which implies the preservation of the rates of the algebraic decay of the Kolmogorov r-width. Furthermore, if we have an exponential decay

$$d_r(\mathcal{M}) \leq ce^{ar^{\alpha}}$$

for some constants c, a and α, then the greedy algorithm will converge as $Ce^{Ar^{\frac{\alpha}{\alpha+1}}}$ with constants C and A independent of r.

Galerkin projection

An approximation $\mathbf{u}_r(\mu)$ of $\mathbf{u}(\mu)$ in U_r can be obtained by Galerkin projection. Galerkin projection is such that the residual associated with the approximate solution is orthogonal to the approximation space:

$$\langle \mathbf{r}(\mathbf{u}_r(\mu),\mu), \mathbf{v}\rangle = 0, \ \forall \mathbf{v} \in U_r,$$

where $\mathbf{r}(\mathbf{u}_r(\mu),\mu) = \mathbf{b}(\mu) - \mathbf{A}(\mu)\mathbf{u}_r(\mu)$. For each parameter value the computation of the Galerkin projection requires the solution of a small $r \times r$ system of equations, called reduced system, which can be efficiently done in the online stage for each parameter value.

For coercive problems the quasi-optimality of the Galerkin projection can be characterized by the Cea's lemma, which states that (under the condition that $\theta(\mu) > 0$)

$$\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \leq \frac{\beta(\mu)}{\theta(\mu)} \min_{\mathbf{w}\in U_r} \|\mathbf{u}(\mu) - \mathbf{w}\|_U,$$

where

$$\theta(\mu) := \min_{\mathbf{x}\in U\backslash\{\mathbf{0}\}} \frac{\langle \mathbf{A}(\mu)\mathbf{x}, \mathbf{x}\rangle}{\|\mathbf{x}\|_U^2},$$
$$\beta(\mu) := \max_{\mathbf{x}\in U\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}}{\|\mathbf{x}\|_U},$$

are the operator's coercivity and continuity constants, respectively. Moreover, if the matrix $\mathbf{A}(\mu)$ is self-adjoint, then one has an improved quasi-optimality result,

$$\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \leq \sqrt{\frac{\beta(\mu)}{\theta(\mu)}} \min_{\mathbf{w}\in U_r} \|\mathbf{u}(\mu) - \mathbf{w}\|_U.$$

Error estimation

When the approximate solution $\mathbf{u}_r(\mu)$ has been obtained, one can provide a certification of its accuracy with an a posteriori upper bound of the error $\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U$. For this, we can proceed with a classical residual-based error estimator defined as

$$\Delta_r(\mathbf{u}_r(\mu), \mu) = \frac{\|\mathbf{r}(\mathbf{u}_r(\mu), \mu)\|_{U'}}{\eta(\mu)},$$

where $\eta(\mu)$ is a computable lower bound of the minimal singular value or the coercivity constant of $\mathbf{A}(\mu)$ (seen as an operator from U to U'), which can be obtained theoretically [82] or with the Successive Constraint Method [93].

Primal-dual correction

For problems where $l(\cdot, \mu)$ is a linear functional, i.e., $l(\cdot, \mu) = \langle \mathbf{l}(\mu), \cdot\rangle$ with $\mathbf{l}(\mu) \in U'$, we have the following error bound for the approximate output quantity $l(\mathbf{u}_r(\mu), \mu)$:

$$|l(\mathbf{u}(\mu), \mu) - l(\mathbf{u}_r(\mu), \mu)| \leq \|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \|\mathbf{l}(\mu)\|_{U'}.$$

The accuracy of the reduced model's prediction of the quantity of interest can be improved with a primal-dual correction. This approach consists in (besides approximation of the solution to the primal problem (1.1)) finding an approximation to the solution of the adjoint (or dual) problem, defined as

$$\mathbf{A}(\mu)^{\mathrm{H}} \mathbf{u}^{\mathrm{du}}(\mu) = -\mathbf{l}(\mu).$$

Having an approximate dual solution $\mathbf{u}_r^{\mathrm{du}}(\mu)$, one can improve the primal output quantity $l(\mathbf{u}_r(\mu), \mu)$ with the estimate

$$l(\mathbf{u}_r(\mu), \mu) - \langle \mathbf{u}_r^{\mathrm{du}}(\mu), \mathbf{r}(\mathbf{u}_r(\mu); \mu)\rangle,$$

so that,

$$|l(\mathbf{u}(\mu),\mu) - \Big(l(\mathbf{u}_r(\mu),\mu) - \langle \mathbf{u}_r^{\mathrm{du}}(\mu), \mathbf{r}(\mathbf{u}_r(\mu);\mu)\rangle\Big)| \\ \leq \|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \|\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{u}_r^{\mathrm{du}}(\mu) + \mathbf{l}(\mu)\|_{U'},$$

in this way improving the error (bound) by a factor equal to the dual residual error scaled by $\|\mathbf{l}(\mu)\|_{U'}$.

Parameter-separability

The online efficiency of the classical projection-based MOR methods is usually based on parameter separability. We say that a parameter-dependent quantity $\mathbf{v}(\mu)$ with values in a vector space V over a field $\mathbb{K}$ admits an affine representation (or is parameter-separable) if

$$\mathbf{v}(\mu) = \sum_{i=1}^{d} \mathbf{v}_i \lambda_i(\mu) \tag{1.4}$$

where $\lambda_i(\mu) \in \mathbb{K}$ and $\mathbf{v}_i \in V$. The offline precomputation of an affine representation of the reduced system of equations and the quantities needed for the evaluation of the residual error and the output quantity, allows rapid online computation of the quantity of interest and the residual error for each parameter value with a cost independent of the high dimension n. Such affine decompositions can be obtained from the affine decompositions of $\mathbf{A}(\mu)$, $\mathbf{b}(\mu)$ and $l(\cdot,\mu)$. For instance, if $\mathbf{U}_r$ is the matrix whose columns are the basis vectors for U_r and

$$\mathbf{A}(\mu) = \sum_{i=1}^{d_{\mathbf{A}}} \lambda_{\mathbf{A}}^{(i)}(\mu)\mathbf{A}^{(i)}, \quad \mathbf{b}(\mu) = \sum_{i=1}^{d_{\mathbf{b}}} \lambda_{\mathbf{b}}^{(i)}(\mu)\mathbf{b}^{(i)} \text{ and } l(\cdot,\mu) = \sum_{i=1}^{d_l} \lambda_l^{(i)}(\mu)\langle \mathbf{l}^{(i)}, \cdot\rangle,$$

with small $d_{\mathbf{A}}, d_{\mathbf{b}}, d_l \ll n$, then for each parameter value the primal output $l(\mathbf{u}_r(\mu),\mu)$ associated with the Galerkin projection on U_r can be computed as

$$l(\mathbf{u}_r(\mu),\mu) = \mathbf{a}_r(\mu)^{\mathrm{H}}\mathbf{l}_r(\mu),$$

with $\mathbf{a}_r(\mu)$ being the solution to the reduced system of equations

$$\mathbf{A}_r(\mu)\mathbf{a}_r(\mu) = \mathbf{b}_r(\mu),$$

where the quantities $\mathbf{l}_r(\mu) = \sum_{i=1}^{d_l} \lambda_l^{(i)}(\mu)[\mathbf{U}_r^{\mathrm{H}}\mathbf{l}_i]$, $\mathbf{A}_r(\mu) = \sum_{i=1}^{d_{\mathbf{A}}} \lambda_{\mathbf{A}}^{(i)}(\mu)[\mathbf{U}_r^{\mathrm{H}}\mathbf{A}^{(i)}\mathbf{U}_r]$ and $\mathbf{b}_r(\mu) = \sum_{i=1}^{d_{\mathbf{A}}} \lambda_{\mathbf{b}}^{(i)}(\mu)[\mathbf{U}_r^{\mathrm{H}}\mathbf{b}^{(i)}]$ can be efficiently assembled online for each parameter value by precomputing the terms $\mathbf{U}_r^{\mathrm{H}}\mathbf{l}^{(i)}$, $\mathbf{U}_r^{\mathrm{H}}\mathbf{A}^{(i)}\mathbf{U}_r$ and $\mathbf{U}_r^{\mathrm{H}}\mathbf{b}^{(i)}$ in the offline stage. Similar considerations also hold for computation of the residual norm and the primal-dual correction.

The affine decomposition of a parameter-dependent quantity can be derived theoretically or approximately obtained with the Empirical Interpolation Method (EIM) [16, 106].

1.1.3 Empirical interpolation method

Let us outline the EIM method since it often plays a crucial role for making the MOR algorithms non-intrusive and online-efficient. Since we here consider a discrete setting we shall present a discrete version of EIM noting that similar considerations also apply for infinite dimensional function spaces. Note that the essential ingredients of the discrete empirical interpolation method (DEIM) described below are that of the technique introduced in [72] for image reconstruction termed the gappy POD. This technique was used in the context of model order reduction for PDEs in [41, 67, 77, 84, 85] and further developed for (discrete) POD in [45].

Let $\mathbf{v}(\mu) \in V := \mathbb{K}^s$ be a parameter-dependent vector for which an affine representation (1.4) is desired. The DEIM consists of two components. First, a low-dimensional space $V = \mathrm{span}\{\mathbf{v}_i : 1 \leq i \leq d\}$ is obtained that approximates well the manifold $\{\mathbf{v}(\mu) : \mu \in \mathcal{P}\}$. This space is typically constructed with a greedy algorithm or POD.

Further an empirical interpolation is used for online-efficient computation of the coefficients $\lambda_1(\mu), \dots, \lambda_d(\mu)$ in (1.4). The affine representation of $\mathbf{v}(\mu)$ can be written as

$$\mathbf{v}(\mu) \approx \mathbf{V}_d \boldsymbol{\lambda}(\mu),$$

where $\mathbf{V}_d = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_d]$ and $\boldsymbol{\lambda}(\mu) = [\lambda_1(\mu), \lambda_2(\mu), \dots, \lambda_d(\mu)]^{\mathrm{T}}$. The vector of coefficients $\boldsymbol{\lambda}(\mu)$ is obtained by solving a sub-system, constructed from the (least linearly dependent) d rows of $\mathbf{V}_d \boldsymbol{\lambda}(\mu) = \mathbf{v}(\mu)$, of the form

$$\mathbf{S}_d \mathbf{v}(\mu) = \mathbf{S}_d \mathbf{V}_d \boldsymbol{\lambda}(\mu), \tag{1.5}$$

where $\mathbf{S}_d$ is a sampling matrix whose rows are d (disjoint) rows of the identity matrix. The Equation (1.5) can be efficiently assembled and then solved for each parameter-value in the online stage since it requires identification and operation with only d entries of the vector $\mathbf{v}(\mu)$ (and the vectors $\mathbf{v}_i, 1 \leq i \leq d$) rather than the entire (high-dimensional) vector. The sampling matrix $\mathbf{S}_d$ is usually obtained with a greedy algorithm, at the i-th iteration augmenting $\mathbf{S}_i$ with the k_i-th row of the identity matrix, where k_i is the index of the maximal entry of the vector

$$|\mathbf{v}_{i+1} - \mathbf{V}_i (\mathbf{S}_i \mathbf{V}_i)^{-1} \mathbf{S}_i \mathbf{v}_{i+1}|.$$

Note that the greedy construction of the sampling matrix $\mathbf{S}_d$ and the approximation space V can be performed simultaneously. In this case $\mathbf{v}_{i+1}$ would be the snapshot at the parameter value μ where the error of the reconstruction of $\mathbf{v}(\mu)$ was the maximal at the previous iteration. One may think of various improvements of the DEIM for finding an approximate affine representation. For instance, we can select the sampling matrix with the best points interpolation method [115] consisting in the solution of an optimization problem, rather than the greedy algorithm. Furthermore, the coefficients $\boldsymbol{\lambda}(\mu)$ can be obtained by an orthogonal projection of $\mathbf{v}(\mu)$ onto $V_d = \mathrm{span}(\mathbf{V}_d)$ with respect to a certain semi-inner product $\langle \cdot, \cdot \rangle_{V^*}$ chosen depending on the problem. Note that such an approximation is a generalization of DEIM, since it is reduced to DEIM if $\langle\cdot,\cdot\rangle_{V^*} = \langle\mathbf{S}_d\cdot,\mathbf{S}_d\cdot\rangle$. If V is equipped with inner product $\langle\cdot,\cdot\rangle_V = \langle\mathbf{R}_V\cdot,\cdot\rangle$, where $\mathbf{R}_V$ is some self-adjoint positive-definite matrix, then choosing $\langle\cdot,\cdot\rangle_{V^*} = \langle\mathbf{R}_V\mathbf{D}_k\cdot,\mathbf{D}_k\cdot\rangle$, where $\mathbf{D}_k$ is a diagonal matrix with $k \geq d$ nonzero entries (i.e., $\mathbf{D}_k = \mathbf{S}_k^{\mathrm{H}}\mathbf{S}_k$) can provide a better accuracy and numerical stability of approximation than the standard DEIM. For numerical methods for PDEs such an approach is equivalent to reducing the integration domain as we have in the so-called hyper-reduction methods [75, 137].

The DEIM can be used as a tool for finding approximate affine representations of parameter-dependent vectors such as $\mathbf{b}(\mu)$ and $\mathbf{l}(\mu)$. It can also be applied to parameter-dependent operators, such as $\mathbf{A}(\mu)$, by representing them as vectors in $\mathbb{K}^{n^2}$ (using a reshaping operation) [114]. Note that in some situations $\mathbf{A}(\mu)$, $\mathbf{b}(\mu)$ or $\mathbf{l}(\mu)$ cannot be accurately approximated with an affine representation with a small number of terms. A remedy can be to apply the DEIM (with a slight modification) to projections (or restrictions) of the vectors and operators onto the approximation space U_r.

1.1.4 Partitioning of the parameter domain

Classical RB approach becomes ineffective if the solution manifold $\mathcal{M}$ cannot be well approximated by a single low-dimensional subspace, i.e., if the manifold's Kolmogorov r-width does not have a fast decay with the dimension r. One can extend the classical RB method by considering a reduced subspace $U_r(\mu)$ depending on a parameter μ. One way to obtain $U_r(\mu)$ is to use a *hp*-refinement method as in [68, 69], which consists in partitioning (adaptively) the parameter set $\mathcal{P}$ into subsets $\{\mathcal{P}_i : 1 \leq i \leq M\}$ and in associating to each subset $\mathcal{P}_i$ a subspace $U_r^i \subset U$ of dimension at most r, therefore resulting in

$$U_r(\mu) := U_r^i, \text{ if } \mu \in \mathcal{P}_i, \ 1 \leq i \leq M. \tag{1.6}$$

For the method to be efficient (in particular, to outperform the classical approximation with a single approximation space), the value for M should not be very large compared to r and is usually chosen as $M = \mathcal{O}(r^\nu)$, for some small number ν, say $\nu = 2$ or 3.

Formally speaking, the *hp*-refinement method aims to approximate the solution manifold with a set, called library, of low-dimensional spaces. The associated nonlinear width of approximation is the following [144]

$$d_r(\mathcal{M}; M) = \inf_{\#\mathcal{L}_r = M} \sup_{\mathbf{u}\in\mathcal{M}} \min_{W_r\in\mathcal{L}_r} \min_{\mathbf{w}\in U_r} \|\mathbf{u}-\mathbf{w}\|_U, \tag{1.7}$$

where the infimum is taken over all libraries of M r-dimensional spaces. Clearly, a projection $\mathbf{u}_r(\mu)$ of $\mathbf{u}(\mu)$ onto approximation space $U_r(\mu)$ defined by (1.6) satisfies

$$d_r(\mathcal{M}; M) \leq \max_{\mu\in\mathcal{P}} \|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U. \tag{1.8}$$

Therefore, for the *hp*-refinement method to be effective, the solution manifold is required to be well approximable in terms of the measure $d_r(\mathcal{M}; M)$. As was revealed in [13] the decay of $d_r(\mathcal{M}; M)$, where $M = \mathcal{O}(r^\eta)$, of several parameter-dependent vectors may not be preserved by their sum. This property of the approximation width, however, can be crucial for problems where the solution is composed as a superposition of several contributions, which is a quite typical situation. For instance, this happens in PDEs with (multiple) transport phenomena. The *hp*-refinement method can also be sensitive to the parametrization and require a large number of subdomains in $\mathcal{P}$ for high-dimensional parameter domains. A modification of the *hp*-refinement method as in [83, 109] can lead to partial reduction of the aforementioned drawbacks. Furthermore, the dictionary-based approximation proposed in [13, 64, 97] can be seen as a promising alternative to the partitioning of the parameter domain. It provides an approximation of the solution manifold which is insensitive (or only weakly sensitive) to the parametrization. Moreover, in contrast to the partitioning methods, the decay of the dictionary-based r-width of several vectors is preserved by their sum, as was shown in [13] (see Chapter 3).

1.1.5 Minimal residual methods

The Galerkin projection can be very inaccurate for non-coercive or ill-conditioned problems. Indeed, for such problems choosing the test space (with respect to which the orthogonality of the residual is prescribed) as the approximation space can lead to dramatic instabilities. Therefore, a better choice of the test space is needed. In the context of numerical methods for PDEs this is a particularly important topic for convection-diffusion-reaction, wave and heterogeneous problems. Various approaches for the selection of suitable test spaces have been proposed in the context of reduced basis approximation methods [59, 107, 134, 160]. One of the simplest ways is to use minimal residual methods, where the test space is chosen to minimize the residual error $\|\mathbf{r}(\mathbf{u}_r(\mu); \mu)\|_{U'}$. More precisely, the minres projection is a Petrov-Galerkin projection defined by

$$\langle \mathbf{r}(\mathbf{u}_r(\mu), \mu), \mathbf{v} \rangle = 0, \ \forall \mathbf{v} \in V_r(\mu), \tag{1.9}$$

where $V_r(\mu) = \{\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{w} : \mathbf{w} \in U_r\}$. The major benefits of the minres methods over the classical Galerkin methods are the already mentioned improved stability of the projection for non-coercive problems and more effective residual-based error bounds of an approximation (see e.g. [38]). Furthermore, minres methods are better suited for combination with random sketching technique. There are few drawbacks of the minres methods. The first drawback is the increased online computational cost, since the reduced system of equations associated with (1.9) can contain much more terms in its affine expansion than the one associated with Galerkin projection. In addition, the orthogonalization of the reduced basis is not as effective for minres methods as for Galerkin methods to guarantee the numerical stability of the reduced system of equations. Another drawback is the high computational cost associated with the precomputation of the affine decomposition of the reduced system of equations from the basis vectors. Note that all these drawbacks can be circumvented with the random sketching technique proposed in Chapter 3 of this manuscript.

1.1.6 Parameter-dependent preconditioners for MOR

As was said, the performance of the projection-based methods for an approximate solution of (1.1) highly depends on the properties of $\mathbf{A}(\mu)$ such as the condition number. These properties can be improved with preconditioning. Let $\mathbf{P}(\mu)$ be an approximate inverse of $\mathbf{A}(\mu)$. Then the (approximate) solution of (1.1) can be obtained from

$$\mathbf{B}(\mu)\mathbf{u}(\mu) = \mathbf{f}(\mu), \ \mu \in \mathcal{P}, \tag{1.10}$$

where $\mathbf{B}(\mu) := \mathbf{R}_U\mathbf{P}(\mu)\mathbf{A}(\mu)$ and $\mathbf{f}(\mu) := \mathbf{R}_U\mathbf{P}(\mu)\mathbf{b}(\mu)$. If $\mathbf{P}(\mu)\mathbf{A}(\mu)$ is close to the identity matrix, then $\mathbf{B}(\mu)$ should have better properties than the original operator $\mathbf{A}(\mu)$, which implies a better effectiveness of projection-based methods. In particular, if $\mathbf{P}(\mu) = \mathbf{A}(\mu)^{-1}$ then (1.10) is perfectly conditioned. Furthermore, preconditioning can be used for effective error certification, which does not require computation of expensive stability constants as in the classical methods. The preconditioner can be taken as the inverse of $\mathbf{A}(\mu)^{-1}$ computed at some parameter value. A better choice is to construct $\mathbf{P}(\mu)$ by an interpolation of matrix inverse as in [160]. This approach is addressed in Chapter 4 of this manuscript. Note that for the Galerkin projection the preconditioning can be interpreted as an improvement of the test space. The minimal residual projection can also be viewed as a preconditioned Galerkin projection, where $\mathbf{P}(\mu)$ is taken as $\mathbf{R}_U^{-1}\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{R}_U^{-1}$.

1.1.7 Recent advances and the limitations of model reduction

In recent years projection-based MOR methods have received a substantial development [121]. A particular interest was dedicated to the theoretical analysis of convergence properties of the methods [62]. The results on Kolmogorov r-widths for parameter-dependent elliptic problems [57, 108] were extended to a wider class of problems in [56]. The optimality of greedy algorithms for reduced basis methods was analyzed in [28, 35, 81]. The Galerkin methods with an improved stability for ill-conditioned and non-coercive problems were developed in [1, 145, 160]. Moreover, new nonlinear approximation methods were introduced to tackle problems with slow decay of the Kolmogorov r-width [39, 64, 73, 97, 120, 130, 142]. Along with the analysis, a great effort was spent on improving the efficiency of the algorithms and enlarging the area of applicability of projection-based approximation methods. A special attention was given to online-efficient methods for problems with nonlinear operators by exploiting EIM [42, 45, 67, 156]. An important topic of effective, numerically stable, but yet efficient, error estimation/certification of a solution of parametric equations was addressed in [12, 43, 141, 143]. Finally, classical MOR algorithms executed in modern computational environments were considered in [5, 33, 90, 91, 99, 122].

Although considerably improved over the past decades, the applicability of projection-based MOR methods still remains highly limited. The central bottleneck is that parameter-dependent problems often do not show fast decays of the Kolmogorov r-width of the solution manifold, which is the essential condition for most projection-based MOR methods. In the context of PDEs this issue is particularly dramatic for problems with (moving) discontinuities. Such problems are typically tackled with so-called freezing (or space transformation) methods. The freezing approach [39, 120, 155] can be interpreted as preconditioning, where one first obtains a map (here assumed to be linear) $\mathbf{P}(\mu)$ among a certain family of mappings such that the manifold $\{\mathbf{v}(\mu) := \mathbf{P}(\mu)^{-1}\mathbf{u}(\mu) : \mu \in \mathcal{P}\}$ shows a better convergence of the Kolmogorov r-width, and then considers the solution of the right-preconditioned system of equations

$$\mathbf{A}(\mu)\mathbf{P}(\mu)\mathbf{v}(\mu) = \mathbf{b}(\mu).$$

The existing approaches for finding a suitable $\mathbf{P}(\mu)$, however, are highly intrusive or require very heavy offline computations, and are applicable only for the cases with one-dimensional or (to some extend) two-dimensional parameter domains. They also can involve computations in the online stage that depend on the high-dimension n, which is prohibited by many computational architectures. The techniques based on dictionary learning and compressed sensing have recently revealed their great potential for dealing with parameter-dependent problems with a slow decay of the Kolmogorov r-width [13, 64]. These techniques, however, also require further development to tackle complex problems.

1.2 Randomized linear algebra

Randomized linear algebra (RLA) is a popular approach for reduction of the computational cost of basic problems in numerical linear algebra such as products and low-rank approximations of high-dimensional matrices, and the solution of least-squares problems [88, 157]. It is broadly used in such fields as theoretical computer science, data analysis, statistical learning and compressed sensing. Open source routines for efficient RLA can be found for instance in [112, 153].

The random sketching technique is based on the dimension reduction by embedding a set (or a subspace) of high-dimensional vectors into a low-dimensional space without altering much the geometry (i.e., the inner products between vectors). The geometry can then be efficiently analyzed in the low-dimensional space without appealing to high-dimensional vectors. One way to perform an embedding of a set of vectors is the adaptive selection of the most relevant coordinates of the vectors either with a greedy selection or a leverage scores sampling [65, 66, 110, 136]. This approach, however, is not as efficient and robust for parameter-dependent problems or problems where the set of vectors is unknown a priori, as oblivious embeddings, discussed further.

The idea of oblivious sketching comes from the observation of Johnson and Lindenstrauss in their groundbreaking paper [95] stating that any set of N points in a Euclidean space can be randomly embedded into a Euclidean space of dimension $\mathcal{O}(\log(N))$, so that all pairwise distances between the points are nearly preserved with high probability. In [94, 103] the Johnson-Lindenstrauss type embeddings were applied to nearest-neighbor algorithms and in [74, 123] used for numerical linear algebra. Thereafter Achlioptas [2] showed that a similar performance as with standard Gaussian matrices can be attained also by considering more efficient discrete random matrices. The next breakthrough was done by Ailon and Chazelle [3] who proposed to construct the random embeddings with structured matrices, which lead to considerable improvements of the complexities of the previous algorithms. In the numerical linear algebra context, the random sketching technique with structured embeddings was firstly used by Sarlós in [139] to reduce the complexity of SVD, least-squares problem and the computation of products of matrices. These algorithms were then improved in [58, 87, 88, 131, 157] and recently developed even further [50, 150]. A rigorous theoretical analysis of linear random sketching can be found in [31, 88, 146, 147, 148, 150, 151, 157].

As was revealed in [15], the Johnson-Lindenstrauss type embeddings are strongly connected to the Restricted Isometry Property (RIP) introduced by Candes and Tao in [40] for sparse signal recovery [21]. The RIP property can be interpreted as quasi-isometry property of a linear map when restricted to sparse vectors. The RIP for Johnson-Lindenstrauss type embeddings can be shown to hold simply by arguing that this property is equivalent to the (almost) preservation of inner products between all vectors from low-dimensional subspaces. In [149] Tropp provided an empirical argument for an exact recovery of a sparse signal from its random projection with a greedy algorithm called Orthogonal Matching Pursuit (OMP). The theory of sparse signal recovery can be translated into best k-term approximation framework [53, 54]. In this manuscript we adapted some ideas from compressed sensing.

Further we introduce the random sketching technique and discuss its role for improving the efficiency and stability of projection-based MOR methods.

1.2.1 Random sketching: a least-squares problem

The introduction of the random sketching technique will be performed on a weighted least-squares problem. Assume that the goal is to obtain a vector $\mathbf{a} \in \mathbb{K}^d$ of coefficients, which minimizes

$$\|\mathbf{V}\mathbf{a}-\mathbf{b}\|_X, \tag{1.11}$$

where $\mathbf{V} \in \mathbb{K}^{n\times d}$ is a large-scale matrix with $d \ll n$, $\mathbf{b} \in \mathbb{K}^n$ is a large-scale vector, and $\|\cdot\|_X$ is a norm induced by a weighted Euclidean inner product $\langle\cdot,\cdot\rangle_X = \langle \mathbf{R}_X\cdot,\cdot\rangle$. The matrix $\mathbf{R}_X \in \mathbb{K}^{n\times n}$ is self-adjoint positive definite. The least-squares problem can be solved by considering the normal equation

$$\mathbf{V}^{\mathrm{H}}\mathbf{R}_X\mathbf{V}\mathbf{a} = \mathbf{V}^{\mathrm{H}}\mathbf{R}_X\mathbf{b}, \tag{1.12}$$

which is a small $d \times d$ system of equations. Assembling this system of equations, however, requires computation of a product $\mathbf{V}^{\mathrm{H}}\mathbf{R}_X\mathbf{V}$ of large-scale matrices, which needs (if $\mathbf{R}_X$ is sparse and has $\mathcal{O}(n)$ nonzero entries or is stored in an efficient implicit format) $\mathcal{O}(nd^2)$ flops and 2 passes over $\mathbf{V}$. This can be very expensive or even lead to a computational burden, e.g., if $\mathbf{V}$ does not fit into the RAM or its columns are maintained on multiple workstations.

The computational cost of minimizing (1.11) can be drastically reduced if one agrees to sacrifice a little of quality of the solution. Suppose that we seek a quasi-optimal solution satisfying

$$\|\mathbf{V}\mathbf{a}-\mathbf{b}\|_X \leq (1+\tau)\min_{\mathbf{x}\in\mathbb{K}^d}\|\mathbf{V}\mathbf{x}-\mathbf{b}\|_X, \tag{1.13}$$

for some τ. The coefficient τ has to be picked depending on the particular scenario. For the typical applications choosing $\tau < 10^{-1}$ should be enough to provide a good estimation of the solution. A solution satisfying (1.13) can be obtained by the sketching technique, which consists in constructing a suitable sketching matrix $\mathbf{\Theta} \in \mathbb{K}^{k\times n}$ with $k \ll n$, which allows efficient precomputation of products $\mathbf{\Theta}\mathbf{V}$ and $\mathbf{\Theta}\mathbf{b}$, and considering the solution to a small regression problem

$$\min_{\mathbf{x}\in\mathbb{K}^d}\|[\mathbf{\Theta}\mathbf{V}]\mathbf{x}-[\mathbf{\Theta}\mathbf{b}]\|. \tag{1.14}$$

The matrix $\mathbf{\Theta}$ typically has $k = \mathcal{O}(\tau^{-2}d)$ (up to a logarithmic factor in n or d) rows. The dominant computations are the products $\mathbf{\Theta}\mathbf{V}$ and $\mathbf{\Theta}\mathbf{b}$, which should be efficiently performed exploiting the specific structure of $\mathbf{\Theta}$. The structure for $\mathbf{\Theta}$ has to be chosen according to the particular computational architecture.

It is important to note that the solution of (1.11) with normal equation (1.12) can suffer from round-off errors. Indeed, the condition number of the normal matrix is equal to the square of the condition number of $\mathbf{R}_X^{1/2}\mathbf{V}$, which can lead to a dramatic numerical instability when $\mathbf{V}$ is ill-conditioned. On the other hand, the sketched least-squares problem (1.14) can be efficiently solved directly with a standard routine such as QR factorization or SVD, without forming the normal equation. The matrix $\mathbf{\Theta}$ is typically chosen such that the condition number of $\mathbf{\Theta}\mathbf{V}$ is at most (up to a small factor) the condition number of $\mathbf{R}_X^{1/2}\mathbf{V}$. Consequently, the factorization of

$\mathbf{\Theta V}$ can be much less sensitive to round-off errors than the solution of the normal equation (1.12).

A guarantee of uniqueness and quasi-optimality of the solution in (1.14) can be obtained by requiring $\mathbf{\Theta}$ to satisfy a $X \to \ell_2$ ε-subspace embedding property for $W := \mathrm{range}(\mathbf{V}) + \mathrm{span}(\mathbf{b})$, which states that for any two vectors $\mathbf{x}, \mathbf{y} \in W$,

$$|\langle \mathbf{x}, \mathbf{y} \rangle_X - \langle \mathbf{\Theta x}, \mathbf{\Theta y} \rangle| \leq \varepsilon \|\mathbf{x}\|_X \|\mathbf{y}\|_X.$$

Next, the question of the construction of ε-embeddings for the subspace of interest W is addressed. Matrices $\mathbf{\Theta}$ satisfying the ε-embedding property for W can be constructed with RLA. Such probabilistic techniques have a user-specified probability of failure δ. The computational cost (mainly associated with the size of $\mathbf{\Theta}$) depends only logarithmically on the probability of failure, therefore allowing prescription of very small values for δ (say, $\delta = 10^{-10}$) without considerable impact on the overall computational costs.

Let $\mathbf{W}$ be a matrix whose column vectors form a basis for W. One way is to choose $\mathbf{\Theta}$ according to the leverage scores sampling of rows of $\mathbf{W}$. The leverage scores sampling, however, has very limited applicability to parameter-dependent problems or other problems with varying or unknown a priori subspaces of interest W. For these scenarios, the distribution of sketching matrices has to be chosen such that $\mathbf{\Theta}$ is an ε-embedding for W with high probability for any low-dimensional subspace W. This approach is referred to as an oblivious construction of $\mathbf{\Theta}$ since it does not require any a priori knowledge about W. An ε-embedding for W with a user-specified high probability can be obtained as a realization of an oblivious subspace embedding with sufficiently large number of rows. The number of rows for $\mathbf{\Theta}$ can be chosen using theoretical a priori bounds from [12] or adaptively with the procedure from [13].

Oblivious $\ell_2 \to \ell_2$ subspace embeddings (defined by taking $X = \ell_2$) include the rescaled Gaussian distribution, the rescaled Rademacher distribution, the Subsampled Randomized Hadamard Transform (SRHT), the Subsampled Randomized Fourier Transform (SRFT), CountSketch matrix, SRFT combined with sequences of random Givens rotations, and others [12, 88, 131, 157]. In this manuscript we shall only rely on the Gaussian distribution, the Rademacher distribution and SRHT.

A rescaled Gaussian matrix in $\mathbb{K}^{k \times n}$ has entries that are independent Gaussian random variables with mean 0 and variance k^{-1}. This is the most common choice for random sketching algorithms. The computational cost reduction with these matrices is usually attained due to an exploitation of the computational architecture, which can be characterized by the efficiency of the data flow or its maintenance, or by the usage of high-level linear algebra routines. For instance the products of Gaussian matrices with vectors are embarrassingly parallelizable.

For the rescaled Rademacher distribution over $\mathbb{K}^{k \times n}$, the entries of the random matrix are independent random variables that are equal to $k^{-1/2}$ or $-k^{-1/2}$ with probabilities 1/2. It was shown in [2] that a Rademacher matrix has the same guarantees for performance as a Gaussian matrix, yet it can be more efficiently implemented using standard SQL primitives.

Assuming that n is a power of 2, the SRHT distribution is defined as $k^{-1/2}\mathbf{R}\mathbf{H}_n\mathbf{D}$, where

- $\mathbf{R} \in \mathbb{K}^{k\times n}$ is a sampling matrix defined as the first k rows of a random permutation of rows of an identity matrix,
- $\mathbf{H}_n \in \mathbb{R}^{n\times n}$ is a Walsh-Hadamard matrix, which is a structured matrix defined recursively by $\mathbf{H}_n = \mathbf{H}_{n/2} \otimes \mathbf{H}_2$, with
$$\mathbf{H}_2 := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$
Note that a product of $\mathbf{H}_n$ with a vector can be computed with $n\log_2(n)$ flops by using the fast Walsh-Hadamard transform,
- $\mathbf{D} \in \mathbb{R}^{n\times n}$ is a random diagonal matrix with random entries such that $\mathbb{P}([\mathbf{D}]_{i,i} = \pm 1) = 1/2$.

The algorithms designed with random sketching using SRHT can largely outperform the deterministic techniques in terms of the classical metric of efficiency, which is the number of flops, and are particularly interesting from the theoretical point of view. In practice the usage of SRHT is the most beneficial in basic computational environments such as computations on portable devices. The partial SRHT (P-SRHT) is used when n is not a power of 2, and is defined as the first n columns of a SRHT matrix of size s, were s is the power of 2 such that $n \leq s < 2n$.

An oblivious $X \to \ell_2$ subspace embedding for a general inner product $\langle\cdot,\cdot\rangle_X$ can be constructed as

$$\mathbf{\Theta} = \mathbf{\Omega}\mathbf{Q}, \tag{1.15}$$

where $\mathbf{\Omega}$ is an oblivious $\ell_2 \to \ell_2$ subspace embedding and matrix $\mathbf{Q}$ is such that $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{R}_X$. One way to compute the matrix $\mathbf{Q}$ is to employ the (sparse) Cholesky decomposition. As was discussed in [12], $\mathbf{R}_X$ can have a block structure that can be exploited for more efficient computation of $\mathbf{Q}$. Typically, the multiplication of $\mathbf{Q}$ by a vector is expected to have log-linear complexity, i.e., $\mathcal{O}(n\log(n)^{\nu})$ for some small $\nu \geq 1$. In this case, if $\mathbf{\Omega}$ is a P-SRHT matrix, the sketched matrices and vectors in the least-squares problem (1.14) shall require only $\mathcal{O}(nr\log(n)^{\nu})$ flops for their computation, which can be considerably less than the complexity $\mathcal{O}(nr^2)$ required for assembling the normal equation (1.13). Moreover, using a seeded random number generator allows computation of the sketched matrices and vectors with only one pass over $\mathbf{V}$ and $\mathbf{b}$ (compared to 2 passes required by the normal equation), which can be crucial if these quantities may not be efficiently stored.

From [12, 157] it follows that the oblivious $X \to \ell_2$ subspace embedding constructed with (1.15) using for $\mathbf{\Omega}$ a Gaussian or Rademacher matrix with $k = \mathcal{O}(\varepsilon^{-2}(\dim(W) + \log(1/\delta))$ rows, is an ε-subspace embedding for W with probability at least $1-\delta$. The P-SRHT has worse theoretical bounds, namely $k = \mathcal{O}(\varepsilon^{-2}(\dim(W) + \log(n/\delta))\log(\dim(W)/\delta))$, for the sufficient number of rows to satisfy the ε-subspace embedding property for W. Yet, in our experiments it has been revealed that Gaussian matrices, Rademacher matrices and P-SRHT show in practice the same performance.

1.2.2 Randomized linear algebra for MOR

The techniques based on RLA started to be used in the MOR community only recently. Perhaps one of the earliest works considering RLA in the context of MOR is [160], where the authors proposed to use RLA for interpolation of the matrix inverse for constructing parameter-dependent preconditioners. More specifically, they constructed a preconditioner of the form

$$\mathbf{P}(\mu) = \sum_{i=1}^{d} \lambda_i(\mu)\mathbf{A}(\mu_i)^{-1},$$

with the coefficients $\lambda_i(\mu)$ obtained for each parameter value by (approximate) minimization of the Frobenius norm of the residual matrix $\mathbf{I}-\mathbf{P}(\mu)\mathbf{A}(\mu)$, which requires solution of a small least-squares problem and can be efficiently performed online. The feasibility of the heavy offline computations was attained by estimation of $\|\mathbf{I}-\mathbf{P}(\mu)\mathbf{A}(\mu)\|_F$ via $\|(\mathbf{I}-\mathbf{P}(\mu)\mathbf{A}(\mu))\mathbf{\Omega}^{\mathrm{H}}\|_F$, where $\mathbf{\Omega}$ is a small rescaled Rademacher or P-SRHT matrix. It was then shown that the minimization of the sketched norm over $\lambda_i(\mu)$ yields a quasi-optimal solution with high probability. This principle is taken as the starting point of Chapter 4 of this manuscript and is improved in several ways (see Section 1.4 for more details).

In [141] the authors developed a probabilistic error estimator based on RLA. They proposed to estimate the error $\|\mathbf{u}(\mu)-\mathbf{u}_r(\mu)\|_U$ of the approximate solution $\mathbf{u}_r(\mu) \in U$ by

$$\|\mathbf{Y}(\mu)^{\mathrm{H}}\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|,$$

where $\mathbf{Y}(\mu) = [\mathbf{y}_1(\mu), \mathbf{y}_2(\mu), \ldots, \mathbf{y}_k(\mu)]$ is a matrix whose column vectors are approximate solutions to the dual problems with random right-hand-sides

$$\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{y}_i(\mu) = \mathbf{z}_i, \quad 1 \le i \le k.$$

The random right-hand-side vectors can be generated as $\mathbf{z}_i = \mathbf{Q}\boldsymbol{\omega}_i$, where $\mathbf{Q}$ is a matrix such that $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{R}_U$ and $\boldsymbol{\omega}_i$ is a standard (rescaled) Gaussian vector. Such an error estimation is closely related to the preconditioned residual-based error estimation presented in Chapter 4. The difference is that in [141], the dual problems are tackled separately (with projection-based MOR), while in Chapter 4 we obtain the whole solution matrix $\mathbf{Y}(\mu)$ by considering a single equation

$$\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{Y}(\mu) = \mathbf{Z},$$

where $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k]$, with a solution obtained by interpolation of the matrix inverse:

$$\mathbf{Y}(\mu) \approx \sum_{i=1}^{p} \lambda_i(\mu)\mathbf{A}(\mu^i)^{-\mathrm{H}}\mathbf{Z}.$$

Note that our approach has several important advantages over the one from [141] as is discussed in Chapter 4.

The RLA has also been employed for reduction of the computational cost of MOR in [5, 91], where the authors considered random sketching only as a tool for efficient evaluation of low-rank approximations of large matrices (using randomized versions of SVDs), which in principle could be done with any other efficient algorithm. In [37] a probabilistic range finder based on random sketching has been used for combining RB method with domain decomposition. Probabilistic methods from compressed sensing were applied to numerical approximation of partial differential equations in [34].

Throughout the thesis we employ the random sketching technique for drastic improvement of the efficiency (and numerical stability) of projection-based MOR based on Galerkin or minimal residual methods. This is done by expressing MOR methodology in a non-conventional (semi-)discrete form well-suited for combination with the random sketching technique. The central ingredient of our approach is the approximation of the original inner products $\langle\cdot,\cdot\rangle_U$ and $\langle\cdot,\cdot\rangle_{U'}$ by their efficient sketched versions, respectively, defined as

$$\langle\cdot,\cdot\rangle_U^{\mathbf{\Theta}} := \langle\mathbf{\Theta}\cdot,\mathbf{\Theta}\cdot\rangle, \text{ and } \langle\cdot,\cdot\rangle_{U'}^{\mathbf{\Theta}} := \langle\mathbf{\Theta}\mathbf{R}_U^{-1}\cdot,\mathbf{\Theta}\mathbf{R}_U^{-1}\cdot\rangle, \tag{1.16}$$

where $\mathbf{\Theta}$ is an oblivious $U \to \ell_2$ subspace embedding, which yields efficient and accurate approximations of the projection, residual-based error estimation and primal-dual correction. Instead of operating in a high-dimensional space, we embed the model into a low-dimensional space by using the projection $\mathbf{\Theta}$ and then construct the reduced model there. We use the fact that the construction of the reduced model requires operation with vectors lying only in some subspaces $W \subset U$ and $W' \subset U'$ of moderate dimensions. Letting $\{\mathbf{w}_i\}$ and $\{\mathbf{w}'_i\}$ be bases of W and W', respectively, the sketched inner products between any two vectors in W or W' can be efficiently computed from small projections $\{\mathbf{\Theta}\mathbf{w}'_i\}$ and $\{\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{w}'_i\}$ that compose *a sketch of a reduced model*. It follows that the reduced model can be constructed from its sketch with a negligible computational cost, while the sketch can be efficiently computed in any computational environment by using suitable random matrices.

As discussed in [12, 13], random sketching does not only provide efficient approximation of a reduced model, but also can be used to improve numerical stability of the estimation (or minimization for minres projection) of the residual norm. This is attained due to direct computation of the (sketched) residual norm and not its square (with the normal matrix) as in classical methods.

1.3 Computational architectures

Development of effective numerical algorithms requires clear characterization of the factors which define the *computational costs*.

Classical computational environment

The computational cost of an algorithm can be classically measured with the number of floating point operations (flops). This metric of efficiency is relevant for not very large problems and is traditional for the theoretical analysis of the computational cost. The pure sequential computational environment can be rarely found in modern applications. Usually it is mixed with other environments implying the necessity to consider other metrics of efficiency. For modern problems the number of flops started to become less relevant, with memory consumption and communication costs emerging as the central factors for efficiency.

Limited-memory and streaming computational environment

Over the years the out-of-core computations became drastically slower than the operations on data stored in the fast memory (RAM) with the increased latency of hard disks. Since the typical computational setup has a limited RAM, the RAM consumption can be the major constraint for the development of numerical methods. If the method entails operations with large data sets which do not fit into RAM, the efficiency of an algorithm shall be mainly affected by the number of passes over the data. In this case the pass-efficiency of an algorithm has to be of the primary concern. In some scenarios one can work in a streaming environment where each entry of the data can be accessed only once, that requires usage of single-pass algorithms.

Parallel computing

The Moore's law suggests that the transistor count in an integrated circuit doubles every two years, yielding exponential improvement of the computational capabilities of the hardware with time. In the last years keeping up the Moore's pace required changing to heterogeneous architectures. Nowadays, multi-core chips are used in every computing device. Furthermore, practically every contemporary engineer and scientist has an access to parallel computing clusters allowing performance of very heavy computations. The introduction of heterogeneous computational environments resulted in the need of corresponding changes in the design of numerical algorithms.

The computational cost of most modern algorithms is measured by the efficiency of the data flow between cores.

Distributed computing

Computations using a network of workstations (with a slow communication between each other) is a quite standard situation. For such a scenario the overall runtime of an algorithm shall be mainly dominated by the amount of data exchanged between the computers, and one has to design the algorithm based on this factor.

Other computational environments

The performance of an algorithm can be characterized not only by the runtime but also by the required computational resources. Indeed, when the algorithms are executed on servers with limited or on a pay-as-go basis budgets, one should use as few resources from the server as possible. This can be done by performing most computations beyond the server and appeal to the server only when necessary. In this situation, with the reduction of computations on the server, one also has to avoid transfers of large data sets that can become a bottleneck.

The adaptation of classical projection-based MOR methods to modern computational architectures was addressed in [5, 33, 90, 91, 99, 122]. In these works the authors did not propose a new methodology but rather exploited the key opportunities for the computational cost reduction, by direct appealing to the particular architecture or by using modern numerical algorithms for low-rank approximation of a large-scale matrix.

The algorithms designed with RLA are often universal and can be easily adapted to practically any computational architecture [32, 51, 88, 153, 157, 159]. The methodology proposed in the present thesis is not an exception. The (offline) computations involved in our algorithms mainly consist in evaluating random matrix-vector products and solving systems of equations. The former operation is known to be cheap in any computational environment (with a good choice of random matrices), while the latter one is very well-studied and can be efficiently performed with state-of-the-art routines. Furthermore, the exceptional feature of our methodology is no maintenance and operation with large-scale vectors but only with their small sketches. Since our algorithms do not require transfers of large data sets, they are particularly well-suited for computations with a network of workstations or a server with limited budget.

1.4 Contributions and overview of the manuscript

In this section we summarize the main contributions of the thesis and present an overview of the manuscript. Each chapter following the introductory chapter is presented in a form of an article and can be considered separately from the rest of the thesis.

Galerkin methods and error estimation

Chapter 2 presents a methodology based on random sketching technique for drastic reduction of the computational cost of classical projection-based MOR methods. We introduce the concept of a sketch of a model and propose new and efficient randomized sketched versions of Galerkin projection, residual-based error estimation, and primal-dual correction, with the precise conditions on the dimension of the random sketch for the resulting reduced order model to be quasi-optimal with high probability. We then present and discuss randomized sketched greedy algorithm and POD for the efficient generation of reduced approximation spaces. The methodology is experimentally validated on two benchmarks.

The proposed approach can be beneficial in basically any classical or modern computational environment. It can reduce both complexity and memory requirements. Furthermore, the reduced order model can be constructed under extreme memory constraints. All major operations in our algorithms, except solving linear systems of equations, are embarrassingly parallel. Our version of POD can be computed on multiple computational devices with the total communication cost independent of the dimension of the full order model.

Minimal residual methods and dictionary-based approximation

In Chapter 3 we introduce a sketched version of the minimal residual projection as well as a novel nonlinear approximation method, where for each parameter value, the solution is approximated by minimal residual projection onto a low-dimensional space with a basis (adaptively) selected from a dictionary of vectors. It is shown that in addition to enhancement of efficiency, random sketching technique can also offer improvement of numerical stability. We provide the conditions on the random sketch to obtain a given accuracy. These conditions may be ensured a priori with high probability by considering for the sketching matrix an oblivious embedding of sufficiently large size. In contrast to Galerkin methods, with minimal residual methods the quality of the sketching matrix can be characterized regardless of the operator's properties. Furthermore, a simple and reliable way for a posteriori verification of the quality of the sketch is provided. This approach can be used for certification of the approximation or for adaptive selection of an optimal size of random sketching matrices. We also propose a randomized procedure for an efficient approximation of an inner product between parameter-dependent vectors having affine decompositions with many (expensive to operate with) terms. This procedure can be used for efficient extraction of the quantity of interest and the primal-dual correction from the reduced model's solution.

Parameter-dependent preconditioners for model order reduction

In Chapter 4 we develop an effective methodology for the construction of a parameter-dependent preconditioner. Here only theoretical results are discussed without numerical validation, which is left for the future. The starting point of our approach is [160], which is improved in several ways mainly thanks to the framework from Chapter 2. In addition, we here propose some fundamentally novel ideas.

The preconditioner can be constructed by interpolation of matrix inverse based on minimization of an error indicator measuring a discrepancy between the preconditioned operator and the identity. In [160] the authors considered a single general error indicator for any context. However, obtaining a good quality of a preconditioner in terms of this error indicator may be an intractable task. Consequently we propose three different strategies for choosing the error indicator. The most pertinent strategy should be chosen depending on the particular scenario: multi-purpose context, when one is interested in minimizing the condition number; residual-based error estimation and certification; and Petrov-Galerkin projection onto a given approximation space.

For the multi-purpose context, the quality of the preconditioner is characterized with respect to a general norm represented by a self-adjoint positive define matrix instead of the ℓ_2-norm as in [160]. This is important, for instance, in the context of numerical methods for PDEs to control the quality of an approximation regardless the used discretization.

All proposed error indicators are online-efficient. For feasibility of the offline stage a (semi-)norm of a large-scale matrix is approximated by the ℓ_2-norm of its random sketch. The computational cost is considerably improved compared to the algorithms in [160] by using a three-phase sketching scheme instead of a sketching with a single random matrix. Thanks to the framework introduced in Chapter 2 we derive better theoretical bounds for the size of a sketch for the quasi-optimality of minimization of an error indicator, compared to the bounds derived in [160]. Moreover, we provide rigorous conditions for the quasi-optimality of the preconditioned Galerkin projection and residual-based error estimation based on the error indicators.

Chapter 2

Random sketching for Galerkin methods and error estimation

This chapter is based on the article [12] accepted for publication in journal “Advances in Computational Mathematics”. Here we propose a probabilistic way for reducing the cost of classical projection-based model order reduction methods for parameter-dependent linear equations. A reduced order model is here approximated from its random sketch, which is a set of low-dimensional random projections of the reduced approximation space and the spaces of associated residuals. This approach exploits the fact that the residuals associated with approximations in low-dimensional spaces are also contained in low-dimensional spaces. We provide conditions on the dimension of the random sketch for the resulting reduced order model to be quasi-optimal with high probability. Our approach can be used for reducing both complexity and memory requirements. The provided algorithms are well suited for any modern computational environment. Major operations, except solving linear systems of equations, are embarrassingly parallel. Our version of proper orthogonal decomposition can be computed on multiple workstations with a communication cost independent of the dimension of the full order model. The reduced order model can even be constructed in a so-called streaming environment, i.e., under extreme memory constraints. In addition, we provide an efficient way for estimating the error of the reduced order model, which is not only more efficient than the classical approach but is also less sensitive to round-off errors. Finally, the methodology is validated on benchmark problems.

Contents

2.1 Introduction

Projection-based model order reduction (MOR) methods, including the reduced basis (RB) method or proper orthogonal decomposition (POD), are popular approaches for approximating large-scale parameter-dependent equations (see the recent surveys and monographs [24, 25, 89, 128]). They can be considered in the contexts of optimization, uncertainty quantification, inverse problems, real-time simulations, etc. An essential feature of MOR methods is offline/online splitting of the computations. The construction of the reduced order (or surrogate) model, which is usually the most computationally demanding task, is performed during the offline stage. This stage consists of (i) the generation of a reduced approximation space with a greedy algorithm for RB method or a principal component analysis of a set of samples of the solution for POD and (ii) the efficient representation of the reduced system of equations, usually obtained through (Petrov-)Galerkin projection, and of all the quantities needed for evaluating output quantities of interest and error estimators. In the online stage, the reduced order model is evaluated for each value of the parameter and provides prediction of the output quantity of interest with a small computational cost, which is independent of the dimension of the initial system of equations.

In this chapter, we address the reduction of computational costs for both offline and online stages of projection-based model order reduction methods by adapting random sketching methods [2, 139] to the context of RB and POD. These methods were proven capable of significant complexity reduction for basic problems in numerical linear algebra such as computing products or factorizations of matrices [88, 157]. We show how a reduced order model can be approximated from a small set, called a sketch, of efficiently computable random projections of the reduced basis vectors and the vectors involved in the affine expansion[1] of the residual, which is assumed to contain a small number of terms. Standard algebraic operations are performed on the sketch, which avoids heavy operations on large-scale matrices and vectors. Sufficient conditions on the dimension of the sketch for quasi-optimality of approximation of the reduced order model can be obtained by exploiting the fact that the residuals associated with reduced approximation spaces are contained in low-dimensional spaces. Clearly, the randomization inevitably implies a probability of failure. This probability, however, is a user-specified parameter that can be chosen extremely small without affecting considerably the computational costs. Even though this chapter is concerned only with linear equations, similar considerations should also apply to a wide range of nonlinear problems.

Note that deterministic techniques have also been proposed for adapting POD

[1] A parameter-dependent quantity $\mathbf{v}(\mu)$ with values in vector space V over a field $\mathbb{K}$ is said to admit an affine representation (or be parameter-separable) if $\mathbf{v}(\mu) = \sum_{i=1}^{d} \mathbf{v}_i \lambda_i(\mu)$ with $\lambda_i(\mu) \in \mathbb{K}$ and $\mathbf{v}_i \in V$. Note that for V of finite dimension, $\mathbf{v}(\mu)$ always admits an affine representation with a finite number of terms.

methods to modern (e.g., multi-core or limited-memory) computational architectures [33, 90, 122]. Compared to the aforementioned deterministic approaches, our randomized version of POD (see Section 2.5.2) has the advantage of not requiring the computation of the full reduced basis vectors, but only of their small random sketches. In fact, maintaining and operating with large vectors can be completely avoided. This remarkable feature makes our algorithms particularly well suited for distributed computing and streaming contexts.

Randomized linear algebra has been employed for reducing the computational cost of MOR in [5, 91], where the authors considered random sketching only as a tool for efficient evaluation of low-rank approximations of large matrices (using randomized versions of SVDs). They, however, did not adapt the MOR methodology itself and therefore did not fully exploit randomization techniques. In [37] a probabilistic range finder based on random sketching has been used for combining the RB method with domain decomposition. Randomized linear algebra was also used for building parameter-dependent preconditioners for projection-based MOR in [160].

The rest of the chapter is organized as follows. Section 2.1.1 presents the main contributions and discusses the benefits of the proposed methodology. In Section 2.2 we introduce the problem of interest and present the ingredients of standard projection-based model order reduction methods. In Section 2.3, we extend the classical sketching technique in Euclidean spaces to a more general framework. In Section 2.4, we introduce the concept of a *sketch of a model* and propose new and efficient randomized versions of Galerkin projection, residual-based error estimation, and primal-dual correction. In Section 2.5, we present and discuss the randomized greedy algorithm and POD for the efficient generation of reduced approximation spaces. In Section 2.6, the methodology is validated on two benchmarks. Finally, in Section 2.7, we provide conclusions and perspectives.

Proofs of propositions and theorems are provided in the Appendix.

2.1.1 Main contributions

Our methodology can be used for the efficient construction of a reduced order model. In classical projection-based methods, the cost of evaluating samples (or snapshots) of the solution for a training set of parameter values can be much smaller than the cost of other computations. This is the case when the samples are computed using a sophisticated method for solving linear systems of equations requiring log-linear complexity, or beyond the main routine, e.g., using a highly optimized commercial solvers or a server with limited budget, and possibly obtained using multiple workstations. This is also the case when, due to memory constraints, the computational time of algorithms for constructing the reduced order model are greatly affected by the number of passes taken over the data. In all these cases the cost of the offline stage is dominated by the post-processing of samples but not their computation. We here assume that the cost of solving high-dimensional systems is irreducible and focus on the reduction of other computational costs. The metric for efficiency depends on the computational environment and how data is presented to us. Our algorithms can be beneficial in basically all computational environments.

Complexity reduction

Consider a parameter-dependent linear system of equations $\mathbf{A}(\mu)\mathbf{u}(\mu) = \mathbf{b}(\mu)$ of dimension n and assume that the parameter-dependent matrix $\mathbf{A}(\mu)$ and vector $\mathbf{b}(\mu)$ are parameter-separable with m_A and m_b terms, respectively (see Section 2.2 for more details). Let $r \ll n$ be the dimension of the reduced approximation space. Given a basis of this space, the classical construction of a reduced order model requires the evaluation of inner products between high-dimensional vectors. More precisely, it consists in multiplying each of the $rm_A + m_b$ vectors in the affine expansion of the residual by r vectors for constructing the reduced systems and by $rm_A + m_b$ other vectors for estimating the error. These two operations result in $\mathcal{O}(nr^2m_A + nrm_b)$ and $\mathcal{O}(nr^2m_A^2 + nm_b^2)$ flops respectively. It can be argued that the aforementioned complexities can dominate the total complexity of the offline stage (see Section 2.4.4). With the methodology presented in this work the complexities can be reduced to $\mathcal{O}(nrm_A \log k + nm_b \log k)$, where $r \leq k \ll n$.

Let m be the number of samples in the training set. The computation of the POD basis using a direct eigenvalue solver requires multiplication of two $n \times m$ matrices, i.e., $\mathcal{O}(nm \min(n,m))$ flops, while using a Krylov solver it requires multiplications of a $n \times m$ matrix by $\mathcal{O}(r)$ adaptively chosen vectors, i.e., $\mathcal{O}(nmr)$ flops. In the prior work [5] on randomized algorithms for MOR, the authors proposed to use a randomized version of SVD introduced in [88] for the computation of the POD basis. More precisely, the SVD can be performed by applying Algorithms 4.5 and 5.1 in [88] with complexities $\mathcal{O}(nm \log k + nk^2)$ and $\mathcal{O}(nmk)$, respectively. However, the authors in [5] did not take any further advantage of random sketching methods, besides the SVD, and did not provide any theoretical analysis. In addition, they considered the Euclidean norm for the basis construction, which can be far from optimal. Here we reformulate the classical POD and obtain an algebraic form (see Proposition 2.2.5) well suited for the application of efficient low-rank approximation algorithms, e.g., randomized or incremental SVDs [11]. We consider a general inner product associated with a self-adjoint positive definite matrix. More importantly, we provide a new version of POD (see Section 2.5.2) which does not require evaluation of high-dimensional basis vectors. In this way, the complexity of POD can be reduced to only $\mathcal{O}(nm \log k)$.

Restricted memory and streaming environments

Consider an environment where the memory consumption is the primary constraint. The classical offline stage involves evaluations of inner products of high-dimensional vectors. These operations require many passes over large data sets, e.g., a set of samples of the solution or the reduced basis, and can result in a computational burden. We show how to build the reduced order model with only one pass over the data. In extreme cases our algorithms may be employed in a streaming environment, where samples of the solution are provided as data-streams and storage of only a few large vectors is allowed. Moreover, with our methodology one can build a reduced order model without storing any high-dimensional vector.

Distributed computing

The computations involved in our version of POD can be efficiently distributed among multiple workstations. Each sample of the solution can be evaluated and processed on a different machine with absolutely no communication. Thereafter, small sketches of the samples can be sent to the master workstation for building the reduced order model. The total amount of communication required by our algorithm is proportional to k (the dimension of the sketch) and is independent of the dimension of the initial full order model.

Parallel computing

Recently, parallelization was considered as a workaround to address large-scale computations [99]. The authors did not propose a new methodology but rather exploited the key opportunities for parallelization in a standard approach. We, on the other hand, propose a new methodology which can be better suited for parallelization than the classical one. The computations involved in our algorithms mainly consist in evaluating random matrix-vector products and solving high-dimensional systems of equations. The former operation is embarrassingly parallel (with a good choice of random matrices), while the latter one can be efficiently parallelized with state-of-the-art algorithms.

Online-efficient and robust error estimation

In addition, we provide a new way for estimating the error associated with a solution of the reduced order model, the error being defined as some norm of the residual. It does not require any assumption on the way to obtain the approximate solution and can be employed separately from the rest of the methodology. For example, it could be used for the efficient estimation of the error associated with a classical Galerkin projection. Our approach yields cost reduction for the offline stage but it is also online-efficient. Given the solution of the reduced order model, it requires only $\mathcal{O}(rm_A+m_b)$ flops for estimating the residual-based error while a classical procedure takes $\mathcal{O}(r^2m_A^2+m_b^2)$ flops. Moreover, compared to the classical approach, our method is less sensitive to round-off errors.

2.2 Projection-based model order reduction methods

In this section, we introduce the problem of interest and present the basic ingredients of classical MOR algorithms in a form well suited for random sketching methods. We consider a discrete setting, e.g, a problem arising after discretization of a parameter-dependent PDE or integral equation. We use notations that are standard in the context of variational methods for PDEs. However, for models simply described by algebraic equations, the notions of solution spaces, dual spaces, etc., can be disregarded.

Let $U := \mathbb{K}^n$ (with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) denote the solution space equipped with inner product $\langle\cdot,\cdot\rangle_U := \langle\mathbf{R}_U\cdot,\cdot\rangle$, where $\langle\cdot,\cdot\rangle$ is the canonical inner product on $\mathbb{K}^n$ and $\mathbf{R}_U \in \mathbb{K}^{n\times n}$ is some self-adjoint (symmetric if $\mathbb{K} = \mathbb{R}$ and Hermitian if $\mathbb{K} = \mathbb{C}$) positive definite matrix. The dual space of U is identified with $U' := \mathbb{K}^n$, which is endowed with inner product $\langle\cdot,\cdot\rangle_{U'} := \langle\cdot,\mathbf{R}_U^{-1}\cdot\rangle$. For a matrix $\mathbf{M} \in \mathbb{K}^{n\times n}$ we denote by $\mathbf{M}^{\mathrm{H}}$ its adjoint (transpose if $\mathbb{K} = \mathbb{R}$ and Hermitian transpose if $\mathbb{K} = \mathbb{C}$).

Remark 2.2.1. *The matrix $\mathbf{R}_U$ is seen as a map from U to U'. In the framework of numerical methods for PDEs, the entries of $\mathbf{R}_U$ can be obtained by evaluating inner products of corresponding basis functions. For example, if the PDE is defined on a space equipped with H^1 inner product, then $\mathbf{R}_U$ is equal to the stiffness (discrete Laplacian) matrix. For algebraic parameter-dependent equations, $\mathbf{R}_U$ can be taken as identity.*

Let μ denote parameters taking values in a set $\mathcal{P}$ (which is typically a subset of $\mathbb{K}^p$, but could also be a subset of function spaces, etc.). Let parameter-dependent linear forms $\mathbf{b}(\mu) \in U'$ and $\mathbf{l}(\mu) \in U'$ represent the right-hand side and the extractor of a quantity of interest, respectively, and let $\mathbf{A}(\mu) : U \to U'$ represent the parameter-dependent operator. The problem of interest can be formulated as follows: for each given $\mu \in \mathcal{P}$ find the quantity of interest $s(\mu) := \langle\mathbf{l}(\mu),\mathbf{u}(\mu)\rangle$, where $\mathbf{u}(\mu) \in U$ is such that

$$\mathbf{A}(\mu)\mathbf{u}(\mu) = \mathbf{b}(\mu). \tag{2.1}$$

Further, we suppose that the solution manifold $\{\mathbf{u}(\mu) : \mu \in \mathcal{P}\}$ can be well approximated by some low dimensional subspace of U. Let $U_r \subseteq U$ be such a subspace and $\mathbf{U}_r \in \mathbb{K}^{n\times r}$ be a matrix whose column vectors form a basis for U_r. The question of finding a good U_r is addressed in Sections 2.2.4 and 2.2.4. In projection-based MOR methods, $\mathbf{u}(\mu)$ is approximated by a projection $\mathbf{u}_r(\mu) \in U_r$.

2.2.1 Galerkin projection

Usually, a Galerkin projection $\mathbf{u}_r(\mu)$ is obtained by imposing the following orthogonality condition to the residual [128]:

$$\langle \mathbf{r}(\mathbf{u}_r(\mu);\mu), \mathbf{w}\rangle = 0, \ \forall \mathbf{w} \in U_r, \tag{2.2}$$

where $\mathbf{r}(\mathbf{x};\mu) := \mathbf{b}(\mu) - \mathbf{A}(\mu)\mathbf{x}$, $\mathbf{x} \in U$. This condition can be expressed in a different form that will be particularly handy in further sections. For this we define the following semi-norm over U':

$$\|\mathbf{y}\|_{U_r'} := \max_{\mathbf{w}\in U_r\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{y},\mathbf{w}\rangle|}{\|\mathbf{w}\|_U}, \ \mathbf{y} \in U'. \tag{2.3}$$

Note that replacing U_r by U in definition (2.3) yields a norm consistent with the one induced by $\langle\cdot,\cdot\rangle_{U'}$. The relation (2.2) can now be rewritten as

$$\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U_r'} = 0. \tag{2.4}$$

Let us define the following parameter-dependent constants characterizing quasi-optimality of Galerkin projection:

$$\alpha_r(\mu) := \min_{\mathbf{x}\in U_r\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'}}{\|\mathbf{x}\|_U}, \tag{2.5a}$$

$$\beta_r(\mu) := \max_{\mathbf{x}\in(\mathrm{span}\{\mathbf{u}(\mu)\}+U_r)\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'}}{\|\mathbf{x}\|_U}. \tag{2.5b}$$

It has to be mentioned that $\alpha_r(\mu)$ and $\beta_r(\mu)$ can be bounded by the coercivity constant $\theta(\mu)$ and the continuity constant (the maximal singular value) $\beta(\mu)$ of $\mathbf{A}(\mu)$, respectively defined by

$$\theta(\mu) := \min_{\mathbf{x}\in U\backslash\{\mathbf{0}\}} \frac{\langle \mathbf{A}(\mu)\mathbf{x},\mathbf{x}\rangle}{\|\mathbf{x}\|_U^2} \leq \alpha_r(\mu), \tag{2.6a}$$

$$\beta(\mu) := \max_{\mathbf{x}\in U\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}}{\|\mathbf{x}\|_U} \geq \beta_r(\mu). \tag{2.6b}$$

For some problems it is possible to provide lower and upper bounds for $\theta(\mu)$ and $\beta(\mu)$ [82].

If $\alpha_r(\mu)$ is positive, then the reduced problem (2.2) is well-posed. For given $V \subseteq U$, let $\mathbf{P}_V : U \to V$ denote the orthogonal projection on V with respect to $\|\cdot\|_U$, i.e.,

$$\forall \mathbf{x} \in U, \ \mathbf{P}_V\mathbf{x} = \arg\min_{\mathbf{w}\in V} \|\mathbf{x}-\mathbf{w}\|_U. \tag{2.7}$$

We now provide a quasi-optimality characterization for the projection $\mathbf{u}_r(\mu)$.

Proposition 2.2.2 (modified Cea's lemma). *If* $\alpha_r(\mu) > 0$*, then the solution* $\mathbf{u}_r(\mu)$ *of (2.2) is such that*

$$\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \leq (1 + \frac{\beta_r(\mu)}{\alpha_r(\mu)}) \|\mathbf{u}(\mu) - \mathbf{P}_{U_r}\mathbf{u}(\mu)\|_U. \tag{2.8}$$

Proof. See appendix. □

Note that Proposition 2.2.2 is a slightly modified version of the classical Cea's lemma with the continuity constant $\beta(\mu)$ replaced by $\beta_r(\mu)$.

The coordinates of $\mathbf{u}_r(\mu)$ in the basis $\mathbf{U}_r$, i.e., $\mathbf{a}_r(\mu) \in \mathbb{K}^r$ such that $\mathbf{u}_r(\mu) = \mathbf{U}_r\mathbf{a}_r(\mu)$, can be found by solving the following system of equations

$$\mathbf{A}_r(\mu)\mathbf{a}_r(\mu) = \mathbf{b}_r(\mu), \tag{2.9}$$

where $\mathbf{A}_r(\mu) = \mathbf{U}_r^{\mathrm{H}}\mathbf{A}(\mu)\mathbf{U}_r \in \mathbb{K}^{r \times r}$ and $\mathbf{b}_r(\mu) = \mathbf{U}_r^{\mathrm{H}}\mathbf{b}(\mu) \in \mathbb{K}^r$. The numerical stability of (2.9) is usually obtained by orthogonalization of $\mathbf{U}_r$.

Proposition 2.2.3. *If* $\mathbf{U}_r$ *is orthogonal with respect to* $\langle\cdot,\cdot\rangle_U$*, then the condition number of* $\mathbf{A}_r(\mu)$ *is bounded by* $\frac{\beta_r(\mu)}{\alpha_r(\mu)}$.

Proof. See appendix. □

2.2.2 Error estimation

When an approximation $\mathbf{u}_r(\mu) \in U_r$ of the exact solution $\mathbf{u}(\mu)$ has been evaluated, it is important to be able to certify how close they are. The error $\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U$ can be bounded by the following error indicator

$$\Delta(\mathbf{u}_r(\mu);\mu) := \frac{\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}}{\eta(\mu)}, \tag{2.10}$$

where $\eta(\mu)$ is such that

$$\eta(\mu) \leq \min_{\mathbf{x} \in U \backslash \{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}}{\|\mathbf{x}\|_U}. \tag{2.11}$$

In its turn, the certification of the output quantity of interest $s_r(\mu) := \langle \mathbf{l}(\mu), \mathbf{u}_r(\mu)\rangle$ is provided by

$$|s(\mu) - s_r(\mu)| \leq \|\mathbf{l}(\mu)\|_{U'}\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \leq \|\mathbf{l}(\mu)\|_{U'}\Delta(\mathbf{u}_r(\mu);\mu). \tag{2.12}$$

2.2.3 Primal-dual correction

The accuracy of the output quantity obtained by the aforementioned methodology can be improved by goal-oriented correction [132] explained below. A dual problem can be formulated as follows: for each $\mu \in \mathcal{P}$, find $\mathbf{u}^{\mathrm{du}}(\mu) \in U$ such that

$$\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{u}^{\mathrm{du}}(\mu) = -\mathbf{l}(\mu). \tag{2.13}$$

The dual problem can be tackled in the same manner as the primal problem. For this we can use a Galerkin projection onto a certain r^{du}-dimensional subspace $U_r^{\mathrm{du}} \subseteq U$.

Now suppose that besides approximation $\mathbf{u}_r(\mu)$ of $\mathbf{u}(\mu)$, we also have obtained an approximation of $\mathbf{u}^{\mathrm{du}}(\mu)$ denoted by $\mathbf{u}_r^{\mathrm{du}}(\mu) \in U_r^{\mathrm{du}}$. The quantity of interest can be estimated by

$$s_r^{\mathrm{pd}}(\mu) := s_r(\mu) - \langle \mathbf{u}_r^{\mathrm{du}}(\mu), \mathbf{r}(\mathbf{u}_r(\mu);\mu)\rangle. \tag{2.14}$$

Proposition 2.2.4. *The estimation $s_r^{\mathrm{pd}}(\mu)$ of $s(\mu)$ is such that*

$$|s(\mu) - s_r^{\mathrm{pd}}(\mu)| \leq \|\mathbf{r}^{\mathrm{du}}(\mathbf{u}_r^{\mathrm{du}}(\mu);\mu)\|_{U'} \Delta(\mathbf{u}_r(\mu);\mu), \tag{2.15}$$

where $\mathbf{r}^{\mathrm{du}}(\mathbf{u}_r^{\mathrm{du}}(\mu);\mu) := -\mathbf{l}(\mu) - \mathbf{A}(\mu)^{\mathrm{H}}\mathbf{u}_r^{\mathrm{du}}(\mu)$.

Proof. See appendix. □

We observe that the error bound (2.15) of the quantity of interest is now quadratic in the residual norm in contrast to (2.12).

2.2.4 Reduced basis generation

Until now we have assumed that the reduced subspaces U_r and U_r^{du} were given. Let us briefly outline the standard procedure for the reduced basis generation with the greedy algorithm and POD. The POD is here presented in a general algebraic form, which allows a non-intrusive use of any low-rank approximation algorithm. Below we consider only the primal problem noting that similar algorithms can be used for the dual one. We also assume that a training set $\mathcal{P}_{\mathrm{train}} \subseteq \mathcal{P}$ with finite cardinality m is provided.

Greedy algorithm

The approximation subspace U_r can be constructed recursively with a (weak) greedy algorithm. At iteration i, the basis of U_i is enriched by snapshot $\mathbf{u}(\mu^{i+1})$, i.e.,

$$U_{i+1} := U_i + \mathrm{span}(\mathbf{u}(\mu^{i+1})),$$

evaluated at a parameter value μ^{i+1} that maximizes a certain error indicator $\widetilde{\Delta}(U_i;\mu)$ over the training set. Note that for efficient evaluation of $\arg\max_{\mu \in \mathcal{P}_{\mathrm{train}}} \widetilde{\Delta}(U_i;\mu)$ a provisional online solver associated with U_i has to be provided.

The error indicator $\widetilde{\Delta}(U_i;\mu)$ for the greedy selection is typically chosen as an upper bound or estimator of $\|\mathbf{u}(\mu)-\mathbf{P}_{U_i}\mathbf{u}(\mu)\|_U$. One can readily take $\widetilde{\Delta}(U_i;\mu) := \Delta(\mathbf{u}_i(\mu);\mu)$, where $\mathbf{u}_i(\mu)$ is the Galerkin projection defined by (2.4). The quasi-optimality of such $\widetilde{\Delta}(U_i,\mu)$ can then be characterized by using Proposition 2.2.2 and definitions (2.6b) and (2.10).

Proper Orthogonal Decomposition

In the context of POD we assume that the samples (snapshots) of $\mathbf{u}(\mu)$, associated with the training set, are available. Let them be denoted as $\{\mathbf{u}(\mu^i)\}_{i=1}^m$, where $\mu^i \in \mathcal{P}_{\text{train}}$, $1 \leq i \leq m$. Further, let us define $\mathbf{U}_m := \left[\mathbf{u}(\mu^1), \mathbf{u}(\mu^2), ..., \mathbf{u}(\mu^m)\right] \in \mathbb{K}^{n\times m}$ and $U_m := \text{range}(\mathbf{U}_m)$. POD aims at finding a low dimensional subspace $U_r \subseteq U_m$ for the approximation of the set of vectors $\{\mathbf{u}(\mu^i)\}_{i=1}^m$.

For each $r \leq \dim(U_m)$ we define

$$POD_r(\mathbf{U}_m, \|\cdot\|_U) := \arg \min_{\substack{W_r \subseteq U_m \\ \dim(W_r)=r}} \sum_{i=1}^m \|\mathbf{u}(\mu^i) - \mathbf{P}_{W_r}\mathbf{u}(\mu^i)\|_U^2. \tag{2.16}$$

The standard POD consists in choosing U_r as $POD_r(\mathbf{U}_m, \|\cdot\|_U)$ and using the method of snapshots [140], or SVD of matrix $\mathbf{R}_U^{1/2}\mathbf{U}_m$, for computing the basis vectors. For large-scale problems, however, performing the method of snapshots or the SVD can become a computational burden. In such a case the standard eigenvalue decomposition and SVD have to be replaced by other low-rank approximations, e.g., incremental SVD, randomized SVD, hierarchical SVD, etc. For each of them it can be important to characterize quasi-optimality of the approximate POD basis. Below we provide a generalized algebraic version of POD well suited for a combination with low-rank approximation algorithms as well as state-of-the-art SVD. Note that obtaining (e.g., using a spectral decomposition) and operating with $\mathbf{R}_U^{1/2}$ can be expensive and should be avoided for large-scale problems. The usage of this matrix for constructing the POD basis can be easily circumvented (see Remark 2.2.7).

Proposition 2.2.5. *Let $\mathbf{Q} \in \mathbb{K}^{s\times n}$ be such that $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{R}_U$. Let $\mathbf{B}_r^* \in \mathbb{K}^{s\times m}$ be a best rank-r approximation of $\mathbf{Q}\mathbf{U}_m$ with respect to the Frobenius norm $\|\cdot\|_F$. Then for any rank-r matrix $\mathbf{B}_r \in \mathbb{K}^{s\times m}$, it holds*

$$\frac{1}{m}\|\mathbf{Q}\mathbf{U}_m - \mathbf{B}_r^*\|_F^2 \leq \frac{1}{m}\sum_{i=1}^m \|\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}\mathbf{u}(\mu^i)\|_U^2 \leq \frac{1}{m}\|\mathbf{Q}\mathbf{U}_m - \mathbf{B}_r\|_F^2, \tag{2.17}$$

where $U_r := \{\mathbf{R}_U^{-1}\mathbf{Q}^{\mathrm{H}}\mathbf{b} : \mathbf{b} \in \text{span}(\mathbf{B}_r)\}$.

Proof. See appendix. □

Corollary 2.2.6. *Let $\mathbf{Q} \in \mathbb{K}^{s \times n}$ be such that $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{R}_U$. Let $\mathbf{B}_r^* \in \mathbb{K}^{s \times m}$ be a best rank-r approximation of $\mathbf{Q}\mathbf{U}_m$ with respect to the Frobenius norm $\| \cdot \|_F$. Then*

$$POD_r(\mathbf{U}_m, \| \cdot \|_U) = \{\mathbf{R}_U^{-1}\mathbf{Q}^{\mathrm{H}}\mathbf{b} : \mathbf{b} \in \mathrm{range}(\mathbf{B}_r^*)\}. \tag{2.18}$$

It follows that the approximation subspace U_r for $\{\mathbf{u}(\mu^i)\}_{i=1}^m$ can be obtained by computing a low-rank approximation of $\mathbf{Q}\mathbf{U}_m$. According to Proposition 2.2.5, for given r, quasi-optimality of U_r can be guaranteed by quasi-optimality of $\mathbf{B}_r$.

Remark 2.2.7. *The matrix $\mathbf{Q}$ in Proposition 2.2.5 and Corollary 2.2.6 can be seen as a map from U to $\mathbb{K}^s$. Clearly, it can be computed with a Cholesky (or spectral) decomposition of $\mathbf{R}_U$. For large-scale problems, however, it might be a burden to obtain, store or operate with such a matrix. We would like to underline that $\mathbf{Q}$ does not have to be a square matrix. It can be easily obtained in the framework of numerical methods for PDEs (e.g., finite elements, finite volumes, etc.). Suppose that $\mathbf{R}_U$ can be expressed as an assembly of smaller self-adjoint positive semi-definite matrices $\mathbf{R}_U^{(i)}$ each corresponding to the contribution, for example, of a finite element or subdomain. In other words,*

$$\mathbf{R}_U = \sum_{i=1}^{l} \mathbf{E}^{(i)}\mathbf{R}_U^{(i)}[\mathbf{E}^{(i)}]^{\mathrm{T}},$$

where $\mathbf{E}^{(i)}$ is an extension operator mapping a local vector to the global one (usually a boolean matrix). Since $\mathbf{R}_U^{(i)}$ are small matrices, their Cholesky (or spectral) decompositions are easy to compute. Let $\mathbf{Q}^{(i)}$ denote the adjoint of the Cholesky factor of $\mathbf{R}_U^{(i)}$. It can be easily verified that

$$\mathbf{Q} := \begin{bmatrix} \mathbf{Q}^{(1)}[\mathbf{E}^{(1)}]^{\mathrm{T}} \\ \mathbf{Q}^{(2)}[\mathbf{E}^{(2)}]^{\mathrm{T}} \\ \dots \\ \mathbf{Q}^{(l)}[\mathbf{E}^{(l)}]^{\mathrm{T}} \end{bmatrix}$$

satisfies $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{R}_U$.

The POD procedure using low-rank approximations is depicted in Algorithm 2.

2.3 Random sketching

In this section, we adapt the classical sketching theory in Euclidean spaces [157] to a slightly more general framework. The sketching technique is seen as a modification of inner product for a given subspace. The modified inner product is approximately equal to the original one but it is much easier to operate with. Thanks to such interpretation of the methodology, integration of the sketching technique to the context of projection-based MOR will become straightforward.

Algorithm 2 Approximate Proper Orthogonal Decomposition

```
Given: $\mathcal{P}_{\text{train}}$, $\mathbf{A}(\mu)$, $\mathbf{b}(\mu)$, $\mathbf{R}_U$
Output: $\mathbf{U}_r$ and $\Delta^{\text{POD}}$
1. Compute the snapshot matrix $\mathbf{U}_m$.
2. Determine $\mathbf{Q}$ such that $\mathbf{Q}^{\text{H}}\mathbf{Q} = \mathbf{R}_U$.
3. Compute a rank-$r$ approximation, $\mathbf{B}_r$, of $\mathbf{Q}\mathbf{U}_m$.
4. Compute an upper bound, $\Delta^{\text{POD}}$, of $\frac{1}{m}\|\mathbf{Q}\mathbf{U}_m - \mathbf{B}_r\|_F^2$.
5. Find a matrix, $\mathbf{C}_r$ whose column space is span$(\mathbf{B}_r)$.
6. Evaluate $\mathbf{U}_r := \mathbf{R}_U^{-1}\mathbf{Q}^{\text{H}}\mathbf{C}_r$.
```

2.3.1 ℓ_2-embeddings

Let $X := \mathbb{K}^n$ be endowed with inner product $\langle\cdot,\cdot\rangle_X := \langle \mathbf{R}_X\cdot,\cdot\rangle$ for some self-adjoint positive definite matrix $\mathbf{R}_X \in \mathbb{K}^{n\times n}$, and let Y be a subspace of X of moderate dimension. The dual of X is identified with $X' := \mathbb{K}^n$ and the dual of Y is identified with $Y' := \{\mathbf{R}_X\mathbf{y} : \mathbf{y} \in Y\}$. X' and Y' are both equipped with inner product $\langle\cdot,\cdot\rangle_{X'} := \langle\cdot,\mathbf{R}_X^{-1}\cdot\rangle$. The inner products $\langle\cdot,\cdot\rangle_X$ and $\langle\cdot,\cdot\rangle_{X'}$ can be very expensive to evaluate. The computational cost can be reduced drastically if we are interested solely in operating with vectors lying in subspaces Y or Y'. For this we introduce the concept of $X \to \ell_2$ subspace embeddings.

Let $\mathbf{\Theta} \in \mathbb{K}^{k\times n}$ with $k \leq n$. Further, $\mathbf{\Theta}$ is seen as an embedding for subspaces of X. It maps vectors from the subspaces of X to vectors from $\mathbb{K}^k$ equipped with the canonical inner product $\langle\cdot,\cdot\rangle$, so $\mathbf{\Theta}$ is referred to as an $X \to \ell_2$ subspace embedding. Let us now introduce the following semi-inner products on X and X':

$$\langle\cdot,\cdot\rangle_X^{\mathbf{\Theta}} := \langle\mathbf{\Theta}\cdot,\mathbf{\Theta}\cdot\rangle, \text{ and } \langle\cdot,\cdot\rangle_{X'}^{\mathbf{\Theta}} := \langle\mathbf{\Theta}\mathbf{R}_X^{-1}\cdot,\mathbf{\Theta}\mathbf{R}_X^{-1}\cdot\rangle. \tag{2.19}$$

Let $\|\cdot\|_X^{\mathbf{\Theta}}$ and $\|\cdot\|_{X'}^{\mathbf{\Theta}}$ denote the associated semi-norms. In general, $\mathbf{\Theta}$ is chosen so that $\langle\cdot,\cdot\rangle_X^{\mathbf{\Theta}}$ approximates well $\langle\cdot,\cdot\rangle_X$ for all vectors in Y or, in other words, $\mathbf{\Theta}$ is $X \to \ell_2$ ε-subspace embedding for Y, as defined below.

Definition 2.3.1. *If $\mathbf{\Theta}$ satisfies*

$$\forall \mathbf{x},\mathbf{y} \in Y, \ \left|\langle\mathbf{x},\mathbf{y}\rangle_X - \langle\mathbf{x},\mathbf{y}\rangle_X^{\mathbf{\Theta}}\right| \leq \varepsilon\|\mathbf{x}\|_X\|\mathbf{y}\|_X, \tag{2.20}$$

for some $\varepsilon \in [0,1)$, then it is called a $X \to \ell_2$ ε-subspace embedding (or simply, ε-embedding) for Y.

Corollary 2.3.2. *If $\mathbf{\Theta}$ is a $X \to \ell_2$ ε-subspace embedding for Y, then*

$$\forall \mathbf{x}',\mathbf{y}' \in Y', \ \left|\langle\mathbf{x}',\mathbf{y}'\rangle_{X'} - \langle\mathbf{x}',\mathbf{y}'\rangle_{X'}^{\mathbf{\Theta}}\right| \leq \varepsilon\|\mathbf{x}'\|_{X'}\|\mathbf{y}'\|_{X'}.$$

Proposition 2.3.3. *If $\mathbf{\Theta}$ is a $X \to \ell_2$ ε-subspace embedding for Y, then $\langle\cdot,\cdot\rangle_X^{\mathbf{\Theta}}$ and $\langle\cdot,\cdot\rangle_{X'}^{\mathbf{\Theta}}$ are inner products on Y and Y', respectively.*

Proof. See appendix. □

Let $Z \subseteq Y$ be a subspace of Y. A semi-norm $\|\cdot\|_{Z'}$ over Y' can be defined by

$$\|\mathbf{y}'\|_{Z'} := \max_{\mathbf{x}\in Z\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{y}', \mathbf{x}\rangle|}{\|\mathbf{x}\|_X} = \max_{\mathbf{x}\in Z\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{R}_X^{-1}\mathbf{y}', \mathbf{x}\rangle_X|}{\|\mathbf{x}\|_X},\ \mathbf{y}' \in Y'. \tag{2.21}$$

We propose to approximate $\|\cdot\|_{Z'}$ by the semi norm $\|\cdot\|^{\boldsymbol{\Theta}}_{Z'}$ given by

$$\|\mathbf{y}'\|^{\boldsymbol{\Theta}}_{Z'} := \max_{\mathbf{x}\in Z\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{R}_X^{-1}\mathbf{y}', \mathbf{x}\rangle^{\boldsymbol{\Theta}}_X|}{\|\mathbf{x}\|^{\boldsymbol{\Theta}}_X},\ \mathbf{y}' \in Y'. \tag{2.22}$$

Observe that letting $Z = Y$ in Equations (2.21) and (2.22) leads to norms on Y' which are induced by $\langle\cdot,\cdot\rangle_{X'}$ and $\langle\cdot,\cdot\rangle^{\boldsymbol{\Theta}}_{X'}$.

Proposition 2.3.4. *If $\boldsymbol{\Theta}$ is a $X \to \ell_2$ ε-subspace embedding for Y, then for all $\mathbf{y}' \in Y'$,*

$$\frac{1}{\sqrt{1+\varepsilon}}(\|\mathbf{y}'\|_{Z'} - \varepsilon\|\mathbf{y}'\|_{X'}) \leq \|\mathbf{y}'\|^{\boldsymbol{\Theta}}_{Z'} \leq \frac{1}{\sqrt{1-\varepsilon}}(\|\mathbf{y}'\|_{Z'} + \varepsilon\|\mathbf{y}'\|_{X'}). \tag{2.23}$$

Proof. See appendix. □

2.3.2 Data-oblivious embeddings

Here we show how to build a $X \to \ell_2$ ε-subspace embedding $\boldsymbol{\Theta}$ as a realization of a carefully chosen probability distribution over matrices. A reduction of the complexity of an algorithm can be obtained when $\boldsymbol{\Theta}$ is a structured matrix (e.g., sparse or hierarchical) [157] so that it can be efficiently multiplied by a vector. In such a case $\boldsymbol{\Theta}$ has to be operated implicitly with matrix-vector multiplications performed in a black-box manner. For environments where the memory consumption or the cost of communication between cores is the primary constraint, unstructured $\boldsymbol{\Theta}$ can still provide drastic reductions and be more expedient [88].

Definition 2.3.5. *$\boldsymbol{\Theta}$ is called a (ε,δ,d) oblivious $X \to \ell_2$ subspace embedding if for any d-dimensional subspace V of X it holds*

$$\mathbb{P}(\boldsymbol{\Theta} \textit{ is a } X \to \ell_2 \textit{ subspace embedding for } V) \geq 1-\delta. \tag{2.24}$$

Corollary 2.3.6. *If $\boldsymbol{\Theta}$ is a (ε,δ,d) oblivious $X \to \ell_2$ subspace embedding, then $\boldsymbol{\Theta}\mathbf{R}_X^{-1}$ is a (ε,δ,d) oblivious $X' \to \ell_2$ subspace embedding.*

The advantage of oblivious embeddings is that they do not require any a priori knowledge of the embedded subspace. In this work we shall consider three well-known oblivious $\ell_2 \to \ell_2$ subspace embeddings: the rescaled Gaussian distribution, the rescaled Rademacher distribution, and the partial Subsampled Randomized Hadamard Transform (P-SRHT). The rescaled Gaussian distribution is such that the entries of $\mathbf{\Theta}$ are independent normal random variables with mean 0 and variance k^{-1}. For the rescaled Rademacher distribution, the entries of $\mathbf{\Theta}$ are independent random variables satisfying $\mathbb{P}\left([\mathbf{\Theta}]_{i,j} = \pm k^{-1/2}\right) = 1/2$. Next we recall a standard result that states that the rescaled Gaussian and Rademacher distributions with sufficiently large k are (ε, δ, d) oblivious $\ell_2 \to \ell_2$ subspace embeddings. This can be found in [139, 157]. The authors, however, provided the bounds for k in $\mathcal{O}$ (asymptotic) notation with no concern about the constants. These bounds can be impractical for certification (both a priori and a posteriori) of the solution. Below we provide explicit bounds for k.

Proposition 2.3.7. *Let ε and δ be such that $0 < \varepsilon < 0.572$ and $0 < \delta < 1$. The rescaled Gaussian and the rescaled Rademacher distributions over $\mathbb{R}^{k \times n}$ with $k \geq 7.87\varepsilon^{-2}(6.9d + \log(1/\delta))$ for $\mathbb{K} = \mathbb{R}$ and $k \geq 7.87\varepsilon^{-2}(13.8d + \log(1/\delta))$ for $\mathbb{K} = \mathbb{C}$ are (ε, δ, d) oblivious $\ell_2 \to \ell_2$ subspace embeddings.*

Proof. See appendix. □

Remark 2.3.8. *For $\mathbb{K} = \mathbb{C}$, an embedding with a better theoretical bound for k than the one in Proposition 2.3.7 can be obtained by taking $\mathbf{\Theta} := \frac{1}{\sqrt{2}}(\mathbf{\Theta}_{\text{Re}} + j\mathbf{\Theta}_{\text{Im}})$, where $j = \sqrt{-1}$ and $\mathbf{\Theta}_{\text{Re}}, \mathbf{\Theta}_{\text{Im}} \in \mathbb{R}^{k \times n}$ are rescaled Gaussian matrices. It can be shown that such $\mathbf{\Theta}$ is an (ε, δ, d) oblivious $\ell_2 \to \ell_2$ subspace embedding for $k \geq 3.94\varepsilon^{-2}(13.8d + \log(1/\delta))$. A detailed proof of this fact is provided in the supplementary material. In this work, however, we shall consider only real-valued embeddings.*

For the P-SRHT distribution, $\mathbf{\Theta}$ is taken to be the first n columns of the matrix $k^{-1/2}(\mathbf{R}\mathbf{H}_s\mathbf{D}) \in \mathbb{R}^{k \times s}$, where s is the power of 2 such that $n \leq s < 2n$, $\mathbf{R} \in \mathbb{R}^{k \times s}$ are the first k rows of a random permutation of rows of the identity matrix, $\mathbf{H}_s \in \mathbb{R}^{s \times s}$ is a Walsh-Hadamard matrix[2], and $\mathbf{D} \in \mathbb{R}^{s \times s}$ is a random diagonal matrix with random entries such that $\mathbb{P}([\mathbf{D}]_{i,i} = \pm 1) = 1/2$.

Proposition 2.3.9. *Let ε and δ be such that $0 < \varepsilon < 1$ and $0 < \delta < 1$. The P-SRHT distribution over $\mathbb{R}^{k \times n}$ with $k \geq 2(\varepsilon^2 - \varepsilon^3/3)^{-1}\left[\sqrt{d} + \sqrt{8\log(6n/\delta)}\right]^2 \log(3d/\delta)$ is a (ε, δ, d) oblivious $\ell_2 \to \ell_2$ subspace embedding.*

Proof. See appendix. □

[2] The Walsh-Hadamard matrix $\mathbf{H}_s$ of dimension s, with s being a power of 2, is a structured matrix defined recursively by $\mathbf{H}_s = \mathbf{H}_{s/2} \otimes \mathbf{H}_2$, with $\mathbf{H}_2 := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. A product of $\mathbf{H}_s$ with a vector can be computed with $s\log_2(s)$ flops by using the fast Walsh-Hadamard transform.

Remark 2.3.10. *A product of P-SRHT and Gaussian (or Rademacher) matrices can lead to oblivious $\ell_2 \to \ell_2$ subspace embeddings that have better theoretical bounds for k than P-SRHT but still have low complexity of multiplication by a vector.*

We observe that the lower bounds in Propositions 2.3.7 and 2.3.9 are independent or only weakly (logarithmically) dependent on the dimension n and the probability of failure δ. In other words, $\mathbf{\Theta}$ with a moderate k can be guaranteed to satisfy (2.24) even for extremely large n and small δ. Note that the theoretical bounds for k shall be useful only for problems with rather high initial dimension, say with $n/r > 10^4$. Furthermore, in our experiments we revealed that the presented theoretical bounds are pessimistic. Another way for selecting the size for the random sketching matrix $\mathbf{\Theta}$ such that it is an ε-embedding for a given subspace V is the adaptive procedure proposed in Chapter 3 of this manuscript.

The rescaled Rademacher distribution and P-SRHT provide database-friendly matrices, which are easy to operate with. The rescaled Rademacher distribution is attractive from the data structure point of view and it can be efficiently implemented using standard SQL primitives [2]. The P-SRHT has a hierarchical structure allowing multiplications by vectors with only $s\log_2(s)$ flops, where s is a power of 2 and $n \leq s < 2n$, using the fast Walsh-Hadamard transform or even $2s\log_2(k+1)$ flops using a more sophisticated procedure proposed in [3]. In the algorithms P-SRHT distribution shall be preferred. However for multi-core computing, where the hierarchical structure of P-SRHT cannot be fully exploited, Gaussian or Rademacher matrices can be more expedient. Finally, we would like to point out that a random sequence needed for constructing a realization of Gaussian, Rademacher or P-SRHT distribution can be generated using a seeded random number generator. In this way, an embedding can be efficiently maintained with negligible communication (for parallel and distributed computing) and storage costs.

The following proposition can be used for constructing oblivious $X \to \ell_2$ subspace embeddings for general inner product $\langle \mathbf{R}_X \cdot, \cdot \rangle$ from classical $\ell_2 \to \ell_2$ subspace embeddings.

Proposition 2.3.11. *Let $\mathbf{Q} \in \mathbb{K}^{s\times n}$ be any matrix such that $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{R}_X$. If $\mathbf{\Omega} \in \mathbb{K}^{k\times s}$ is a (ε,δ,d) oblivious $\ell_2 \to \ell_2$ subspace embedding, then $\mathbf{\Theta} = \mathbf{\Omega}\mathbf{Q}$ is a (ε,δ,d) oblivious $X \to \ell_2$ subspace embedding.*

Proof. See appendix. □

Note that the matrix $\mathbf{Q}$ in Proposition 2.3.11 can be efficiently obtained block-wise (see Remark 2.2.7). In addition, there is no need to evaluate $\mathbf{\Theta} = \mathbf{\Omega}\mathbf{Q}$ explicitly.

2.4 ℓ_2-embeddings for projection-based MOR

In this section we integrate the sketching technique in the context of model order reduction methods from Section 2.2. Let us define the following subspace of U:

$$Y_r(\mu) := U_r + \mathrm{span}\{\mathbf{R}_U^{-1}\mathbf{r}(\mathbf{x};\mu) : \mathbf{x} \in U_r\}, \tag{2.25}$$

where $\mathbf{r}(\mathbf{x};\mu) = \mathbf{b}(\mu) - \mathbf{A}(\mu)\mathbf{x}$, and identify its dual space with $Y_r(\mu)' := \mathrm{span}\{\mathbf{R}_U\mathbf{x} : \mathbf{x} \in Y_r(\mu)\}$. Furthermore, let $\mathbf{\Theta} \in \mathbb{K}^{k\times n}$ be a certain sketching matrix seen as an $U \to \ell_2$ subspace embedding.

2.4.1 Galerkin projection

We propose to use random sketching for estimating the Galerkin projection. For any $\mathbf{x} \in U_r$ the residual $\mathbf{r}(\mathbf{x};\mu)$ belongs to $Y_r(\mu)'$. uently, taking into account Proposition 2.3.4, if $\mathbf{\Theta}$ is a $U \to \ell_2$ ε-subspace embedding for $Y_r(\mu)$, then for all $\mathbf{x} \in U_r$ the semi-norm $\|\mathbf{r}(\mathbf{x};\mu)\|_{U_r'}$ in (2.4) can be well approximated by $\|\mathbf{r}(\mathbf{x};\mu)\|^{\mathbf{\Theta}}_{U_r'}$. This leads to the sketched version of the Galerkin orthogonality condition:

$$\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}}_{U_r'} = 0. \tag{2.26}$$

The quality of projection $\mathbf{u}_r(\mu)$ satisfying (2.26) can be characterized by the following coefficients:

$$\alpha_r^{\mathbf{\Theta}}(\mu) := \min_{\mathbf{x}\in U_r\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|^{\mathbf{\Theta}}_{U_r'}}{\|\mathbf{x}\|_U}, \tag{2.27a}$$

$$\beta_r^{\mathbf{\Theta}}(\mu) := \max_{\mathbf{x}\in(\mathrm{span}\{\mathbf{u}(\mu)\}+U_r)\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|^{\mathbf{\Theta}}_{U_r'}}{\|\mathbf{x}\|_U}. \tag{2.27b}$$

Proposition 2.4.1 (Cea's lemma for sketched Galerkin projection). *Let $\mathbf{u}_r(\mu)$ satisfy (2.26). If $\alpha_r^{\mathbf{\Theta}}(\mu) > 0$, then the following relation holds*

$$\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \leq \left(1 + \frac{\beta_r^{\mathbf{\Theta}}(\mu)}{\alpha_r^{\mathbf{\Theta}}(\mu)}\right)\|\mathbf{u}(\mu) - \mathbf{P}_{U_r}\mathbf{u}(\mu)\|_U. \tag{2.28}$$

Proof. See appendix. □

Proposition 2.4.2. *Let*

$$a_r(\mu) := \max_{\mathbf{w}\in U_r\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{w}\|_{U'}}{\|\mathbf{A}(\mu)\mathbf{w}\|_{U_r'}}.$$

If $\mathbf{\Theta}$ *is a* $U \to \ell_2$ ε*-embedding for* $Y_r(\mu)$*, then*

$$\alpha_r^{\mathbf{\Theta}}(\mu) \geq \frac{1}{\sqrt{1+\varepsilon}}(1-\varepsilon a_r(\mu))\alpha_r(\mu), \tag{2.29a}$$

$$\beta_r^{\mathbf{\Theta}}(\mu) \leq \frac{1}{\sqrt{1-\varepsilon}}(\beta_r(\mu)+\varepsilon\beta(\mu)). \tag{2.29b}$$

Proof. See appendix. □

There are two ways to select a random distribution for $\mathbf{\Theta}$ such that it is guaranteed to be a $U \to \ell_2$ ε-embedding for $Y_r(\mu)$ for all $\mu \in \mathcal{P}$, simultaneously, with probability at least $1-\delta$. A first way applies when $\mathcal{P}$ is of finite cardinality. We can choose $\mathbf{\Theta}$ such that it is a $(\varepsilon, \delta \# \mathcal{P}^{-1}, d)$ oblivious $U \to \ell_2$ subspace embedding, where $d := \max_{\mu \in \mathcal{P}} \dim(Y_r(\mu))$ and apply a union bound for the probability of success. Since $d \leq 2r+1$, $\mathbf{\Theta}$ can be selected of moderate size. When $\mathcal{P}$ is infinite, we make a standard assumption that $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$ admit affine representations. It then follows directly from the definition of $Y_r(\mu)$ that $\bigcup_{\mu \in \mathcal{P}} Y_r(\mu)$ is contained in a low-dimensional space Y_r^*. Let d^* be the dimension of this space. By definition, if $\mathbf{\Theta}$ is a $(\varepsilon, \delta, d^*)$ oblivious $U \to \ell_2$ subspace embedding, then it is a $U \to \ell_2$ ε-embedding for Y_r^*, and hence for every $Y_r(\mu)$, simultaneously, with probability at least $1-\delta$.

The lower bound for $\alpha_r^{\mathbf{\Theta}}(\mu)$ in Proposition 2.4.2 depends on the product $\varepsilon a_r(\mu)$. In particular, to guarantee positivity of $\alpha_r^{\mathbf{\Theta}}(\mu)$ and ensure well-posedness of (2.26), condition $\varepsilon a_r(\mu) < 1$ has to be satisfied. The coefficient $a_r(\mu)$ is bounded from above by $\frac{\beta(\mu)}{\alpha_r(\mu)}$. Consequently, $a_r(\mu)$ for coercive well-conditioned operators is expected to be lower than for non-coercive ill-conditioned $\mathbf{A}(\mu)$. The condition number and coercivity of $\mathbf{A}(\mu)$, however, do not fully characterize $a_r(\mu)$. This coefficient rather reflects how well U_r corresponds to its image $\{\mathbf{A}(\mu)\mathbf{x} : \mathbf{x} \in U_r\}$ through the map $\mathbf{A}(\mu)$. For example, if the basis for U_r is formed from eigenvectors of $\mathbf{A}(\mu)$ then $a_r(\mu) = 1$. We also would like to note that the performance of the random sketching technique depends on the operator, only when it is employed for estimating the Galerkin projection. The accuracy of estimation of the residual error and the goal-oriented correction depends on the quality of sketching matrix $\mathbf{\Theta}$ but not on $\mathbf{A}(\mu)$. In addition, to make the performance of random sketching completely insensitive to the operator's properties, one can consider another type of projection (randomized minimal residual projection) for $\mathbf{u}_r(\mu)$ as is discussed in Chapter 3.

The coordinates of the solution $\mathbf{u}_r(\mu)$ of (2.26) can be found by solving

$$\mathbf{A}_r(\mu)\mathbf{a}_r(\mu) = \mathbf{b}_r(\mu), \tag{2.30}$$

where $\mathbf{A}_r(\mu) := \mathbf{U}_r^{\mathrm{H}}\mathbf{\Theta}^{\mathrm{H}}\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_r \in \mathbb{K}^{r\times r}$ and $\mathbf{b}_r(\mu) := \mathbf{U}_r^{\mathrm{H}}\mathbf{\Theta}^{\mathrm{H}}\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{b}(\mu) \in \mathbb{K}^r$.

Proposition 2.4.3. *Let $\mathbf{\Theta}$ be a $U \to \ell_2$ ε-embedding for U_r, and let $\mathbf{U}_r$ be orthogonal with respect to $\langle \cdot, \cdot \rangle_U^{\mathbf{\Theta}}$. Then the condition number of $\mathbf{A}_r(\mu)$ in (2.30) is bounded by $\sqrt{\frac{1+\varepsilon}{1-\varepsilon} \frac{\beta_r^{\mathbf{\Theta}}(\mu)}{\alpha_r^{\mathbf{\Theta}}(\mu)}}$.*

Proof. See appendix. □

2.4.2 Error estimation

Let $\mathbf{u}_r(\mu) \in U_r$ be an approximation of $\mathbf{u}(\mu)$. Consider the following error estimator:

$$\Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu); \mu) := \frac{\|\mathbf{r}(\mathbf{u}_r(\mu); \mu)\|_{U'}^{\mathbf{\Theta}}}{\eta(\mu)}, \tag{2.31}$$

where $\eta(\mu)$ is defined by (2.11). Below we show that under certain conditions, $\Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu); \mu)$ is guaranteed to be close to the classical error indicator $\Delta(\mathbf{u}_r(\mu); \mu)$.

Proposition 2.4.4. *If $\mathbf{\Theta}$ is a $U \to \ell_2$ ε-embedding for* $\operatorname{span}\{\mathbf{R}_U^{-1}\mathbf{r}(\mathbf{u}_r(\mu); \mu)\}$, *then*

$$\sqrt{1-\varepsilon}\Delta(\mathbf{u}_r(\mu); \mu) \leq \Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu); \mu) \leq \sqrt{1+\varepsilon}\Delta(\mathbf{u}_r(\mu); \mu). \tag{2.32}$$

Proof. See appendix. □

Corollary 2.4.5. *If $\mathbf{\Theta}$ is a $U \to \ell_2$ ε-embedding for $Y_r(\mu)$, then relation (2.32) holds.*

2.4.3 Primal-dual correction

The sketching technique can be applied to the dual problem in exactly the same manner as to the primal problem.

Let $\mathbf{u}_r(\mu) \in U_r$ and $\mathbf{u}_r^{\mathrm{du}}(\mu) \in U_r^{\mathrm{du}}$ be approximations of $\mathbf{u}(\mu)$ and $\mathbf{u}^{\mathrm{du}}(\mu)$, respectively. The sketched version of the primal-dual correction (2.14) can be expressed as follows

$$s_r^{\mathrm{spd}}(\mu) := s_r(\mu) - \langle \mathbf{u}_r^{\mathrm{du}}(\mu), \mathbf{R}_U^{-1}\mathbf{r}(\mathbf{u}_r(\mu); \mu)\rangle_U^{\mathbf{\Theta}}. \tag{2.33}$$

Proposition 2.4.6. *If $\mathbf{\Theta}$ is $U \to \ell_2$ ε-embedding for* $\operatorname{span}\{\mathbf{u}_r^{\mathrm{du}}(\mu), \mathbf{R}_U^{-1}\mathbf{r}(\mathbf{u}_r(\mu); \mu)\}$, *then*

$$|s(\mu) - s_r^{\mathrm{spd}}(\mu)| \leq \frac{\|\mathbf{r}(\mathbf{u}_r(\mu); \mu)\|_{U'}}{\eta(\mu)}((1+\varepsilon)\|\mathbf{r}^{\mathrm{du}}(\mathbf{u}_r^{\mathrm{du}}(\mu); \mu)\|_{U'} + \varepsilon\|\mathbf{l}(\mu)\|_{U'}). \tag{2.34}$$

Proof. See appendix. □

Remark 2.4.7. *We observe that the new version of primal-dual correction (2.33) and its error bound (2.34) are no longer symmetric in terms of the primal and dual solutions. When the residual error of $\mathbf{u}_r^{\mathrm{du}}(\mu)$ is smaller than the residual error of $\mathbf{u}_r(\mu)$, it can be more beneficial to consider the dual problem as the primal one and vice versa.*

Remark 2.4.8. *Consider the so called "compliant case", i.e., $\mathbf{A}(\mu)$ is self-adjoint, and $\mathbf{b}(\mu)$ is equal to $\mathbf{l}(\mu)$ up to a scaling factor. In such a case the same solution (up to a scaling factor) should be used for both the primal and the dual problems. If the approximation $\mathbf{u}_r(\mu)$ of $\mathbf{u}(\mu)$ is obtained with the classical Galerkin projection then the primal-dual correction is automatically included to the primal output quantity, i.e., $s_r(\mu) = s_r^{\mathrm{pd}}(\mu)$. A similar scenario can be observed for the sketched Galerkin projection. If $\mathbf{u}_r(\mu)$ satisfies (2.26) and the same $\boldsymbol{\Theta}$ is considered for both the projection and the inner product in (2.33), then $s_r(\mu) = s_r^{\mathrm{spd}}(\mu)$.*

It follows that if ε is of the order of $\|\mathbf{r}^{\mathrm{du}}(\mathbf{u}_r^{\mathrm{du}}(\mu);\mu)\|_{U'}/\|\mathbf{l}(\mu)\|_{U'}$, then the quadratic dependence in residual norm of the error bound is preserved. For relatively large ε, however, the error is expected to be proportional to $\varepsilon\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}$. Note that ε can decrease slowly with k (typically $\varepsilon = \mathcal{O}(k^{-1/2})$, see Propositions 2.3.7 and 2.3.9). Consequently, preserving high precision of the primal-dual correction can require large sketching matrices.

More accurate but yet efficient estimation of $s^{\mathrm{pd}}(\mu)$ can be obtained by introducing an approximation $\mathbf{w}_r^{\mathrm{du}}(\mu)$ of $\mathbf{u}_r^{\mathrm{du}}(\mu)$ such that the inner products with $\mathbf{w}_r^{\mathrm{du}}(\mu)$ are efficiently computable. Such approximation does not have to be very precise. As it will become clear later, it is sufficient to have $\mathbf{w}_r^{\mathrm{du}}(\mu)$ such that $\|\mathbf{u}_r^{\mathrm{du}}(\mu) - \mathbf{w}_r^{\mathrm{du}}(\mu)\|_U$ is of the order of $\varepsilon^{-1}\|\mathbf{u}_r^{\mathrm{du}}(\mu) - \mathbf{u}^{\mathrm{du}}(\mu)\|_U$. A possible choice is to let $\mathbf{w}_r^{\mathrm{du}}(\mu)$ be the orthogonal projection of $\mathbf{u}_r^{\mathrm{du}}(\mu)$ on a certain subspace $W_r^{\mathrm{du}} \subset U$, where W_r^{du} is such that it approximates well $\{\mathbf{u}_r^{\mathrm{du}}(\mu) : \mu \in \mathcal{P}\}$ but is much cheaper to operate with than U_r^{du}, e.g., if it has a smaller dimension. One can simply take $W_r^{\mathrm{du}} = U_i^{\mathrm{du}}$ (the subspace spanned by the first i^{du} basis vectors obtained during the generation of U_r^{du}), for some small $i^{\mathrm{du}} < r^{\mathrm{du}}$. A better approach consists in using a greedy algorithm or the POD method with a training set $\{\mathbf{u}_r^{\mathrm{du}}(\mu) : \mu \in \mathcal{P}_{\mathrm{train}}\}$. We could also choose W_r^{du} as the subspace associated with a coarse-grid interpolation of the solution. In this case, even if W_r^{du} has a high dimension, it can be operated with efficiently because its basis vectors are sparse. Strategies for the efficient construction of approximation spaces for $\mathbf{u}_r^{\mathrm{du}}(\mu)$ (or $\mathbf{u}_r(\mu)$) are provided in Chapter 3. Now, let us assume that $\mathbf{w}_r^{\mathrm{du}}(\mu)$ is given and consider the following estimation of $s_r^{\mathrm{pd}}(\mu)$:

$$s_r^{\mathrm{spd+}}(\mu) := s_r(\mu) - \langle \mathbf{w}_r^{\mathrm{du}}(\mu), \mathbf{r}(\mathbf{u}_r(\mu);\mu)\rangle - \langle \mathbf{u}_r^{\mathrm{du}}(\mu) - \mathbf{w}_r^{\mathrm{du}}(\mu), \mathbf{R}_U^{-1}\mathbf{r}(\mathbf{u}_r(\mu);\mu)\rangle_U^{\boldsymbol{\Theta}}. \tag{2.35}$$

We notice that $s_r^{\mathrm{spd+}}(\mu)$ can be evaluated efficiently but, at the same time, it has better accuracy than $s_r^{\mathrm{spd}}(\mu)$ in (2.34). By similar consideration as in Proposition 2.4.6

it can be shown that for preserving quadratic dependence in the error for $s_r^{\text{spd}+}(\mu)$, it is sufficient to have ε of the order of $\|\mathbf{u}_r^{\text{du}}(\mu) - \mathbf{u}^{\text{du}}(\mu)\|_{U'}/\|\mathbf{u}_r^{\text{du}}(\mu) - \mathbf{w}_r^{\text{du}}(\mu)\|_{U'}$.

Further, we assume that the accuracy of $s_r^{\text{spd}}(\mu)$ is sufficiently good so that there is no need to consider a corrected estimation $s_r^{\text{spd}+}(\mu)$. For other cases the methodology can be applied similarly.

2.4.4 Computing the sketch

In this section we introduce the concept of a sketch of the reduced order model. A sketch contains all the information needed for estimating the output quantity and certifying this estimation. It can be efficiently computed in basically any computational environment.

We restrict ourselves to solving the primal problem. Similar considerations also apply for the dual problem and primal-dual correction. The $\mathbf{\Theta}$-sketch of a reduced model associated with a subspace U_r is defined as

$$\left\{\left\{\mathbf{\Theta x}, \mathbf{\Theta R}_U^{-1}\mathbf{r}(\mathbf{x};\mu), \langle \mathbf{l}(\mu), \mathbf{x}\rangle\right\}: \quad \mathbf{x} \in U_r\right\} \tag{2.36}$$

In practice, each element of (2.36) can be represented by the coordinates of $\mathbf{x}$ associated with $\mathbf{U}_r$, i.e., a vector $\mathbf{a}_r \in \mathbb{K}^r$ such that $\mathbf{x} = \mathbf{U}_r\mathbf{a}_r$, the sketched reduced basis matrix $\mathbf{U}_r^{\mathbf{\Theta}} := \mathbf{\Theta U}_r$ and the following small parameter-dependent matrices and vectors:

$$\mathbf{V}_r^{\mathbf{\Theta}}(\mu) := \mathbf{\Theta R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_r, \quad \mathbf{b}^{\mathbf{\Theta}}(\mu) := \mathbf{\Theta R}_U^{-1}\mathbf{b}(\mu), \quad \mathbf{l}_r(\mu)^{\mathrm{H}} := \mathbf{l}(\mu)^{\mathrm{H}}\mathbf{U}_r. \tag{2.37}$$

Throughout the chapter, matrix $\mathbf{U}_r^{\mathbf{\Theta}}$ and the affine expansions of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$, $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ and $\mathbf{l}_r(\mu)$ shall be referred to as the $\mathbf{\Theta}$-sketch of $\mathbf{U}_r$. This object should not be confused with the $\mathbf{\Theta}$-sketch associated with a subspace U_r defined by (2.36). The $\mathbf{\Theta}$-sketch of $\mathbf{U}_r$ shall be used for characterizing the elements of the $\mathbf{\Theta}$-sketch associated with U_r similarly as $\mathbf{U}_r$ is used for characterizing the vectors in U_r.

The affine expansions of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$, $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ and $\mathbf{l}_r(\mu)$ can be obtained either by considering the affine expansions of $\mathbf{A}(\mu)$, $\mathbf{b}(\mu)$, and $\mathbf{l}(\mu)$[3] or with the empirical interpolation method (EIM) [106]. Given the sketch, the affine expansions of the quantities (e.g., $\mathbf{A}_r(\mu)$ in (2.30)) needed for efficient evaluation of the output can be computed with negligible cost. Computation of the $\mathbf{\Theta}$-sketch determines the cost of the offline stage and it has to be performed depending on the computational environment. We assume that the affine factors of $\mathbf{l}_r(\mu)$ are cheap to evaluate. Then the remaining computational cost is mainly associated with the following three operations: computing the samples (snapshots) of the solution (i.e., solving the full order problem for several $\mu \in \mathcal{P}$), performing matrix-vector products with $\mathbf{R}_U^{-1}$ and

[3] For instance, if $\mathbf{A}(\mu) = \sum_{i=1}^{m_A}\phi_i(\mu)\mathbf{A}_i$, then $\mathbf{V}_r^{\mathbf{\Theta}}(\mu) = \sum_{i=1}^{m_A}\phi_i(\mu)\left(\mathbf{\Theta R}_U^{-1}\mathbf{A}_i\mathbf{U}_r\right)$. Similar relations can also be derived for $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ and $\mathbf{l}_r(\mu)$.

the affine factors of $\mathbf{A}(\mu)$ (or $\mathbf{A}(\mu)$ evaluated at the interpolation points for EIM), and evaluating matrix-vector products with $\mathbf{\Theta}$.

The cost of obtaining the snapshots is assumed to be low compared to the cost of other offline computations such as evaluations of high dimensional inner and matrix-vector products. This is the case when the snapshots are computed beyond the main routine using highly optimised linear solver or a powerful server with limited budget. This is also the case when the snapshots are obtained on distributed machines with expensive communication costs. Solutions of linear systems of equations should have only a minor impact on the overall cost of an algorithm even when the basic metrics of efficiency, such as the complexity (number of floating point operations) and memory consumption, are considered. For large-scale problems solved in sequential or limited memory environments the computation of each snapshot should have log-linear (i.e., $\mathcal{O}(n(\log n)^d)$, for some small d) complexity and memory requirements. Higher complexity or memory requirements are usually not acceptable with standard architectures. In fact, in recent years there was an extensive development of methods for solving large-scale linear systems of equations [19, 70, 86] allowing computation of the snapshots with log-linear number of flops and bytes of memory (see for instance [29, 71, 105, 111, 158]). On the other hand, for classical model reduction, the evaluation of multiple inner products for the affine terms of reduced systems (2.9) and the quantities for error estimation (see Section 2.4.5) require $\mathcal{O}(nr^2m_A^2+nm_b^2)$ flops, with m_A and m_b being the numbers of terms in affine expansions of $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$, respectively, and $\mathcal{O}(nr)$ bytes of memory. We see that indeed the complexity and memory consumption of the offline stage can be highly dominated by the postprocessing of the snapshots but not their computation.

The matrices $\mathbf{R}_U$ and $\mathbf{A}(\mu)$ should be sparse or maintained in a hierarchical format [86], so that they can be multiplied by a vector using (log-)linear complexity and storage consumption. Multiplication of $\mathbf{R}_U^{-1}$ by a vector should also be an inexpensive operation with the cost comparable to the cost of computing matrix-vector products with $\mathbf{R}_U$. For many problems it can be beneficial to precompute a factorization of $\mathbf{R}_U$ and to use it for efficient multiplication of $\mathbf{R}_U^{-1}$ by multiple vectors. Note that for the typical $\mathbf{R}_U$ (such as stiffness and mass matrices) originating from standard discretizations of partial differential equations in two spatial dimensions, a sparse Cholesky decomposition can be precomputed using $\mathcal{O}(n^{3/2})$ flops and then used for multiplying $\mathbf{R}_U^{-1}$ by vectors with $\mathcal{O}(n\log n)$ flops. For discretized PDEs in higher spatial dimensions, or problems where $\mathbf{R}_U$ is dense, the classical Cholesky decomposition can be more burdensome to obtain and use. For better efficiency, the matrix $\mathbf{R}_U$ can be approximated by $\tilde{\mathbf{Q}}^{\mathrm{H}}\tilde{\mathbf{Q}}$ (with log-linear number of flops) using incomplete or hierarchical [20] Cholesky factorizations. Iterative Krylov methods with good preconditioning are an alternative way for computing products of $\mathbf{R}_U^{-1}$ with vectors with log-linear complexity [29]. Note that although multiplication of $\mathbf{R}_U^{-1}$ by a vector and computation of a snapshot both require solving high-dimensional systems of equations, the cost of the former operation should be considerably less than the cost of the later one due to good properties of $\mathbf{R}_U$ (such as positive-definiteness, symmetry, and parameter-independence providing ability of precomputing a decomposition). In a streaming environment, where the snapshots are provided as data-streams, a special care has to be payed to the memory constraints. It can be important to maintain $\mathbf{R}_U$ and the affine factors (or evaluations at EIM interpolation points) of $\mathbf{A}(\mu)$ with a reduced storage consumption. For discretized PDEs, for example, the entries of these matrices (if they are sparse) can be generated subdomain-by-subdomain on the fly. In such a case the conjugate gradient method can be a good choice for evaluating products of $\mathbf{R}_U^{-1}$ with vectors. In very extreme cases, e.g., where storage of even a single large vector is forbidden, $\mathbf{R}_U$ can be approximated by a block matrix and inverted block-by-block on the fly.

Next we discuss an efficient implementation of $\mathbf{\Theta}$. We assume that

$$\mathbf{\Theta} = \mathbf{\Omega}\mathbf{Q},$$

where $\mathbf{\Omega} \in \mathbb{K}^{k\times s}$ is a classical oblivious $\ell_2 \to \ell_2$ subspace embedding and $\mathbf{Q} \in \mathbb{K}^{s\times n}$ is such that $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{R}_U$ (see Propositions 2.3.7, 2.3.9 and 2.3.11).

The matrix $\mathbf{Q}$ can be expected to have a cost of multiplication by a vector comparable to $\mathbf{R}_U$. If needed, this matrix can be generated block-wise (see Remark 2.2.7) on the fly similarly to $\mathbf{R}_U$.

For environments where the measure of efficiency is the number of flops, a sketching matrix $\mathbf{\Omega}$ with fast matrix-vector multiplications such as P-SRHT is preferable. The complexity of a matrix-vector product for P-SRHT is only $2s\log_2(k+1)$, with s being the power of 2 such that $n \leq s < 2n$ [3, 31][4]. Consequently, assuming that $\mathbf{A}(\mu)$ is sparse, that multiplications of $\mathbf{Q}$ and $\mathbf{R}_U^{-1}$ by a vector take $\mathcal{O}(n(\log n)^d)$ flops, and that $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$ admit affine expansions with m_A and m_b terms respectively, the overall complexity of computation of a $\mathbf{\Theta}$-sketch of $\mathbf{U}_r$, using a P-SRHT matrix as $\mathbf{\Omega}$, from the snapshots is only

$$\mathcal{O}(n[rm_A\log k + m_b\log k + rm_A(\log n)^d]).$$

This complexity can be much less than the complexity of construction of the classical reduced model (including the precomputation of quantities needed for online evaluation of the residual error) from $\mathbf{U}_r$, which is $\mathcal{O}(n[r^2m_A^2 + m_b^2 + rm_A(\log n)^d])$. The efficiency of an algorithm can be also measured in terms of the number of passes taken over the data. Such a situation may arise when there is a restriction on the accessible amount of fast memory. In this scenario, both structured and unstructured matrices may provide drastic reductions of the computational cost. Due to robustness and simplicity of implementation, we suggest using Gaussian or Rademacher matrices over the others. For these matrices a seeded random number

[4] The straightforward implementation of P-SRHT using the fast Walsh-Hadamard transform results in $s\log_2(s)$ complexity of multiplication by a vector, which yields similar computational costs as the procedure from [3].

generator has to be utilized. It allows accessing the entries of $\mathbf{\Omega}$ on the fly with negligible storage costs [88]. In a streaming environment, multiplication of Gaussian or Rademacher matrices by a vector can be performed block-wise.

Note that all aforementioned operations are well suited for parallelization. Regarding distributed computing, a sketch of each snapshot can be obtained on a separate machine with absolutely no communication. The cost of transferring the sketches to the master machine will depend on the number of rows of $\mathbf{\Theta}$ but not the size of the full order problem.

Finally, let us comment on orthogonalization of $\mathbf{U}_r$ with respect to $\langle \cdot, \cdot \rangle_U^{\mathbf{\Theta}}$. This procedure is particularly important for numerical stability of the reduced system of equations (see Proposition 2.4.3). In our applications we are interested in obtaining a sketch of the orthogonal matrix but not the matrix itself. In such a case, operating with large-scale matrices and vectors is not necessary. Let us assume to be given a sketch of $\mathbf{U}_r$ associated with $\mathbf{\Theta}$. Let $\mathbf{T}_r \in \mathbb{K}^{r \times r}$ be such that $\mathbf{U}_r^{\mathbf{\Theta}} \mathbf{T}_r$ is orthogonal with respect to $\langle \cdot, \cdot \rangle$. Such a matrix can be obtained with a standard algorithm, e.g., QR factorization. It can be easily verified that $\mathbf{U}_r^* := \mathbf{U}_r \mathbf{T}_r$ is orthogonal with respect to $\langle \cdot, \cdot \rangle_U^{\mathbf{\Theta}}$. We have,

$$\mathbf{\Theta} \mathbf{U}_r^* = \mathbf{U}_r^{\mathbf{\Theta}} \mathbf{T}_r, \quad \mathbf{\Theta} \mathbf{R}_U^{-1} \mathbf{A}(\mu) \mathbf{U}_r^* = \mathbf{V}_r^{\mathbf{\Theta}}(\mu) \mathbf{T}_r, \text{ and } \mathbf{l}(\mu)^{\mathrm{H}} \mathbf{U}_r^* = \mathbf{l}_r(\mu)^{\mathrm{H}} \mathbf{T}_r.$$

Therefore, the sketch of $\mathbf{U}_r^*$ can be computed, simply, by multiplying $\mathbf{U}_r^{\mathbf{\Theta}}$ and the affine factors of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$, and $\mathbf{l}_r(\mu)^{\mathrm{H}}$, by $\mathbf{T}_r$.

2.4.5 Efficient evaluation of the residual norm

Until now we discussed how random sketching can be used for reducing the offline cost of precomputing factors of affine decompositions of the reduced operator and the reduced right-hand side. Let us now focus on the cost of the online stage. Often, the most expensive part of the online stage is the evaluation of the quantities needed for computing the residual norms for a posteriori error estimation due to many summands in their affine expansions. In addition, as was indicated in [36, 43], the classical procedure for the evaluation of the residual norms can be sensitive to round-off errors. Here we provide a less expensive way of computing the residual norms, which simultaneously offers a better numerical stability.

Let $\mathbf{u}_r(\mu) \in U_r$ be an approximation of $\mathbf{u}(\mu)$, and $\mathbf{a}_r(\mu) \in \mathbb{K}^r$ be the coordinates of $\mathbf{u}_r(\mu)$ associated with $\mathbf{U}_r$, i.e., $\mathbf{u}_r(\mu) = \mathbf{U}_r \mathbf{a}_r(\mu)$. The classical algorithm for evaluating the residual norm $\|\mathbf{r}(\mathbf{u}_r(\mu); \mu)\|_{U'}$ for a large finite set of parameters $\mathcal{P}_{\text{test}} \subseteq \mathcal{P}$ proceeds with expressing $\|\mathbf{r}(\mathbf{u}_r(\mu); \mu)\|_{U'}^2$ in the following form [82]

$$\|\mathbf{r}(\mathbf{u}_r(\mu); \mu)\|_{U'}^2 = \langle \mathbf{a}_r(\mu), \mathbf{M}(\mu) \mathbf{a}_r(\mu) \rangle + 2\mathrm{Re}(\langle \mathbf{a}_r(\mu), \mathbf{m}(\mu) \rangle) + m(\mu), \tag{2.38}$$

where affine expansions of $\mathbf{M}(\mu) := \mathbf{U}_r^{\mathrm{H}} \mathbf{A}(\mu)^{\mathrm{H}} \mathbf{R}_U^{-1} \mathbf{A}(\mu) \mathbf{U}_r$, $\mathbf{m}(\mu) := \mathbf{U}_r^{\mathrm{H}} \mathbf{A}(\mu)^{\mathrm{H}} \mathbf{R}_U^{-1} \mathbf{b}(\mu)$ and $m(\mu) := \mathbf{b}(\mu)^{\mathrm{H}} \mathbf{R}_U^{-1} \mathbf{b}(\mu)$ can be precomputed during the offline stage and used for efficient online evaluation of these quantities for each $\mu \in \mathcal{P}_{\text{test}}$. If $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$ admit affine representations with m_A and m_b terms, respectively, then the associated affine expansions of $\mathbf{M}(\mu)$, $\mathbf{m}(\mu)$ and $m(\mu)$ contain $\mathcal{O}(m_A^2), \mathcal{O}(m_A m_b), \mathcal{O}(m_b^2)$ terms respectively, therefore requiring $\mathcal{O}(r^2 m_A^2 + m_b^2)$ flops for their online evaluations.

An approximation of the residual norm can be obtained in a more efficient and numerically stable way with the random sketching technique. Let us assume that $\mathbf{\Theta} \in \mathbb{K}^{k \times n}$ is a $U \to \ell_2$ embedding such that $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}^{\mathbf{\Theta}}$ approximates well $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}$ (see Proposition 2.4.4). Let us also assume that the factors of affine decompositions of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ have been precomputed and are available. For each $\mu \in \mathcal{P}_{\text{test}}$ an estimation of the residual norm can be provided by

$$\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'} \approx \|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}^{\mathbf{\Theta}} = \|\mathbf{V}_r^{\mathbf{\Theta}}(\mu)\mathbf{a}_r(\mu) - \mathbf{b}^{\mathbf{\Theta}}(\mu)\|. \tag{2.39}$$

We notice that $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ have less terms in their affine expansions than the quantities in (2.38). The sizes of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$, however, can be too large to provide any online cost reduction. In order to improve the efficiency, we introduce an additional $(\varepsilon, \delta, 1)$ oblivious $\ell_2 \to \ell_2$ subspace embedding $\mathbf{\Gamma} \in \mathbb{K}^{k' \times k}$. The theoretical bounds for the number of rows of Gaussian, Rademacher and P-SRHT matrices sufficient to satisfy the $(\varepsilon, \delta, 1)$ oblivious $\ell_2 \to \ell_2$ subspace embedding property can be obtained from [2, Lemmas 4.1 and 5.1] and Proposition 2.3.9. They are presented in Table 2.1. Values are shown for $\varepsilon = 0.5$ and varying probabilities of failure δ. We note that in order to account for the case $\mathbb{K} = \mathbb{C}$ we have to employ [2, Lemmas 4.1 and 5.1] for the real part and the imaginary part of a vector, separately, with a union bound for the probability of success.

Table 2.1: The number of rows of Gaussian (or Rademacher) and P-SRHT matrices sufficient to satisfy the $(1/2, \delta, 1)$ oblivious $\ell_2 \to \ell_2$ ε-subspace embedding property.

	$\delta = 10^{-3}$	$\delta = 10^{-6}$	$\delta = 10^{-12}$	$\delta = 10^{-18}$
Gaussian	200	365	697	1029
P-SRHT	$96.4(8\log k + 69.6)$	$170(8\log k + 125)$	$313(8\log k + 236)$	$454(8\log k + 346)$

Remark 2.4.9. *In practice the bounds provided in Table 2.1 are pessimistic (especially for P-SRHT) and much smaller k' (say, $k' = 100$) may provide desirable results. In addition, in our experiments any significant difference in performance between Gaussian matrices, Rademacher matrices and P-SRHT has not been revealed.*

We observe that the number of rows of $\mathbf{\Gamma}$ can be chosen independent (or weakly dependent) of the number of rows of $\mathbf{\Theta}$. Let $\mathbf{\Phi} := \mathbf{\Gamma}\mathbf{\Theta}$. By definition, for each $\mu \in \mathcal{P}_{\text{test}}$

$$\mathbb{P}\left(\left|(\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}^{\mathbf{\Theta}})^2 - (\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}^{\mathbf{\Phi}})^2\right| \leq \varepsilon(\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}^{\mathbf{\Theta}})^2\right) \geq 1 - \delta; \tag{2.40}$$

which means that $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Phi}}_{U'}$ is an $\mathcal{O}(\varepsilon)$-accurate approximation of $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}$ with high probability. The probability of success for all $\mu \in \mathcal{P}_{\text{test}}$ simultaneously can be guaranteed with a union bound. In its turn, $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Phi}}_{U'}$ can be computed from

$$\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Phi}}_{U'} = \|\mathbf{V}^{\mathbf{\Phi}}_r(\mu)\mathbf{a}_r(\mu) - \mathbf{b}^{\mathbf{\Phi}}(\mu)\|, \tag{2.41}$$

where $\mathbf{V}^{\mathbf{\Phi}}_r(\mu) := \mathbf{\Gamma}\mathbf{V}^{\mathbf{\Theta}}_r(\mu)$ and $\mathbf{b}^{\mathbf{\Phi}}(\mu) := \mathbf{\Gamma}\mathbf{b}^{\mathbf{\Theta}}(\mu)$. The efficient way of computing $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Phi}}_{U'}$ for every $\mu \in \mathcal{P}_{\text{test}}$ consists in two stages. Firstly, we generate $\mathbf{\Gamma}$ and precompute affine expansions of $\mathbf{V}^{\mathbf{\Phi}}_r(\mu)$ and $\mathbf{b}^{\mathbf{\Phi}}(\mu)$ by multiplying each affine factor of $\mathbf{V}^{\mathbf{\Theta}}_r(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ by $\mathbf{\Gamma}$. The cost of this stage is independent of $\#\mathcal{P}_{\text{test}}$ (and n, of course) and becomes negligible for $\mathcal{P}_{\text{test}}$ of moderate size. In the second stage, for each parameter $\mu \in \mathcal{P}_{\text{test}}$, $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Phi}}_{U'}$ is evaluated from (2.41) using precomputed affine expansions. The quantities $\mathbf{V}^{\mathbf{\Phi}}_r(\mu)$ and $\mathbf{b}^{\mathbf{\Phi}}(\mu)$ contain at most the same number of terms as $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$ in their affine expansion. Consequently, if $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$ are parameter-separable with m_A and m_b terms, respectively, then each evaluation of $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Phi}}_{U'}$ from $\mathbf{a}_r(\mu)$ requires only $\mathcal{O}(k'rm_A + k'm_b)$ flops, which can be much less than the $\mathcal{O}(r^2m_A^2 + m_b^2)$ flops required for evaluating (2.38). Note that the classical computation of the residual norm by taking the square root of $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^2_{U'}$ evaluated using (2.38) can suffer from round-off errors. On the other hand, the evaluation of $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Phi}}_{U'}$ using (2.41) is less sensitive to round-off errors since here we proceed with direct evaluation of the (sketched) residual norm but not its square.

Remark 2.4.10. *If $\mathcal{P}_{\text{test}}$ (possibly very large but finite) is provided a priori, then the random matrix $\mathbf{\Gamma}$ can be generated and multiplied by the affine factors of $\mathbf{V}^{\mathbf{\Theta}}_r(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ during the offline stage.*

Remark 2.4.11. *For algorithms where $\mathcal{P}_{\text{test}}$ or U_r are selected adaptively based on a criterion depending on the residual norm (e.g., the classical greedy algorithm outlined in Section 2.2.4), a new realization of $\mathbf{\Gamma}$ has to be generated at each iteration. If the same realization of $\mathbf{\Gamma}$ is used for several iterations of the adaptive algorithm, care must be taken when characterizing the probability of success. This probability can decrease exponentially with the number of iterations, which requires to use considerably larger $\mathbf{\Gamma}$. Such option can be justified only for the cases when the cost of multiplying affine factors by $\mathbf{\Gamma}$ greatly dominates the cost of the second stage, i.e., evaluating $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Phi}}_{U'}$ for all $\mu \in \mathcal{P}_{\text{test}}$.*

2.5 Efficient reduced basis generation

In this section we show how the sketching technique can be used for improving the generation of reduced approximation spaces with greedy algorithm for RB, or a POD. Let $\mathbf{\Theta} \in \mathbb{K}^{k \times n}$ be a $U \to \ell_2$ subspace embedding.

2.5.1 Greedy algorithm

Recall that at each iteration of the greedy algorithm (see Section 2.2.4) the basis is enriched with a new sample (snapshot) $\mathbf{u}(\mu^{i+1})$, selected based on error indicator $\widetilde{\Delta}(U_i;\mu)$. The standard choice is $\widetilde{\Delta}(U_i;\mu) := \Delta(\mathbf{u}_i(\mu);\mu)$ where $\mathbf{u}_i(\mu) \in U_i$ satisfies (2.2). Such error indicator, however, can lead to very expensive computations. The error indicator can be modified to $\widetilde{\Delta}(U_i;\mu) := \Delta^{\mathbf{\Theta}}(\mathbf{u}_i(\mu);\mu)$, where $\mathbf{u}_i(\mu) \in U_i$ is an approximation of $\mathbf{u}(\mu)$ which does not necessarily satisfy (2.2). Further, we restrict ourselves to the case when $\mathbf{u}_i(\mu)$ is the sketched Galerkin projection (2.26). If there is no interest in reducing the cost of evaluating inner products but only reducing the cost of evaluating residual norms, it can be more relevant to consider the classical Galerkin projection (2.2) instead of (2.26).

A quasi-optimality guarantee for the greedy selection with $\widetilde{\Delta}(U_i;\mu) := \Delta^{\mathbf{\Theta}}(\mathbf{u}_i(\mu);\mu)$ can be derived from Propositions 2.4.1 and 2.4.2 and Corollary 2.4.5. At iteration i of the greedy algorithm, we need $\mathbf{\Theta}$ to be a $U \to \ell_2$ ε-subspace embedding for $Y_i(\mu)$ defined in (2.25) for all $\mu \in \mathcal{P}_{\text{train}}$. One way to achieve this is to generate a new realization of an oblivious $U \to \ell_2$ subspace embedding $\mathbf{\Theta}$ at each iteration of the greedy algorithm. Such approach, however, will lead to extra complexities and storage costs compared to the case where the same realization is employed for the entire procedure. In this work, we shall consider algorithms where $\mathbf{\Theta}$ is generated only once. When it is known that the set $\bigcup_{\mu\in\mathcal{P}_{\text{train}}} Y_r(\mu)$ belongs to a subspace Y^*_m of moderate dimension (e.g., when we operate on a small training set), then $\mathbf{\Theta}$ can be chosen such that it is a $U \to \ell_2$ ε-subspace embedding for Y^*_m with high probability. Otherwise, care must be taken when characterizing the probability of success because of the adaptive nature of the greedy algorithm. In such cases, all possible outcomes for U_r should be considered by using a union bound for the probability of success.

Proposition 2.5.1. *Let $U_r \subseteq U$ be a subspace obtained with r iterations of the greedy algorithm with error indicator depending on $\mathbf{\Theta}$. If $\mathbf{\Theta}$ is a $(\varepsilon, m^{-1}\binom{m}{r}^{-1}\delta, 2r+1)$ oblivious $U \to \ell_2$ subspace embedding, then it is a $U \to \ell_2$ ε-subspace embedding for $Y_r(\mu)$ defined in (2.25), for all $\mu \in \mathcal{P}_{\text{train}}$, with probability at least $1-\delta$.*

Proof. See appendix. □

Remark 2.5.2. *Theoretical bounds for the number of rows needed to construct $(\varepsilon, m^{-1}\binom{m}{r}^{-1}\delta, 2r+1)$ oblivious $U \to \ell_2$ subspace embeddings using Gaussian, Rademacher or P-SRHT distributions can be obtained from Propositions 2.3.7, 2.3.9 and 2.3.11. For Gaussian or Rademacher matrices they are proportional to r, while for P-SRHT they are proportional to r^2. In practice, however, embeddings built with P-SRHT, Gaussian or Rademacher distributions perform equally well.*

Evaluating $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}}_{U'}$ for very large training sets can be much more expensive than other costs. The complexity of this step can be reduced using the procedure explained in Section 2.4.5. The efficient sketched greedy algorithm is summarized in Algorithm 3. From Propositions 2.4.1 and 2.4.2, Corollary 2.4.5 and (2.40), we

Algorithm 3 Efficient sketched greedy algorithm

```
Given: $\mathcal{P}_{\text{train}}$, $\mathbf{A}(\mu)$, $\mathbf{b}(\mu)$, $\mathbf{l}(\mu)$, $\mathbf{\Theta}$, $\tau$.
Output: $U_r$
1. Set $i := 0$, $U_0 = \{\mathbf{0}\}$, and pick $\mu^1 \in \mathcal{P}_{\text{train}}$.
while $\max_{\mu \in \mathcal{P}_{\text{train}}} \widetilde{\Delta}(U_i;\mu) \geq \tau$ do
    2. Set $i := i+1$.
    3. Evaluate $\mathbf{u}(\mu^i)$ and set $U_i := U_{i-1} + \text{span}(\mathbf{u}(\mu^i))$.
    4. Update affine factors of $\mathbf{A}_i(\mu)$, $\mathbf{b}_i(\mu)$, $\mathbf{V}_i^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$.
    5. Generate $\mathbf{\Gamma}$ and evaluate affine factors of $\mathbf{V}_i^{\mathbf{\Phi}}(\mu)$ and $\mathbf{b}^{\mathbf{\Phi}}(\mu)$.
    6. Set $\widetilde{\Delta}(U_i;\mu) := \Delta^{\mathbf{\Phi}}(\mathbf{u}_i(\mu);\mu)$.
    7. Use (2.41) to find $\mu^{i+1} := \operatorname{argmax}_{\mu \in \mathcal{P}_{\text{train}}} \widetilde{\Delta}(U_i;\mu)$.
end while
```

can prove the quasi-optimality of the greedy selection in Algorithm 3 with high probability.

2.5.2 Proper Orthogonal Decomposition

Now we introduce the sketched version of POD. We first note that random sketching is a popular technique for obtaining low-rank approximations of large matrices [157]. It can be easily combined with Proposition 2.2.5 and Algorithm 2 for finding POD vectors. For large-scale problems, however, evaluating and storing POD vectors can be too expensive or even unfeasible, e.g., in a streaming or a distributed environment. We here propose a POD where evaluation of the full vectors is not necessary. We give a special attention to distributed computing. The computations involved in our version of POD can be distributed among separate machines with a communication cost independent of the dimension of the full order problem.

We observe that a complete reduced order model can be constructed from a sketch (see Section 2.4). Assume that we are given the sketch of a matrix $\mathbf{U}_m$ containing m solutions samples associated with $\mathbf{\Theta}$, i.e.,

$$\mathbf{U}_m^{\mathbf{\Theta}} := \mathbf{\Theta}\mathbf{U}_m, \;\; \mathbf{V}_m^{\mathbf{\Theta}}(\mu) := \mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_m, \;\; \mathbf{l}_m(\mu)^{\mathrm{H}} := \mathbf{l}(\mu)^{\mathrm{H}}\mathbf{U}_m, \;\; \mathbf{b}^{\mathbf{\Theta}}(\mu) := \mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{b}(\mu).$$

Recall that sketching a set of vectors can be efficiently performed basically in any modern computational environment, e.g., a distributed environment with expensive communication cost (see Section 2.4.4). Instead of computing a full matrix of reduced basis vectors, $\mathbf{U}_r \in \mathbb{K}^{n\times r}$, as in classical methods, we look for a small matrix

$\mathbf{T}_r \in \mathbb{K}^{m\times r}$ such that $\mathbf{U}_r = \mathbf{U}_m\mathbf{T}_r$. Given $\mathbf{T}_r$, the sketch of $\mathbf{U}_r$ can be computed without operating with the whole $\mathbf{U}_m$ but only with its sketch:

$$\mathbf{\Theta}\mathbf{U}_r = \mathbf{U}_m^{\mathbf{\Theta}}\mathbf{T}_r, \;\; \mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_r = \mathbf{V}_m^{\mathbf{\Theta}}(\mu)\mathbf{T}_r, \text{ and } \mathbf{l}(\mu)^{\mathrm{H}}\mathbf{U}_r = \mathbf{l}_m(\mu)^{\mathrm{H}}\mathbf{T}_r.$$

Further we propose an efficient way for obtaining $\mathbf{T}_r$ such that the quality of $U_r := \mathrm{span}(\mathbf{U}_r)$ is close to optimal.

For each $r \leq \mathrm{rank}(\mathbf{U}_m^{\mathbf{\Theta}})$, let U_r be an r-dimensional subspace obtained with the method of snapshots associated with norm $\|\cdot\|_U^{\mathbf{\Theta}}$, presented below.

Definition 2.5.3 (Sketched method of snapshots). *Consider the following eigenvalue problem*

$$\mathbf{G}\mathbf{t} = \lambda\mathbf{t} \tag{2.42}$$

where $\mathbf{G} := (\mathbf{U}_m^{\mathbf{\Theta}})^{\mathrm{H}}\mathbf{U}_m^{\mathbf{\Theta}}$. *Let* $l = \mathrm{rank}(\mathbf{U}_m^{\mathbf{\Theta}}) \geq r$ *and let* $\{(\lambda_i,\mathbf{t}_i)\}_{i=1}^{l}$ *be the solutions to (2.42) ordered such that* $\lambda_1 \geq \ldots \geq \lambda_l$. *Define*

$$U_r := \mathrm{range}(\mathbf{U}_m\mathbf{T}_r), \tag{2.43}$$

where $\mathbf{T}_r := [\mathbf{t}_1,...,\mathbf{t}_r]$.

For given $V \subseteq U_m$, let $\mathbf{P}_V^{\mathbf{\Theta}} : U_m \to V$ denote an orthogonal projection on V with respect to $\|\cdot\|_U^{\mathbf{\Theta}}$, i.e.,

$$\forall \mathbf{x} \in U_m, \;\; \mathbf{P}_V^{\mathbf{\Theta}}\mathbf{x} = \arg\min_{\mathbf{w}\in V} \|\mathbf{x}-\mathbf{w}\|_U^{\mathbf{\Theta}}, \tag{2.44}$$

and define the following error indicator:

$$\Delta^{\mathrm{POD}}(V) := \frac{1}{m}\sum_{i=1}^{m}\left(\|\mathbf{u}(\mu^i) - \mathbf{P}_V^{\mathbf{\Theta}}\mathbf{u}(\mu^i)\|_U^{\mathbf{\Theta}}\right)^2. \tag{2.45}$$

Proposition 2.5.4. *Let* $\{\lambda_i\}_{i=1}^{l}$ *be the set of eigenvalues from Definition 2.5.3. Then*

$$\Delta^{\mathrm{POD}}(U_r) := \frac{1}{m}\sum_{i=r+1}^{l}\lambda_i. \tag{2.46}$$

Moreover, for all $V_r \subseteq U_m$ *with* $\dim(V_r) \leq r$,

$$\Delta^{\mathrm{POD}}(U_r) \leq \Delta^{\mathrm{POD}}(V_r). \tag{2.47}$$

Proof. See appendix. □

Observe that the matrix $\mathbf{T}_r$ (characterizing U_r) can be much cheaper to obtain than the basis vectors for $U_r^* = POD_r(\mathbf{U}_m, \|\cdot\|_U)$. For this, we need to operate only with the sketched matrix $\mathbf{U}_m^{\mathbf{\Theta}}$ but not with the full snapshot matrix $\mathbf{U}_m$. Nevertheless, the quality of U_r can be guaranteed to be close to the quality of U_r^*.

Theorem 2.5.5. *Let $Y \subseteq U_m$ be a subspace of U_m with $\dim(Y) \geq r$, and let*

$$\Delta_Y = \frac{1}{m}\sum_{i=1}^{m} \|\mathbf{u}(\mu^i) - \mathbf{P}_Y \mathbf{u}(\mu^i)\|_U^2.$$

If $\mathbf{\Theta}$ is a $U \to \ell_2$ ε-subspace embedding for Y and every subspace in $\left\{\mathrm{span}(\mathbf{u}(\mu^i) - \mathbf{P}_Y \mathbf{u}(\mu^i))\right\}_{i=1}^{m}$ and $\left\{\mathrm{span}(\mathbf{u}(\mu^i) - \mathbf{P}_{U_r^*} \mathbf{u}(\mu^i))\right\}_{i=1}^{m}$, then

$$\begin{aligned}\frac{1}{m}\sum_{i=1}^{m} \|\mathbf{u}(\mu^i) - \mathbf{P}_{U_r} \mathbf{u}(\mu^i)\|_U^2 &\leq \frac{2}{1-\varepsilon}\Delta^{\mathrm{POD}}(U_r) + \left(\frac{2(1+\varepsilon)}{1-\varepsilon} + 1\right)\Delta_Y \\ &\leq \frac{2(1+\varepsilon)}{1-\varepsilon}\frac{1}{m}\sum_{i=1}^{m} \|\mathbf{u}(\mu^i) - \mathbf{P}_{U_r^*} \mathbf{u}(\mu^i)\|_U^2 + \left(\frac{2(1+\varepsilon)}{1-\varepsilon} + 1\right)\Delta_Y.\end{aligned} \tag{2.48}$$

Moreover, if $\mathbf{\Theta}$ is $U \to \ell_2$ ε-subspace embedding for U_m, then

$$\frac{1}{m}\sum_{i=1}^{m} \|\mathbf{u}(\mu^i) - \mathbf{P}_{U_r} \mathbf{u}(\mu^i)\|_U^2 \leq \frac{1}{1-\varepsilon}\Delta^{\mathrm{POD}}(U_r) \leq \frac{1+\varepsilon}{1-\varepsilon}\frac{1}{m}\sum_{i=1}^{m} \|\mathbf{u}(\mu^i) - \mathbf{P}_{U_r^*} \mathbf{u}(\mu^i)\|_U^2. \tag{2.49}$$

Proof. See appendix. □

By an union bound argument and the definition of an oblivious embedding, the hypothesis in the first part of Theorem 2.5.5 can be satisfied with probability at least $1-3\delta$ if $\mathbf{\Theta}$ is a $(\varepsilon, \delta, \dim(Y))$ and $(\varepsilon, \delta/m, 1)$ oblivious $U \to \ell_2$ embedding. A subspace Y can be taken as U_r^*, or a larger subspace making Δ_Y as small as possible. It is important to note that even if U_r is quasi-optimal, there is no guarantee that $\mathbf{\Theta}$ is a $U \to \ell_2$ ε-subspace embedding for U_r unless it is a $U \to \ell_2$ ε-subspace embedding for the whole U_m. Such guarantee can be unfeasible to achieve for large training sets. One possible solution is to maintain two sketches of $\mathbf{U}_m$: one for the method of snapshots, and one for Galerkin projections and residual norms. Another way (following considerations similar to [88]) is to replace $\mathbf{U}_m$ by its low-rank approximation $\widetilde{\mathbf{U}}_m = \mathbf{P}_W^{\mathbf{\Theta}} \mathbf{U}_m$, with $W = \mathrm{span}(\mathbf{U}_m \mathbf{\Omega}^*)$, where $\mathbf{\Omega}^*$ is a small random matrix (e.g., Gaussian matrix). The latter procedure can be also used for improving the efficiency of the algorithm when m is large. Finally, if $\mathbf{\Theta}$ is a $U \to \ell_2$ ε-subspace embedding for every subspace in $\{\mathrm{span}(\mathbf{u}_i - \mathbf{P}_{U_r}^{\mathbf{\Theta}} \mathbf{u}_i)\}_{i=1}^{m}$ then the error indicator $\Delta^{\mathrm{POD}}(U_r)$ is quasi-optimal. However, if only the first hypothesis of Theorem 2.5.5 is satisfied then the quality of $\Delta^{\mathrm{POD}}(U_r)$ will depend on Δ_Y. In such a case the error can be certified using $\Delta^{\mathrm{POD}}(\cdot)$ defined with a new realization of $\mathbf{\Theta}$.

2.6 Numerical examples

In this section the approach is validated numerically and compared against classical methods. For simplicity in all our experiments, we chose a coefficient $\eta(\mu) = 1$ in Equations (2.10) and (2.31) for the error estimation. The experiments revealed that the theoretical bounds for k in Propositions 2.3.7 and 2.3.9 and Table 2.1 are pessimistic. In practice, much smaller random matrices still provide good estimation of the output. In addition, we did not detect any significant difference in performance between Rademacher matrices, Gaussian matrices and P-SRHT, even though the theoretical bounds for P-SRHT are worse. Finally, the results obtained with Rademacher matrices are not presented. They are similar to those for Gaussian matrices and P-SRHT.

2.6.1 3D thermal block

We use a 3D version of the thermal block benchmark from [82]. This problem describes a heat transfer phenomenon through a domain $\Omega := [0,1]^3$ made of an assembly of blocks, each composed of a different material. The boundary value problem for modeling the thermal block is as follows

$$\begin{cases} -\boldsymbol{\nabla}\cdot(\kappa\boldsymbol{\nabla}T) &= 0, \quad \text{in } \Omega \\ T &= 0, \quad \text{on } \Gamma_D \\ \boldsymbol{n}\cdot(\kappa\boldsymbol{\nabla}T) &= 0, \quad \text{on } \Gamma_{N,1} \\ \boldsymbol{n}\cdot(\kappa\boldsymbol{\nabla}T) &= 1, \quad \text{on } \Gamma_{N,2}, \end{cases} \tag{2.50}$$

where T is the temperature field, $\boldsymbol{n}$ is the outward normal vector to the boundary, κ is the thermal conductivity, and Γ_D, $\Gamma_{N,1}$, $\Gamma_{N,2}$ are parts of the boundary defined by $\Gamma_D := \{(x,y,z) \in \partial\Omega : y = 1\}$, $\Gamma_{N,2} := \{(x,y,z) \in \partial\Omega : y = 0\}$ and $\Gamma_{N,1} := \partial\Omega \backslash (\Gamma_D \cup \Gamma_{N,2})$. Ω is partitioned into $2\times2\times2$ subblocks Ω_i of equal size. A different thermal conductivity κ_i is assigned to each Ω_i, i.e., $\kappa(x) = \kappa_i$, $x \in \Omega_i$. We are interested in estimating the mean temperature in $\Omega_1 := [0,\frac{1}{2}]^3$ for each $\mu := (\kappa_1, ..., \kappa_8) \in \mathcal{P} := [\frac{1}{10}, 10]^8$. The κ_i are independent random variables with log-uniform distribution over $[\frac{1}{10}, 10]$.

Problem (2.50) was discretized using the classical finite element method with approximately $n = 120000$ degrees of freedom. A function w in the finite element approximation space is identified with a vector $\mathbf{w} \in U$. The space U is equipped with an inner product compatible with the H_0^1 inner product, i.e., $\|\mathbf{w}\|_U := \|\boldsymbol{\nabla} w\|_{L_2}$. The training set $\mathcal{P}_{\text{train}}$ and the test set $\mathcal{P}_{\text{test}}$ were taken as 10000 and 1000 independent samples, respectively. The factorization of $\mathbf{R}_U$ was precomputed only once and used for efficient multiplication of $\mathbf{R}_U^{-1}$ by multiple vectors. The sketching matrix $\boldsymbol{\Theta}$ was constructed with Proposition 2.3.11, i.e., $\boldsymbol{\Theta} := \boldsymbol{\Omega}\mathbf{Q}$, where $\boldsymbol{\Omega} \in \mathbb{R}^{k\times s}$ is a classical oblivious $\ell_2 \to \ell_2$ subspace embedding and $\mathbf{Q} \in \mathbb{R}^{s\times n}$ is such that $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{R}_U$.

Furthermore, $\mathbf{Q}$ was taken as the transposed Cholesky factor of $\mathbf{R}_U$. Different distributions and sizes of the matrix $\mathbf{\Omega}$ were considered. The same realizations of $\mathbf{\Omega}$ were used for all parameters and greedy iterations within each experiment. A seeded random number generator was used for memory-efficient operations on random matrices. For P-SRHT, a fast implementation of the fast Walsh-Hadamard transform was employed for multiplying the Walsh-Hadamard matrix by a vector in $s\log_2(s)$ time. In Algorithm 3, we used $\mathbf{\Phi} := \mathbf{\Gamma}\mathbf{\Theta}$, where $\mathbf{\Gamma} \in \mathbb{R}^{k' \times k}$ is a Gaussian matrix and $k' = 100$. The same realizations of $\mathbf{\Gamma}$ were used for all the parameters but it was regenerated at each greedy iteration.

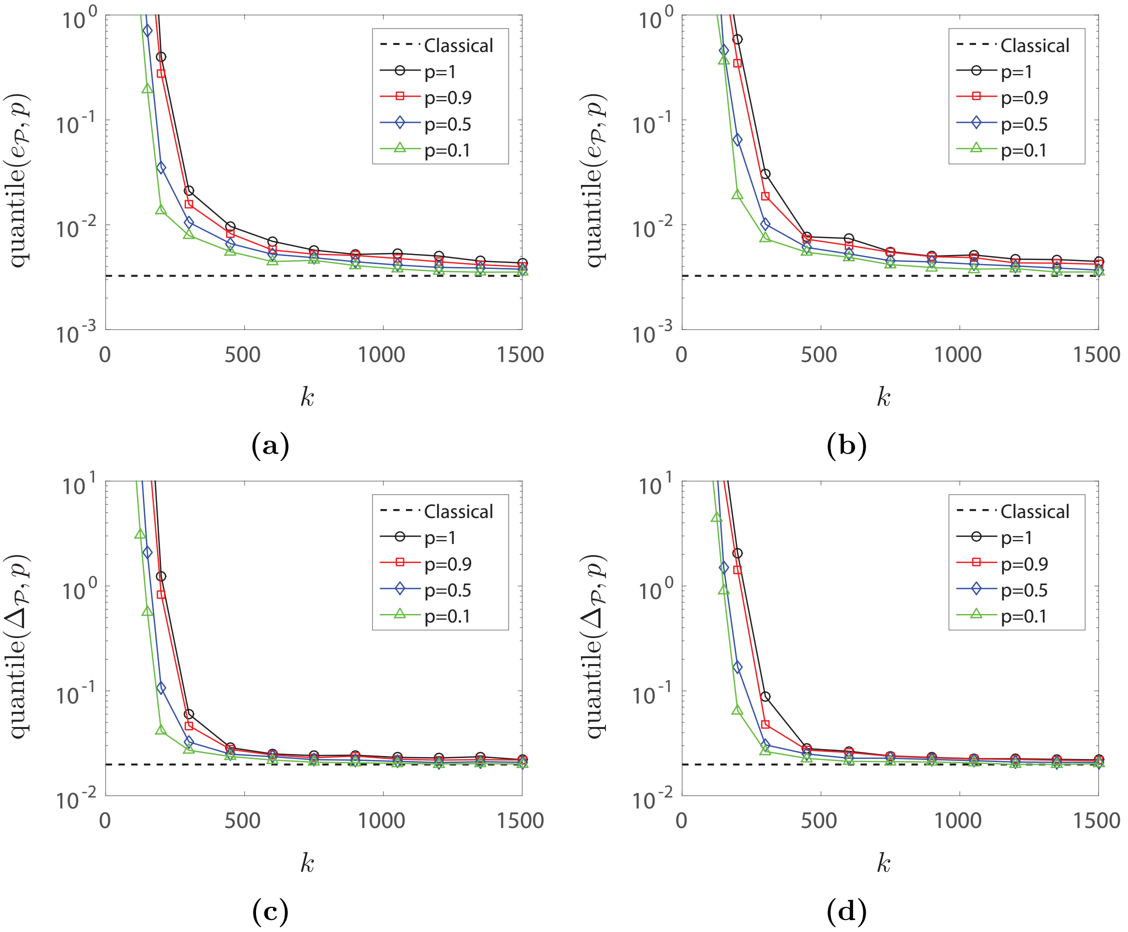
Galerkin projection and primal-dual correction. Let us investigate how the quality of the solution depends on the distribution and size of $\mathbf{\Omega}$. We first generated sufficiently accurate reduced subspaces U_r and U_r^{du} for the primal and the dual problems. The subspaces were spanned by snapshots evaluated at some points in $\mathcal{P}_{\mathrm{train}}$. The interpolation points were obtained by $r = 100$ iterations of the efficient sketched greedy algorithm (Algorithm 3) with P-SRHT and $k = 1000$ rows. Thereafter, $\mathbf{u}(\mu)$ was approximated by a projection $\mathbf{u}_r(\mu) \in U_r$. The classical Galerkin projection (2.2) and its sketched version (2.26) with different distributions and sizes of $\mathbf{\Omega}$ were considered. The quality of a parameter-dependent projection is measured by $e_{\mathcal{P}} := \max_{\mu\in\mathcal{P}_{\mathrm{test}}} \|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U / \max_{\mu\in\mathcal{P}_{\mathrm{test}}} \|\mathbf{u}(\mu)\|_U$ and $\Delta_{\mathcal{P}} := \max_{\mu\in\mathcal{P}_{\mathrm{test}}} \|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'} / \max_{\mu\in\mathcal{P}_{\mathrm{test}}} \|\mathbf{b}(\mu)\|_{U'}$. For each random projection 20 samples of $e_{\mathcal{P}}$ and $\Delta_{\mathcal{P}}$ were evaluated. Figure 2.1 describes how $e_{\mathcal{P}}$ and $\Delta_{\mathcal{P}}$ depend on the number of rows k[5]. We observe that the error associated with the sketched Galerkin projection is large when k is close to r, but as k increases, it asymptotically approaches the error of the classical Galerkin projection. The residual errors of the classical and the sketched projections become almost identical already for $k = 500$ while the exact errors become close for $k = 1000$. We also observe that for the aforementioned k there is practically no deviation of $\Delta_{\mathcal{P}}$ and only a little deviation of $e_{\mathcal{P}}$.

Note that the theoretical bounds for k to preserve the quasi-optimality constants of the classical Galerkin projection can be derived using Propositions 2.3.7 and 2.3.9 combined with Proposition 2.4.1 and a union bound for the probability of success. As was noted in Section 2.3.2, however, the theoretical bounds for k in Propositions 2.3.7 and 2.3.9 shall be useful only for large problems with, say $n/r > 10^4$, which means they should not be applicable here. Indeed, we see that for ensuring that

$$\mathbb{P}(\forall \mu \in \mathcal{P}_{\mathrm{test}} : \varepsilon a_r(\mu) < 1) > 1 - 10^{-6},$$

using the theoretical bounds, we need impractical values $k \geq 39280$ for Gaussian matrices and $k = n \approx 100000$ for P-SRHT. In practice, the value for k can be determined using the adaptive procedure proposed in Chapter 3.

[5] The p-quantile of a random variable X is defined as $\inf\{t : \mathbb{P}(X \leq t) \geq p\}$ and can be estimated by replacing the cumulative distribution function $\mathbb{P}(X \leq t)$ by its empirical estimation. Here we use 20 samples for this estimation.



(a) **(b)** **(c)** **(d)**

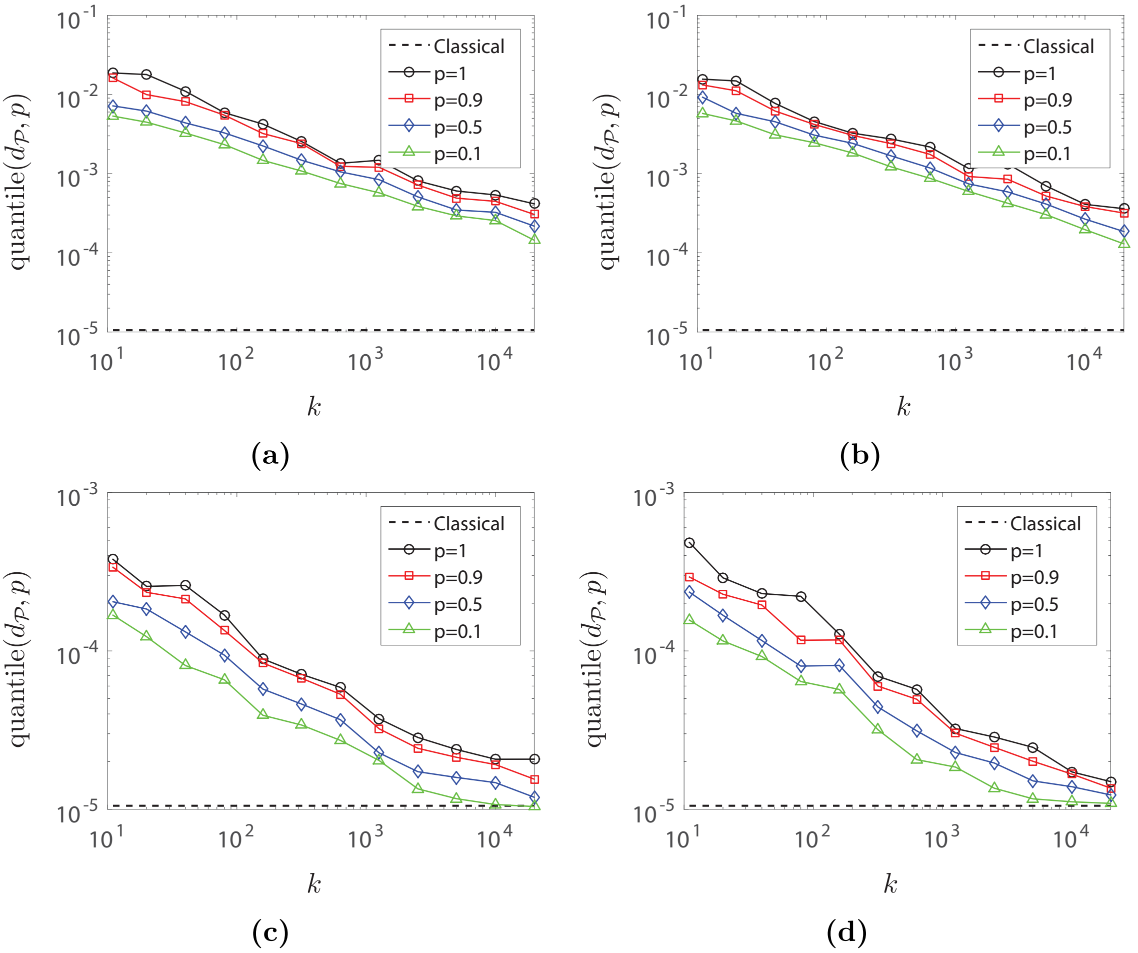
Figure 2.1: Errors $e_{\mathcal{P}}$ and $\Delta_{\mathcal{P}}$ of the classical Galerkin projection and quantiles of probabilities $p = 1, 0.9, 0.5$ and 0.1 over 20 samples of $e_{\mathcal{P}}$ and $\Delta_{\mathcal{P}}$ of the randomized Galerkin projection versus the number of rows of $\mathbf{\Omega}$. (a) The exact error $e_{\mathcal{P}}$ with rescaled Gaussian distribution as $\mathbf{\Omega}$. (b) The exact error $e_{\mathcal{P}}$ with P-SRHT matrix as $\mathbf{\Omega}$. (c) The residual error $\Delta_{\mathcal{P}}$ with rescaled Gaussian distribution as $\mathbf{\Omega}$. (d) The residual error $\Delta_{\mathcal{P}}$ with P-SRHT matrix as $\mathbf{\Omega}$.

Thereafter, we let $\mathbf{u}_r(\mu) \in U_r$ and $\mathbf{u}_r^{\mathrm{du}}(\mu) \in U_r^{\mathrm{du}}$ be the sketched Galerkin projections, where $\mathbf{\Omega}$ was taken as P-SRHT with $k = 500$ rows. For the fixed $\mathbf{u}_r(\mu)$ and $\mathbf{u}_r^{\mathrm{du}}(\mu)$ the classical primal-dual correction $s_r^{\mathrm{pd}}(\mu)$ (2.14), and the sketched primal-dual correction $s_r^{\mathrm{spd}}(\mu)$ (2.33) were evaluated using different sizes and distributions of $\mathbf{\Omega}$. In addition, the approach introduced in Section 2.4.3 for improving the accuracy of the sketched correction was employed. For $\mathbf{w}_r^{\mathrm{du}}(\mu)$ we chose the orthogonal projection of $\mathbf{u}_r^{\mathrm{du}}(\mu)$ on $W_r^{\mathrm{du}} := U_i^{\mathrm{du}}$ with $i^{\mathrm{du}} = 30$ (the subspace spanned by the first $i^{\mathrm{du}} = 30$ basis vectors obtained during the generation of U_r^{du}). With such $\mathbf{w}_r^{\mathrm{du}}(\mu)$ the improved correction $s_r^{\mathrm{spd+}}(\mu)$ defined by (2.35) was computed. It has to be mentioned that $s_r^{\mathrm{spd+}}(\mu)$ yielded additional computations. They, however, are cheaper than the computations required for constructing the classical reduced systems and evaluating the classical output quantities in about 10 times in terms of complexity and 6.67 times in terms of memory. We define the error by $d_{\mathcal{P}} := \max_{\mu \in \mathcal{P}_{\mathrm{test}}} |s(\mu) - \widetilde{s}_r(\mu)| / \max_{\mu \in \mathcal{P}_{\mathrm{test}}} |s(\mu)|$, where $\widetilde{s}_r(\mu) = s_r^{\mathrm{pd}}(\mu), s_r^{\mathrm{spd}}(\mu)$ or $s_r^{\mathrm{spd+}}(\mu)$. For each random correction we computed 20 samples of $d_{\mathcal{P}}$. The errors on the output quantities versus the numbers of rows of $\mathbf{\Theta}$ are presented in Figure 2.2. We see that the error of $s_r^{\mathrm{spd}}(\mu)$ is proportional to $k^{-1/2}$. It can be explained by the fact that for considered sizes of random matrices, ε is large compared to the residual error of the dual solution. As was noted in Section 2.4.3 in such a case the error bound for $s_r^{\mathrm{spd}}(\mu)$ is equal to $\mathcal{O}(\varepsilon \|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'})$. By Propositions 2.3.7 and 2.3.9 it follows that $\varepsilon = \mathcal{O}(k^{-1/2})$, which explains the behavior of the error in Figure 2.2. Note that the convergence of $s_r^{\mathrm{spd}}(\mu)$ is not expected to be reached even for k close to the dimension of the discrete problem. For large enough problems, however, the quality of the classical output will be always attained with $k \ll n$. In general, the error of the sketched primal-dual correction does not depend (or weakly depends for P-SRHT) on the dimension of the full order problem, but only on the accuracies of the approximate solutions $\mathbf{u}_r(\mu)$ and $\mathbf{u}_r^{\mathrm{du}}(\mu)$. On the other hand, we see that $s_r^{\mathrm{spd+}}(\mu)$ reaches the accuracy of the classical primal-dual correction for moderate k.

Further we focus only on the primal problem noting that similar results were observed also for the dual one.

Error estimation. We let U_r and $\mathbf{u}_r(\mu)$ be the subspace and the approximate solution from the previous experiment. The classical error indicator $\Delta(\mathbf{u}_r(\mu);\mu)$ and the sketched error indicator $\Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu);\mu)$ were evaluated for every $\mu \in \mathcal{P}_{\mathrm{test}}$. For $\Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu);\mu)$ different distributions and sizes of $\mathbf{\Omega}$ were considered. The quality of $\Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu);\mu)$ as estimator for $\Delta(\mathbf{u}_r(\mu);\mu)$ can be characterized by $e_{\mathcal{P}}^{\mathrm{ind}} := \max_{\mu \in \mathcal{P}_{\mathrm{test}}} |\Delta(\mathbf{u}_r(\mu);\mu) - \Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu);\mu)| / \max_{\mu \in \mathcal{P}_{\mathrm{test}}} \Delta(\mathbf{u}_r(\mu);\mu)$. For each $\mathbf{\Omega}$, 20 samples of $e_{\mathcal{P}}^{\mathrm{ind}}$ were evaluated. Figure 2.3b shows how $e_{\mathcal{P}}^{\mathrm{ind}}$ depends on k. The convergence of the error is proportional to $k^{-1/2}$, similarly as for the primal-dual correction. In practice, however, $\Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu);\mu)$ does not have to be so accurate as the approximation of the quantity of interest. For many problems, estimating $\Delta(\mathbf{u}_r(\mu);\mu)$ with relative error less than $1/2$ is already good enough. Consequently, $\Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu);\mu)$ employing $\mathbf{\Omega}$ with $k = 100$ or even $k = 10$ rows can be readily used as a reliable error estimator. Note that $\mathcal{P}_{\mathrm{test}}$ and U_r were formed independently of $\mathbf{\Omega}$. Otherwise, a larger $\mathbf{\Omega}$ should be considered with an additional embedding $\mathbf{\Gamma}$ as explained in Section 2.4.5.

To validate the claim that our approach (see Section 2.4.5) for error estimation provides more numerical stability than the classical one, we performed the following experiment. For fixed $\mu \in \mathcal{P}$ such that $\mathbf{u}(\mu) \in U_r$ we picked several vectors $\mathbf{u}_i^* \in U_r$ at different distances of $\mathbf{u}(\mu)$. For each such $\mathbf{u}_i^*$ we evaluated $\Delta(\mathbf{u}_i^*;\mu)$ and $\Delta^{\mathbf{\Theta}}(\mathbf{u}_i^*;\mu)$. The classical error indicator $\Delta(\mathbf{u}_i^*;\mu)$ was evaluated using the traditional procedure, i.e., expressing $\|\mathbf{r}(\mathbf{u}_i^*;\mu)\|_{U'}^2$ in the form (2.38), while $\Delta^{\mathbf{\Theta}}(\mathbf{u}_i^*;\mu)$ was evaluated with relation (2.39). The sketching matrix $\mathbf{\Omega}$ was generated from the P-SRHT or the



(a)

(b)

(c)

(d)

Figure 2.2: The error $d_{\mathcal{P}}$ of the classical primal-dual correction and quantiles of probabilities $p = 1, 0.9, 0.5$ and 0.1 over 20 samples of $d_{\mathcal{P}}$ of the randomized primal-dual corrections with fixed $\mathbf{u}_r(\mu)$ and $\mathbf{u}_r^{\text{du}}(\mu)$ versus the number of rows of $\mathbf{\Omega}$. (a) The errors of $s_r^{\text{pd}}(\mu)$ and $s_r^{\text{spd}}(\mu)$ with Gaussian matrix as $\mathbf{\Omega}$. (b) The errors of $s_r^{\text{pd}}(\mu)$ and $s_r^{\text{spd}}(\mu)$ with P-SRHT distribution as $\mathbf{\Omega}$. (c) The errors of $s_r^{\text{pd}}(\mu)$ and $s_r^{\text{spd+}}(\mu)$ with Gaussian matrix as $\mathbf{\Omega}$ and $W_r^{\text{du}} := U_i^{\text{du}}$, $i^{\text{du}} = 30$. (d) The errors of $s_r^{\text{pd}}(\mu)$ and $s_r^{\text{spd+}}(\mu)$ with P-SRHT distribution as $\mathbf{\Omega}$ and $W_r^{\text{du}} := U_i^{\text{du}}$, $i^{\text{du}} = 30$.

rescaled Gaussian distribution with $k = 100$ rows. Note that μ and $\mathbf{u}_i^*$ were chosen independently of $\mathbf{\Omega}$ so there is no point to use larger $\mathbf{\Omega}$ with additional embedding $\mathbf{\Gamma}$ (see Section 2.4.5). Figure 2.4 clearly reveals the failure of the classical error indicator at $\Delta(\mathbf{u}_i^*;\mu)/\|\mathbf{b}(\mu)\|_{U'} \approx 10^{-7}$. On the contrary, the indicators computed with random sketching technique remain reliable even for $\Delta(\mathbf{u}_i^*;\mu)/\|\mathbf{b}(\mu)\|_{U'}$ close to the machine precision.

Efficient sketched greedy algorithm. Further, we validate the performance of the efficient sketched greedy algorithm (Algorithm 3). For this we generated

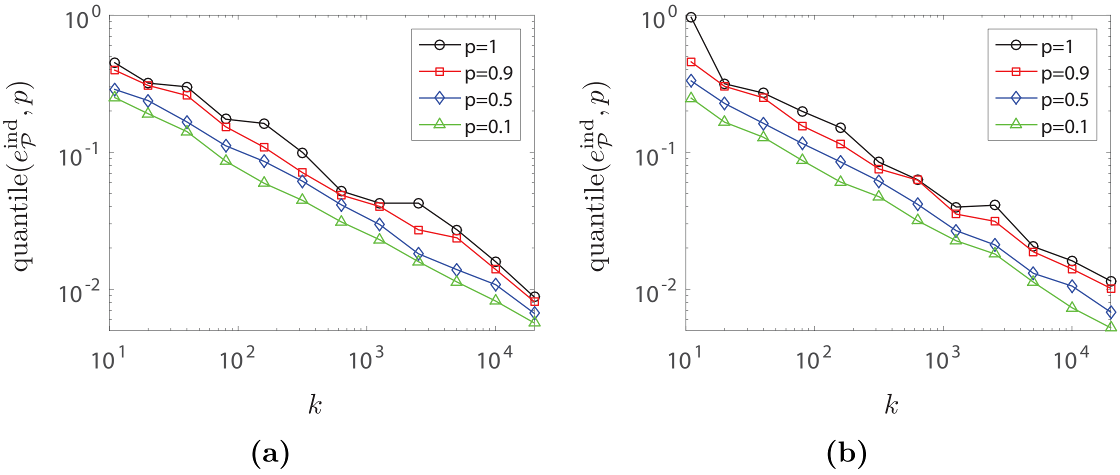


Figure 2.3: Quantiles of probabilities $p = 1, 0.9, 0.5$ and 0.1 over 20 samples of the error $e_{\mathcal{P}}^{\text{ind}}$ of $\Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu);\mu)$ as estimator of $\Delta(\mathbf{u}_r(\mu);\mu)$. (a) The error of $\Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu);\mu)$ with Gaussian distribution. (b) The error of $\Delta^{\mathbf{\Theta}}(\mathbf{u}_r(\mu);\mu)$ with P-SRHT distribution.

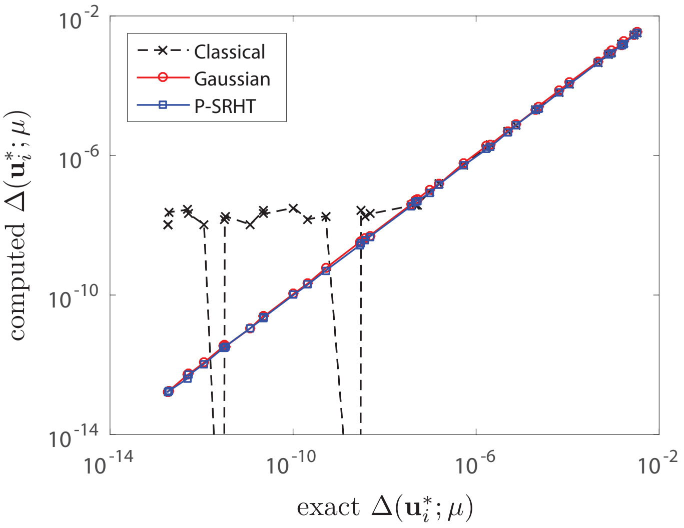


Figure 2.4: Error indicator $\Delta(\mathbf{u}_i^*;\mu)$ (rescaled by $\|\mathbf{b}(\mu)\|_{U'}$) computed with the classical procedure and its estimator $\Delta^{\mathbf{\Theta}}(\mathbf{u}_i^*;\mu)$ computed with relation (2.39) employing P-SRHT or Gaussian distribution with $k = 100$ rows versus the exact value of $\Delta(\mathbf{u}_i^*;\mu)$ (rescaled by $\|\mathbf{b}(\mu)\|_{U'}$).

a subspace U_r of dimension $r = 100$ using the classical greedy algorithm (depicted in Section 2.2.4) and its randomized version (Algorithm 3) employing $\mathbf{\Omega}$ of different types and sizes. In Algorithm 3, $\mathbf{\Gamma}$ was generated from a Gaussian distribution with $k' = 100$ rows. The error at i-th iteration is identified with $\Delta_{\mathcal{P}} := \max_{\mu\in\mathcal{P}_{\text{train}}} \|\mathbf{r}(\mathbf{u}_i(\mu);\mu)\|_{U'} / \max_{\mu\in\mathcal{P}_{\text{train}}} \|\mathbf{b}(\mu)\|_{U'}$. The convergence rates are depicted in Figure 2.5. For the efficient sketched greedy algorithm with $k = 250$ and $k = 500$ a slight difference in performance is detected compared to the classical algorithm. The difference is more evident for $k = 250$ at higher iterations. The behavior of the classical algorithm and Algorithm 3 with $k = 1000$ are almost identical.

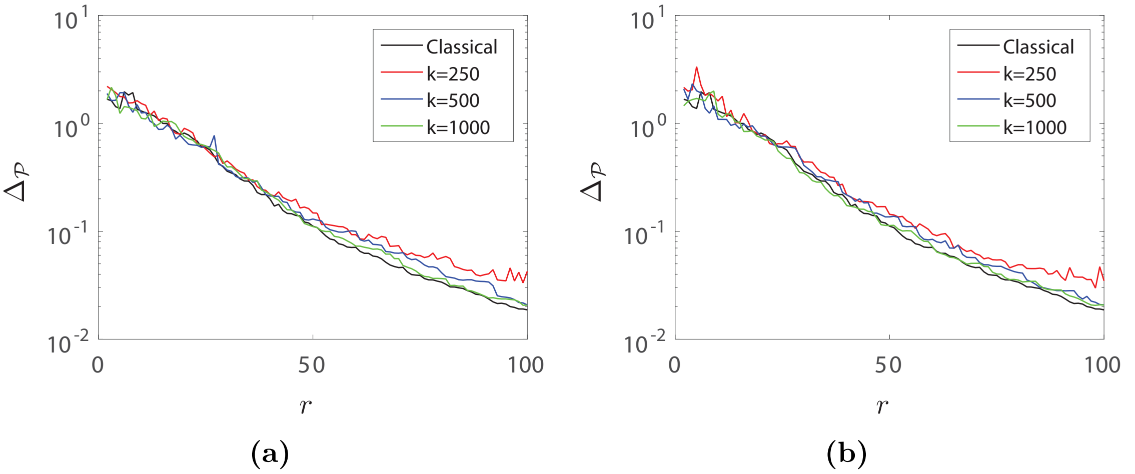


Figure 2.5: Convergence of the classical greedy algorithm (depicted in Section 2.2.4)) and its efficient randomized version (Algorithm 3) using $\mathbf{\Omega}$ drawn from (a) Gaussian distribution or (b) P-SRHT distribution.

Efficient Proper Orthogonal Decomposition. We finish with validation of the efficient randomized version of POD. For this experiment only $m = 1000$ points from $\mathcal{P}_{\text{train}}$ were considered as the training set. The POD bases were obtained with the classical method of snapshots, i.e., Algorithm 2 where $\mathbf{B}_r$ was computed from the SVD of $\mathbf{QU}_m$, or the randomized version of POD introduced in Section 2.5.2. The same $\mathbf{\Omega}$ was used for both the basis generation and the error estimation with $\Delta^{\text{POD}}(U_r)$, defined in (2.45). From Figure 2.6a we observe that for large enough k the quality of the POD basis formed with the new efficient algorithm is close to the quality of the basis obtained with the classical method. Construction of $r = 100$ basis vectors using $\mathbf{\Omega}$ with only $k = 500$ rows provides an almost optimal error. As expected, the error indicator $\Delta^{\text{POD}}(U_r)$ is close to the exact error for large enough k, but it represents the error poorly for small k. Furthermore, $\Delta^{\text{POD}}(U_r)$ is always smaller than the true error and is increasing monotonically with k. Figure 2.6b depicts how the errors of the classical and randomized (with $k = 500$) POD bases depend on the dimension of U_r. We see that the qualities of the basis and the error indicator obtained with the new version of POD remain close to the optimal ones up to dimension $r = 150$. However, as r becomes larger the quasi-optimality of the randomized POD degrades so that for $r \geq 150$ the sketching size $k = 500$ becomes insufficient.

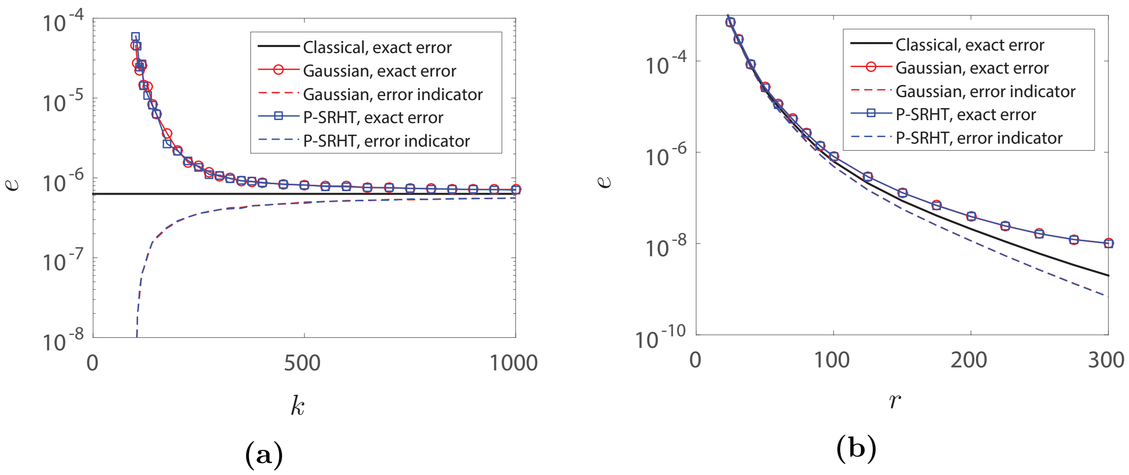


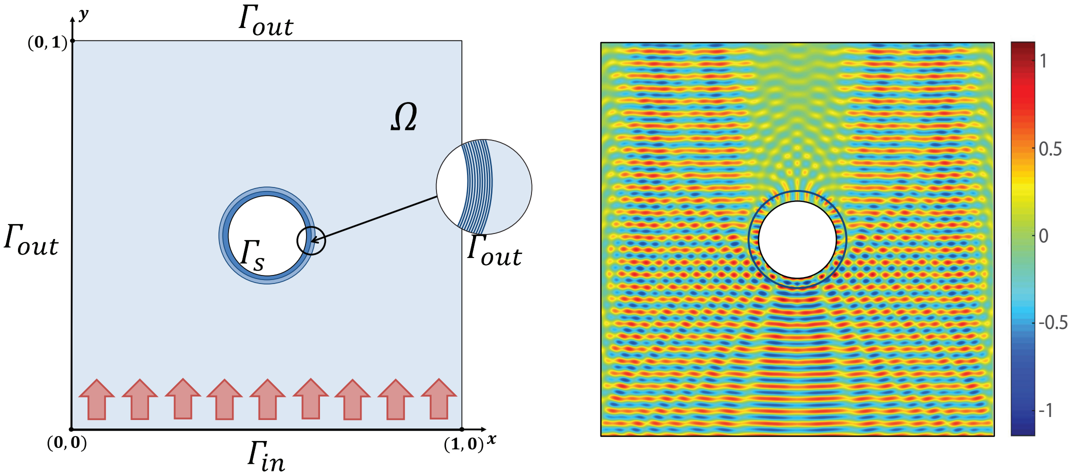
Figure 2.6: Error $e = \frac{1}{m}\sum_{i=1}^{m}\|\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}\mathbf{u}(\mu^i)\|_U^2 / (\frac{1}{m}\sum_{i=1}^{m}\|\mathbf{u}(\mu^i)\|_U^2)$ and error indicator $e = \Delta^{\text{POD}}(U_r)/(\frac{1}{m}\sum_{i=1}^{m}\|\mathbf{u}(\mu^i)\|_U^2)$ associated with U_r computed with traditional POD and its efficient randomized version introduced in Section 2.5.2. (a) Errors and indicators versus the number of rows of $\mathbf{\Omega}$ for $r = 100$. (b) Errors and indicators versus the dimension of U_r for $k = 500$.

2.6.2 Multi-layered acoustic cloak

In the previous numerical example we considered a problem with strongly coercive well-conditioned operator. But as was discussed in Section 2.4.1, random sketching with a fixed number of rows is expected to perform worse for approximating the Galerkin projection with non-coercive ill-conditioned $\mathbf{A}(\mu)$. Further, we would like to validate the methodology on such a problem. The benchmark consists in a scattering problem of a 2D wave with perfect scatterer covered in a multi-layered cloak. For this experiment we solve the following Helmholtz equation with first order absorbing boundary conditions

$$\begin{cases} \Delta u + \kappa^2 u &= 0, & \text{in } \Omega \\ \mathrm{i}\kappa u + \frac{\partial u}{\partial \boldsymbol{n}} &= 0, & \text{on } \Gamma_{out} \\ \mathrm{i}\kappa u + \frac{\partial u}{\partial \boldsymbol{n}} &= 2\mathrm{i}\kappa, & \text{on } \Gamma_{in} \\ \frac{\partial u}{\partial \boldsymbol{n}} &= 0, & \text{on } \Gamma_s, \end{cases} \tag{2.51}$$

where u is the solution field (primal unknown), κ is the wave number and the geometry of the problem is defined in Figure 2.7a. The background has a fixed wave number $\kappa = \kappa_0 := 50$. The cloak consists of 10 layers of equal thicknesses enumerated in the order corresponding to the distance to the scatterer. The i-th layer is composed of a material with wave number $\kappa = \kappa_i$. The quantity of interest is the average of the solution field on Γ_{in}. The aim is to estimate the quantity of interest for each parameter $\mu := (\kappa_1, ..., \kappa_{10}) \in [\kappa_0, \sqrt{2}\kappa_0]^{10} := \mathcal{P}$. The κ_i are considered as independent random variables with log-uniform distribution over $[\kappa_0, \sqrt{2}\kappa_0]$. The solution for a randomly chosen $\mu \in \mathcal{P}$ is illustrated in Figure 2.7b.



(a) Geometry **(b)** Solution at random μ

Figure 2.7: (a) Geometry of acoustic cloak benchmark. (b) The real component of u for randomly picked parameter $\mu = (66.86, 54.21, 61.56, 64.45, 66.15, 58.42, 54.90, 63.79, 58.44, 63.09)$.

The problem has a symmetry with respect to the vertical axis $x = 0.5$. Consequently, only half of the domain has to be considered for discretization. The discretization was performed using quadratic triangular finite elements with approximately 17 complex degrees of freedom per wavelength, i.e., around 200000 complex degrees of freedom in total. A function w in the approximation space is identified with a vector $\mathbf{w} \in U$. The solution space U is equipped with an inner product compatible with the H^1 inner product, i.e.,

$$\|\mathbf{w}\|_U^2 := \|\boldsymbol{\nabla} w\|_{L_2}^2 + \kappa_0^2 \|w\|_{L_2}^2.$$

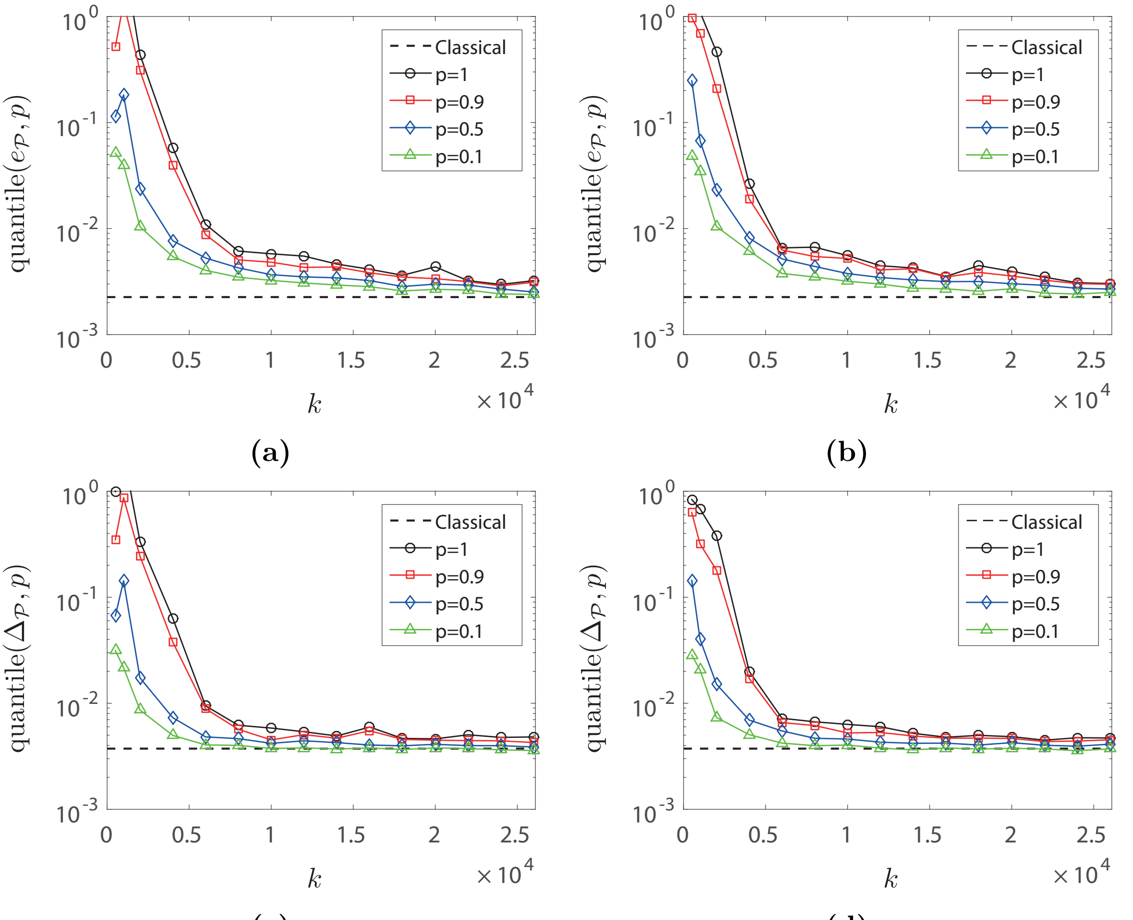
Further, 20000 and 1000 independent samples were considered as the training set $\mathcal{P}_{\text{train}}$ and the test set $\mathcal{P}_{\text{test}}$, respectively. The sketching matrix $\boldsymbol{\Theta}$ was constructed as in the thermal block benchmark, i.e., $\boldsymbol{\Theta} := \boldsymbol{\Omega}\mathbf{Q}$, where $\boldsymbol{\Omega} \in \mathbb{R}^{k \times s}$ is either a Gaussian matrix or P-SRHT and $\mathbf{Q} \in \mathbb{R}^{s \times n}$ is the transposed Cholesky factor of $\mathbf{R}_U$. In addition, we used $\boldsymbol{\Phi} := \boldsymbol{\Gamma}\boldsymbol{\Theta}$, where $\boldsymbol{\Gamma} \in \mathbb{R}^{k' \times k}$ is a Gaussian matrix and $k' = 200$.

Below we present validation of the Galerkin projection and the greedy algorithm only. The performance of our methodology for error estimation and POD does not depend on the operator and is similar to the performance observed in the previous numerical example.

Galerkin projection. A subspace U_r was generated with $r = 150$ iterations of the randomized greedy algorithm (Algorithm 3) with a $\boldsymbol{\Omega}$ drawn from the P-SRHT distribution with $k = 20000$ rows. Such U_r was then used for validation of the

Galerkin projection. We evaluated multiple approximations of $\mathbf{u}(\mu)$ using either the classical projection (2.2) or its randomized version (2.26). Different $\mathbf{\Omega}$ were considered for (2.26). As before, the approximation and residual errors are respectively defined by $e_{\mathcal{P}} := \max_{\mu \in \mathcal{P}_{\text{test}}} \|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U / \max_{\mu \in \mathcal{P}_{\text{test}}} \|\mathbf{u}(\mu)\|_U$ and $\Delta_{\mathcal{P}} := \max_{\mu \in \mathcal{P}_{\text{test}}} \|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'} / \max_{\mu \in \mathcal{P}_{\text{test}}} \|\mathbf{b}(\mu)\|_{U'}$. For each type and size of $\mathbf{\Omega}$, 20 samples of $e_{\mathcal{P}}$ and $\Delta_{\mathcal{P}}$ were evaluated. The errors are presented in Figure 2.8. This experiment reveals that indeed the performance of random sketching is worse than in the thermal block benchmark (see Figure 2.1). For $k = 1000$ the error of the randomized version of Galerkin projection is much larger than the error of the classical projection. Whereas for the same value of k in the thermal block benchmark practically no difference between the qualities of the classical projection and its sketched version was observed. It can be explained by the fact that the quality of randomized Galerkin projection depends on the coefficient $a_r(\mu)$ defined in Proposition 2.4.2, which in its turn depends on the operator. In both numerical examples the coefficient $a_r(\mu)$ was measured over $\mathcal{P}_{\text{test}}$. We observed that $\max_{\mu \in \mathcal{P}_{\text{test}}} a_r(\mu) = 28.3$, while in the thermal block benchmark $\max_{\mu \in \mathcal{P}_{\text{test}}} a_r(\mu) = 2.65$. In addition, here we work on the complex field instead of the real field and consider slightly larger reduced subspaces, which could also have an impact on the accuracy of random sketching. Reduction of performance, however, is not that severe and already starting from $k = 15000$ the sketched version of Galerkin projection has an error close to the classical one. Such size of $\mathbf{\Omega}$ is still very small compared to the dimension of the discrete problem and provides drastic reduction of the computational cost. Let us also note that one could obtain a good approximation of $\mathbf{u}(\mu)$ from the sketch with $k \ll 15000$ by considering another type of projection (a randomized minimal residual projection) proposed in Chapter 3.

Let us further note that we are in the so called "compliant case" (see Remark 2.4.8). Thus, for the classical Galerkin projection we have $s_r(\mu) = s_r^{\text{pd}}(\mu)$ and for the sketched Galerkin projection, $s_r(\mu) = s_r^{\text{spd}}(\mu)$. The output quantity $s_r(\mu)$ was computed with the classical Galerkin projection and with the randomized Galerkin projection employing different $\mathbf{\Omega}$. For each $\mathbf{\Omega}$ we also computed the improved sketched correction $s_r^{\text{spd+}}(\mu)$ (see Section 2.4.3) using $W_r^{\text{du}} := U_i^{\text{du}}$ with $i^{\text{du}} = 30$. It required inexpensive additional computations which are in about 5 times cheaper (in terms of both complexity and memory) than the computations involved in the classical method. The error on the output quantity is measured by $d_{\mathcal{P}} := \max_{\mu \in \mathcal{P}_{\text{test}}} |s(\mu) - \widetilde{s}_r(\mu)| / \max_{\mu \in \mathcal{P}_{\text{test}}} |s(\mu)|$, where $\widetilde{s}_r(\mu) = s_r(\mu)$ or $s_r^{\text{spd+}}(\mu)$. For each random distribution type 20 samples of $d_{\mathcal{P}}$ were evaluated. Figure 2.9 describes how the error of the output quantity depends on k. For small k the error is large because of the poor quality of the projection and lack of precision when approximating the inner product for $s_r^{\text{pd}}(\mu)$ in (2.14) by the one in (2.33). But starting from $k = 15000$ we see that the quality of $s_r(\mu)$ obtained with the random sketching technique becomes close to the quality of the output computed with the classical Galerkin projection. For $k \geq 15000$ the randomized Galerkin projection has practically the same accuracy



(a)

(b)

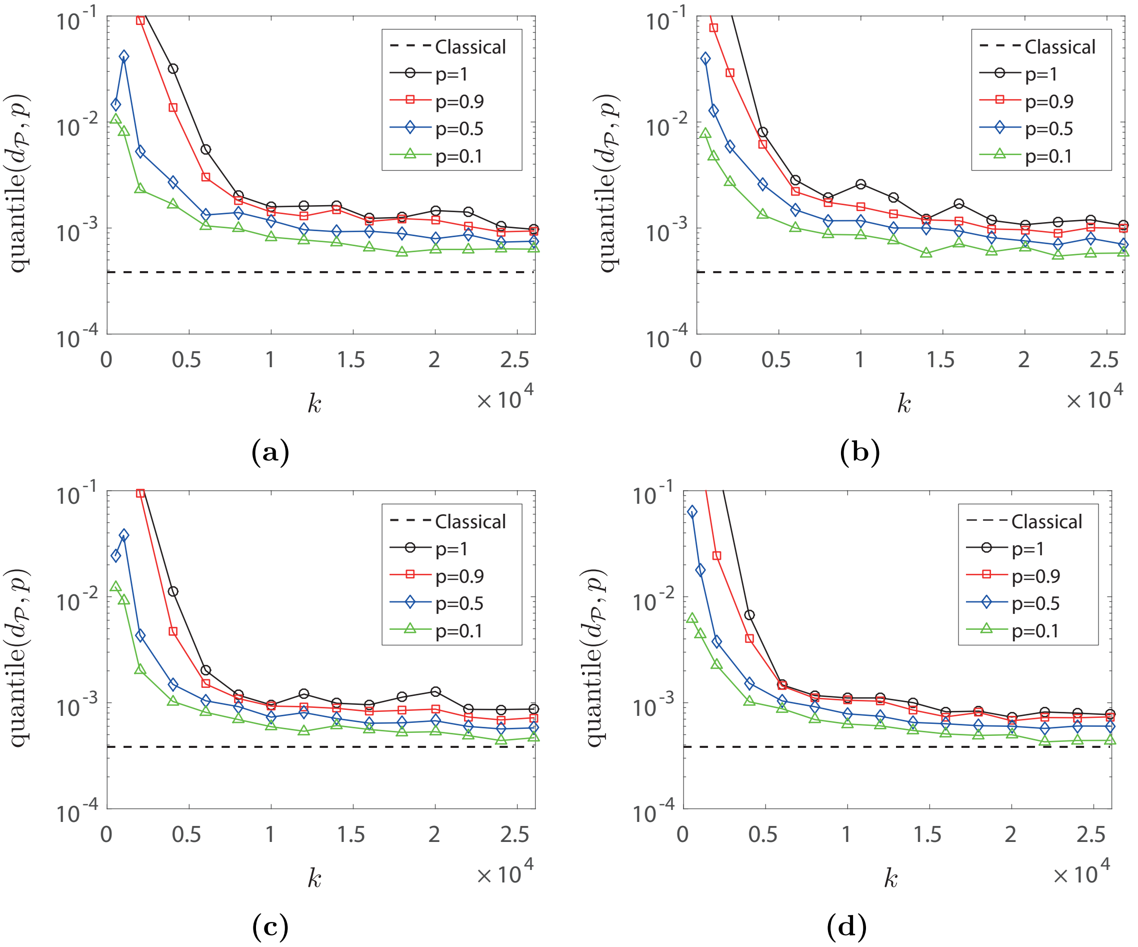
(c)

(d)

Figure 2.8: Error $e_{\mathcal{P}}$ and residual error $\Delta_{\mathcal{P}}$ of the classical Galerkin projection and quantiles of probabilities $p = 1, 0.9, 0.5$ and 0.1 over 20 samples of $e_{\mathcal{P}}$ and $\Delta_{\mathcal{P}}$ of the randomized Galerkin projection versus the number of rows of $\mathbf{\Omega}$. (a) Exact error $e_{\mathcal{P}}$, with rescaled Gaussian distribution as $\mathbf{\Omega}$. (b) Exact error $e_{\mathcal{P}}$, with P-SRHT matrix as $\mathbf{\Omega}$. (c) Residual error $\Delta_{\mathcal{P}}$, with rescaled Gaussian distribution as $\mathbf{\Omega}$. (d) Residual error $\Delta_{\mathcal{P}}$, with P-SRHT matrix as $\mathbf{\Omega}$.

as the classical one. Therefore, for such values of k the error depends mainly on the precision of the approximate inner product for $s_r^{\mathrm{pd}}(\mu)$. Unlike in the thermal block problem (see Figure 2.2), in this experiment the quality of the classical method is attained by $s_r(\mu) = s_r^{\mathrm{spd}}(\mu)$ with $k \ll n$. Consequently, the benefit of employing the improved correction $s_r^{\mathrm{spd+}}(\mu)$ here is not as evident as in the previous numerical example. This experiment only proves that the error associated with approximation of the inner product for $s_r^{\mathrm{pd}}(\mu)$ does not depend on the condition number and the dimension of the operator.

Randomized greedy algorithm. Finally, we performed $r = 150$ iterations of the



(a) (b) (c) (d)

Figure 2.9: The error $d_\mathcal{P}$ of the classical output quantity and quantiles of probabilities $p = 1, 0.9, 0.5$ and 0.1 over 20 samples of $d_\mathcal{P}$ of the output quantities computed with random sketching versus the number of rows of $\mathbf{\Omega}$. (a) The errors of the classical $s_r(\mu)$ and the randomized $s_r(\mu)$ with Gaussian matrix as $\mathbf{\Omega}$. (b) The errors of the classical $s_r(\mu)$ and the randomized $s_r(\mu)$ with P-SRHT distribution as $\mathbf{\Omega}$. (c) The errors of the classical $s_r(\mu)$ and $s_r^{\mathrm{spd+}}(\mu)$ with Gaussian matrix as $\mathbf{\Omega}$ and $W_r^{\mathrm{du}} := U_i^{\mathrm{du}}$, $i^{\mathrm{du}} = 30$. (d) The errors of the classical $s_r(\mu)$ and $s_r^{\mathrm{spd+}}(\mu)$ with P-SRHT distribution as $\mathbf{\Omega}$ and $W_r^{\mathrm{du}} := U_i^{\mathrm{du}}$, $i^{\mathrm{du}} = 30$.

classical greedy algorithm (see Section 2.2.4) and its randomized version (Algorithm 3) using different distributions and sizes for $\mathbf{\Omega}$, and a Gaussian random matrix with $k' = 200$ rows for $\mathbf{\Gamma}$. As in the thermal block benchmark, the error at i-th iteration is measured by $\Delta_\mathcal{P} := \max_{\mu \in \mathcal{P}_{\mathrm{train}}} \|\mathbf{r}(\mathbf{u}_i(\mu);\mu)\|_{U'} / \max_{\mu \in \mathcal{P}_{\mathrm{train}}} \|\mathbf{b}(\mu)\|_{U'}$. For $k = 1000$ we reveal poor performance of Algorithm 3 (see Figure 2.10). It can be explained by the fact that for such size of $\mathbf{\Omega}$ the randomized Galerkin projection has low accuracy. For $k = 20000$, however, the convergence of the classical greedy algorithm is fully preserved.

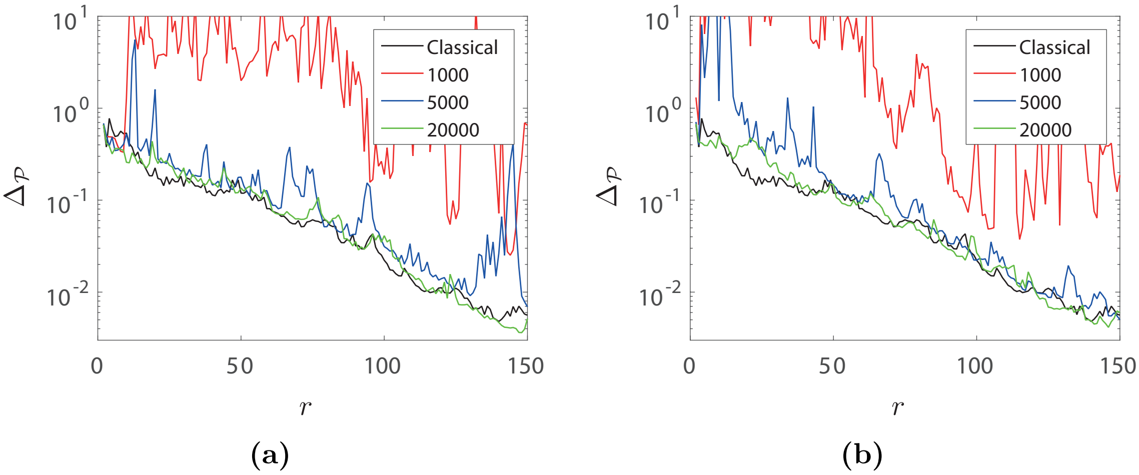


Figure 2.10: Convergences of the classical greedy algorithm (see Section 2.2.4) and its efficient randomized version (Algorithm 3) using $\mathbf{\Omega}$ drawn from (a) Gaussian distribution or (b) P-SRHT distribution.

Comparison of computational costs. Even though the size of $\mathbf{\Omega}$ has to be considered larger than for the thermal block problem, our methodology still yields considerable reduction of the computational costs compared to the classical approach. The implementation was carried out in Matlab® R2015a with an external `C++` function for the fast Walsh-Hadamard transform. Our codes were not designed for a specific problem but rather for a generic multi-query MOR. The algorithms were executed on an Intel® Core$^{\text{TM}}$ i7-7700HQ 2.8GHz CPU, with 16.0GB RAM memory.

Let us start with validation of the computational cost reduction of the greedy algorithm. In Table 2.2 we provide the runtimes of the classical greedy algorithm and Algorithm 3 employing $\mathbf{\Omega}$ drawn from the P-SRHT distribution with $k = 20000$ rows. In Table 2.2 the computations are divided into three basic categories: computing the snapshots (samples of the solution), precomputing the affine expansions for the online solver, and finding $\mu^{i+1} \in \mathcal{P}_{\text{train}}$ which maximizes the error indicator with a provisional online solver. The first category includes evaluation of $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$ using their affine expansions and solving the systems with a built in Matlab® linear solver. The second category consists of evaluating the random sketch in Algorithm 3; evaluating high-dimensional matrix-vector products and inner products for the Galerkin projection; evaluating high-dimensional matrix-vector products and inner products for the error estimation; and the remaining computations, such as precomputing a decomposition of $\mathbf{R}_U$, memory allocations, orthogonalization of the basis, etc. In its turn, the third category of computations includes generating $\mathbf{\Gamma}$ and evaluating the affine factors of $\mathbf{V}_i^{\mathbf{\Phi}}(\mu)$ and $\mathbf{b}^{\mathbf{\Phi}}(\mu)$ from the affine factors of $\mathbf{V}_i^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ at each iteration of Algorithm 3; evaluating the reduced systems from the precomputed affine expansions and solving them with a built in Matlab® linear solver, for all $\mu \in \mathcal{P}_{\text{train}}$, at each iteration; evaluating the residual terms from the affine expansions and using them to evaluate the residual errors of the Galerkin projections, for all $\mu \in \mathcal{P}_{\text{train}}$, at each iteration.

We observe that evaluating the snapshots occupied only 6% of the overall runtime of the classical greedy algorithm. The other 94% could be subject to reduction with the random sketching technique. Due to operating on a large training set, the cost of solving (including estimation of the error) reduced order models on $\mathcal{P}_{\text{train}}$ has a considerable impact on the runtimes of both classical and randomized algorithms. This cost, however, is independent of the dimension of the full system of equations and will become negligible for larger problems. Nevertheless, for $r = 150$ the randomized procedure for error estimation (see Section 2.4.5) yielded reduction of the aforementioned cost in about 2 times. As expected, in the classical method the most expensive computations are numerous evaluations of high-dimensional matrix-vector and inner products. For large problems these computations can become a bottleneck of an algorithm. Their cost reduction by random sketching is drastic. We observe that for the classical algorithm the corresponding runtime grows quadratically with r while for the randomized algorithm it grows only linearly. The cost of this step for $r = 150$ iterations of the greedy algorithm was divided 15. In addition, random sketching helped to reduce memory consumption. The memory required by $r = 150$ iterations of the greedy algorithm has been reduced from 6.1GB (including storage of affine factors of $\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_i$) to only 1GB, from which 0.4GB is meant for the initialization, i.e., defining the discrete problem, precomputing the decomposition of $\mathbf{R}_U$, etc.

Table 2.2: The CPU times in seconds taken by each type of computations in the classical greedy algorithm (see Section 2.2.4) and the randomized greedy algorithm (Algorithm 3).

Category	Computations	Classical			Randomized		
		$r = 50$	$r = 100$	$r = 150$	$r = 50$	$r = 100$	$r = 150$
snapshots		143	286	430	143	287	430
high-dimensional matrix-vector & inner products	sketch	–	–	–	54	113	177
	Galerkin	59	234	525	3	14	31
	error	405	1560	3444	–	–	–
	remaining	27	196	236	7	28	67
	total	491	1899	4205	64	154	275
provisional online solver	sketch	–	–	–	56	127	216
	Galerkin	46	268	779	50	272	783
	error	45	522	2022	43	146	407
	total	91	790	2801	149	545	1406

The improvement of the efficiency of the online stage can be validated by comparing the CPU times of the provisional online solver in the greedy algorithms.

Table 2.2 presents the CPU times taken by the provisional online solver at the i-th iteration of the classical and the sketched greedy algorithms, where the solver is used for efficient computation of the reduced models associated with an i-dimensional approximation space U_i for all parameter's values from the training set. These computations consist of evaluating the reduced systems from the affine expansions and their solutions with the Matlab® linear solver, and computing residual-based error estimates using (2.38) for the classical method or (2.41) for the estimation with random sketching. Moreover, the sketched online stage also involves generation of $\mathbf{\Gamma}$ and computing $\mathbf{V}_i^{\mathbf{\Phi}}(\mu)$ and $\mathbf{b}^{\mathbf{\Phi}}(\mu)$ from the affine factors of $\mathbf{V}_i^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$. Note that random sketching can reduce the online complexity (and improve the stability) associated with residual-based error estimation. The online cost of computation of a solution, however, remains the same for both the classical and the sketched methods. Table 2.3 reveals that for this benchmark the speedups in the online stage are achieved for $i \geq 50$. The computational cost of the error estimation using the classical approach grows quadratically with i, while using the randomized procedure, it grows only linearly. For $i = 150$ we report a reduction of the runtime for error estimation by a factor 5 and a reduction of the total runtime by a factor 2.6.

Table 2.3: The CPU times in seconds taken by each type of computations of the classical and the efficient sketched provisional online solvers during the i-th iteration of the greedy algorithms.

Computations	Classical			Randomized		
	$i = 50$	$i = 100$	$i = 150$	$i = 50$	$i = 100$	$i = 150$
sketch	–	–	–	1.3	1.5	2
Galerkin	2	7	13.5	2.3	7	14
error	2.8	18	45.2	1.3	3.1	7
total	4.8	24.9	58.7	4.8	11.6	22.8

The benefit of using random sketching methods for POD is validated in the context of distributed or limited-memory environments, where the snapshots are computed on distributed workstations or when the storage of snapshots requires too much RAM. For these scenarios the efficiency is characterized by the amount of communication or storage needed for constructing a reduced model. Let us recall that the classical POD requires maintaining and operating with the full basis matrix $\mathbf{U}_m$, while the sketched POD requires the precomputation of a $\mathbf{\Theta}$-sketch of $\mathbf{U}_m$ and then constructs a reduced model from the sketch. In particular, for distributed computing a random sketch of each snapshot should be computed on a separate machine and then efficiently transfered to the master workstation for post-processing. For this experiment, Gaussian matrices of different sizes were tested for $\mathbf{\Omega}$. A seeded random number generator was used for maintaining $\mathbf{\Omega}$ with negligible computational costs. In Table 2.4 we provide the amount of storage needed to maintain a sketch of a single snapshot, which also reflects the required communication for its transfer to the master workstation in the distributed computational environment. We observe that random sketching methods yielded computational costs reductions when $k \leq 17000$. It follows that for $k = 10000$ a $\mathbf{\Theta}$-sketch of a snapshot consumes 1.7 times less memory than the full snapshot. Yet, for $m = \#\mathcal{P}_{\text{train}} \leq 10000$ and $r \leq 150$, the sketched method of snapshots (see Definition 2.5.3) using $\mathbf{\Omega}$ of size $k = 10000$ provides almost optimal approximation of the training set of snapshots with an error which is only at most 1.1 times higher than the error associated with the classical POD approximation. A Gaussian matrix of size $k = 10000$, for $r \leq 150$, also yields with high probability very accurate estimation (up to a factor of 1.1) of the residual error and sufficiently accurate estimation of the Galerkin projection (increasing the residual error by at most a factor of 1.66). For coercive and well-conditioned problems such as the thermal-block benchmark, it can be sufficient to use much smaller sketching matrices than in the present benchmark, say with $k = 2500$ rows. Moreover, this value for k should be pertinent also for ill-conditioned problems, including the considered acoustic cloak benchmark, when the minimal residual methods are used alternatively to the Galerkin methods (see Chapter 3). From Table 2.4 it follows that a random sketch of dimension $k = 2500$ is 6.8 times cheaper to maintain than a full snapshot vector. It has to be mentioned that when the sketch is computed from the affine expansion of $\mathbf{A}(\mu)$ with m_A terms (here $m_A = 11$), its maintenance/transfer costs are proportional to km_A and are independent of the dimension of the initial system of equations. Consequently, for problems with larger n/m_A a better cost reduction is expected.

Table 2.4: The amount of data in megabytes required to maintain/transfer a single snapshot or its $\mathbf{\Theta}$-sketch for post-processing.

full snapshot	$k = 2500$	$k = 5000$	$k = 10000$	$k = 15000$	$k = 17000$	$k = 20000$
1.64	0.24	0.48	0.96	1.44	1.63	1.92

2.7 Conclusions and perspectives

In this chapter we proposed a methodology for reducing the cost of classical projection-based MOR methods such as RB method and POD. The computational cost of constructing a reduced order model is essentially reduced to evaluating the samples (snapshots) of the solution on the training set, which in its turn can be efficiently performed with state-of-the-art routine on a powerful server or distributed machines. Our approach can be beneficial in any computational environment. It improves efficiency of classical MOR methods in terms of complexity (number of flops), memory consumption, scalability, communication cost between distributed machines, etc.

Unlike classical methods, our method does not require maintaining and operating with high-dimensional vectors. Instead, the reduced order model is constructed from a random sketch (a set of random projections), with a negligible computational cost. A new framework was introduced in order to adapt the random sketching technique to the context of MOR. We interpret random sketching as a random estimation of inner products between high-dimensional vectors. The projections are obtained with random matrices (called oblivious subspace embeddings), which are efficient to store and to multiply by. We introduced oblivious subspace embeddings for a general inner product defined by a self-adjoint positive definite matrix. Thereafter, we introduced randomized versions of Galerkin projection, residual-based error estimation, and primal-dual correction. The conditions for preserving the quality of the output of the classical method were provided. In addition, we discussed computational aspects for an efficient evaluation of a random sketch in different computational environments, and introduced a new procedure for estimating the residual norm. This procedure is not only efficient but also is less sensitive to round-off errors than the classical approach. Finally, we proposed randomized versions of POD and greedy algorithm for RB. Again, in both algorithms, standard operations are performed only on the sketch but not on high-dimensional vectors.

The methodology has been validated in a series of numerical experiments. We observed that indeed random sketching can provide a drastic reduction of the computational cost. The experiments revealed that the theoretical bounds for the sizes of random matrices are pessimistic. In practice, it can be pertinent to use much smaller matrices. In such a case it is important to provide a posteriori certification of the solution. In addition, it can be helpful to have an indicator of the accuracy of random sketching, which can be used for an adaptive selection of the random matrices' sizes. The aforementioned issues are addressed in Chapter 3 of this manuscript. It was also observed that the performance of random sketching for estimating the Galerkin projection depends on the operator's properties (more precisely on the constant $a_r(\mu)$ defined in Proposition 2.4.2). Consequently, the accuracy of the output can degrade considerably for problems with ill-conditioned operators. A remedy is to replace Galerkin projection by another type of projection for the approximation of $\mathbf{u}(\mu)$ (and $\mathbf{u}^{\mathrm{du}}(\mu)$). The randomized minimal residual projection proposed in Chapter 3 preserves the quality of the classical minimal residual projection regardless of the operator's properties. Another remedy would be to improve the condition number of $\mathbf{A}(\mu)$ with an affine parameter-dependent preconditioner constructed with an approach from Chapter 4. We also have seen that preserving a high precision for the sketched primal-dual correction (2.33) can require large sketching matrices. A way to overcome this issue was proposed. It consists in obtaining an efficient approximation $\mathbf{w}_r^{\mathrm{du}}(\mu)$ of the solution $\mathbf{u}_r^{\mathrm{du}}(\mu)$ (or $\mathbf{u}_r(\mu)$). Such $\mathbf{w}_r^{\mathrm{du}}(\mu)$ can be also used for reducing the cost of extracting the quantity of interest from $\mathbf{u}_r(\mu)$, i.e., computing $\mathbf{l}_r(\mu)$, which in general can be expensive (but was assumed to have a negligible cost). In addition, this approach can be used for problems with nonlinear quantities of interest.

An approximation $\mathbf{w}_r^{\mathrm{du}}(\mu)$ can be taken as a projection of $\mathbf{u}_r^{\mathrm{du}}(\mu)$ (or $\mathbf{u}_r(\mu)$) on a subspace W_r^{du}. In the experiments W_r^{du} was constructed from the first several basis vectors of the approximation space U_r^{du}. A better subspace can be obtained by approximating the manifold $\{\mathbf{u}_r^{\mathrm{du}}(\mu) : \mu \in \mathcal{P}_{\mathrm{train}}\}$ with a greedy algorithm or POD. Here, random sketching can be again employed for improving the efficiency. The strategies for obtaining both accurate and efficient W_r^{du} with random sketching are discussed in details in Chapter 3.

2.8 Appendix

Here we list the proofs of propositions and theorems from the chapter.

Proof of Proposition 2.2.2 (modified Cea's lemma). For all $\mathbf{x} \in U_r$, it holds

$$\begin{aligned}\alpha_r(\mu)\|\mathbf{u}_r(\mu)-\mathbf{x}\|_U &\leq \|\mathbf{r}(\mathbf{u}_r(\mu);\mu)-\mathbf{r}(\mathbf{x};\mu)\|_{U_r'} \leq \|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U_r'}+\|\mathbf{r}(\mathbf{x};\mu)\|_{U_r'}\\ &= \|\mathbf{r}(\mathbf{x};\mu)\|_{U_r'} \leq \beta_r(\mu)\|\mathbf{u}(\mu)-\mathbf{x}\|_U,\end{aligned}$$

where the first and last inequalities directly follow from the definitions of $\alpha_r(\mu)$ and $\beta_r(\mu)$, respectively. Now,

$$\|\mathbf{u}(\mu)-\mathbf{u}_r(\mu)\|_U \leq \|\mathbf{u}(\mu)-\mathbf{x}\|_U+\|\mathbf{u}_r(\mu)-\mathbf{x}\|_U \leq \|\mathbf{u}(\mu)-\mathbf{x}\|_U+\frac{\beta_r(\mu)}{\alpha_r(\mu)}\|\mathbf{u}(\mu)-\mathbf{x}\|_U,$$

which completes the proof. □

Proof of Proposition 2.2.3. For all $\mathbf{a} \in \mathbb{K}^r$ and $\mathbf{x} := \mathbf{U}_r\mathbf{a}$, it holds

$$\begin{aligned}\frac{\|\mathbf{A}_r(\mu)\mathbf{a}\|}{\|\mathbf{a}\|} &= \max_{\mathbf{z}\in\mathbb{K}^r\setminus\{\mathbf{0}\}}\frac{|\langle\mathbf{z},\mathbf{A}_r(\mu)\mathbf{a}\rangle|}{\|\mathbf{z}\|\|\mathbf{a}\|} = \max_{\mathbf{z}\in\mathbb{K}^r\setminus\{\mathbf{0}\}}\frac{|\mathbf{z}^{\mathrm{H}}\mathbf{U}_r^{\mathrm{H}}\mathbf{A}(\mu)\mathbf{U}_r\mathbf{a}|}{\|\mathbf{z}\|\|\mathbf{a}\|}\\ &= \max_{\mathbf{y}\in U_r\setminus\{\mathbf{0}\}}\frac{|\mathbf{y}^{\mathrm{H}}\mathbf{A}(\mu)\mathbf{x}|}{\|\mathbf{y}\|_U\|\mathbf{x}\|_U} = \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'}}{\|\mathbf{x}\|_U}.\end{aligned}$$

Then the proposition follows directly from definitions of $\alpha_r(\mu)$ and $\beta_r(\mu)$. □

Proof of Proposition 2.2.4. We have

$$\begin{aligned}|s(\mu)-s_r^{\mathrm{pd}}(\mu)| &= |s(\mu)-s_r(\mu)+\langle\mathbf{u}_r^{\mathrm{du}}(\mu),\mathbf{r}(\mathbf{u}_r(\mu);\mu)\rangle|\\ &= |\langle\mathbf{l}(\mu),\mathbf{u}(\mu)-\mathbf{u}_r(\mu)\rangle+\langle\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{u}_r^{\mathrm{du}}(\mu),\mathbf{u}(\mu)-\mathbf{u}_r(\mu)\rangle|\\ &= |\langle\mathbf{r}^{\mathrm{du}}(\mathbf{u}_r^{\mathrm{du}}(\mu);\mu),\mathbf{u}(\mu)-\mathbf{u}_r(\mu)\rangle|\\ &\leq \|\mathbf{r}^{\mathrm{du}}(\mathbf{u}_r^{\mathrm{du}}(\mu);\mu)\|_{U'}\|\mathbf{u}(\mu)-\mathbf{u}_r(\mu)\|_U,\end{aligned}$$

and the result follows from definition (2.10). □

Proof of Proposition 2.2.5. To prove the first inequality we notice that $\mathbf{Q}\mathbf{P}_{U_r}\mathbf{U}_m$ has rank at most r. Consequently,

$$\|\mathbf{Q}\mathbf{U}_m-\mathbf{B}_r^*\|_F^2 \leq \|\mathbf{Q}\mathbf{U}_m-\mathbf{Q}\mathbf{P}_{U_r}\mathbf{U}_m\|_F^2 = \sum_{i=1}^m\|\mathbf{u}(\mu^i)-\mathbf{P}_{U_r}\mathbf{u}(\mu^i)\|_U^2.$$

For the second inequality let us denote the i-th column vector of $\mathbf{B}_r$ by $\mathbf{b}_i$. Since $\mathbf{Q}\mathbf{R}_U^{-1}\mathbf{Q}^{\mathrm{H}} = \mathbf{Q}\mathbf{Q}^\dagger$, with $\mathbf{Q}^\dagger$ the pseudo-inverse of $\mathbf{Q}$, is the orthogonal projection onto range($\mathbf{Q}$), we have

$$\begin{aligned}\|\mathbf{Q}\mathbf{U}_m-\mathbf{B}_r\|_F^2 &\geq \|\mathbf{Q}\mathbf{R}_U^{-1}\mathbf{Q}^{\mathrm{H}}(\mathbf{Q}\mathbf{U}_m-\mathbf{B}_r)\|_F^2 = \sum_{i=1}^m\|\mathbf{u}(\mu^i)-\mathbf{R}_U^{-1}\mathbf{Q}^{\mathrm{H}}\mathbf{b}_i\|_U^2\\ &\geq \sum_{i=1}^m\|\mathbf{u}(\mu^i)-\mathbf{P}_{U_r}\mathbf{u}(\mu^i)\|_U^2.\end{aligned}$$

□

Proof of Proposition 2.3.3. It is clear that $\langle \cdot, \cdot \rangle_X^{\mathbf{\Theta}}$ and $\langle \cdot, \cdot \rangle_{X'}^{\mathbf{\Theta}}$ satisfy (conjugate) symmetry, linearity and positive semi-definiteness properties. The definitenesses of $\langle \cdot, \cdot \rangle_X^{\mathbf{\Theta}}$ and $\langle \cdot, \cdot \rangle_{X'}^{\mathbf{\Theta}}$ on Y and Y', respectively, follow directly from Definition 2.3.1 and Corollary 2.3.2. □

Proof of Proposition 2.3.4. Using Definition 2.3.1, we have

$$\begin{aligned}\|\mathbf{y}'\|_{Z'}^{\mathbf{\Theta}} &= \max_{\mathbf{x}\in Z\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{R}_X^{-1}\mathbf{y}', \mathbf{x}\rangle_X^{\mathbf{\Theta}}|}{\|\mathbf{x}\|_X^{\mathbf{\Theta}}} \leq \max_{\mathbf{x}\in Z\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{R}_X^{-1}\mathbf{y}', \mathbf{x}\rangle_X| + \varepsilon \|\mathbf{y}'\|_{X'}\|\mathbf{x}\|_X}{\|\mathbf{x}\|_X^{\mathbf{\Theta}}} \\ &\leq \max_{\mathbf{x}\in Z\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{R}_X^{-1}\mathbf{y}', \mathbf{x}\rangle_X| + \varepsilon \|\mathbf{y}'\|_{X'}\|\mathbf{x}\|_X}{\sqrt{1-\varepsilon}\|\mathbf{x}\|_X} \\ &\leq \frac{1}{\sqrt{1-\varepsilon}}\left(\max_{\mathbf{x}\in Z\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{y}', \mathbf{x}\rangle|}{\|\mathbf{x}\|_X} + \varepsilon\|\mathbf{y}'\|_{X'}\right),\end{aligned}$$

which yields the right inequality. To prove the left inequality we assume that $\|\mathbf{y}'\|_{Z'} - \varepsilon\|\mathbf{y}'\|_{X'} \geq 0$. Otherwise the relation is obvious because $\|\mathbf{y}'\|_{Z'}^{\mathbf{\Theta}} \geq 0$. By Definition 2.3.1,

$$\begin{aligned}\|\mathbf{y}'\|_{Z'}^{\mathbf{\Theta}} &= \max_{\mathbf{x}\in Z\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{R}_X^{-1}\mathbf{y}', \mathbf{x}\rangle_X^{\mathbf{\Theta}}|}{\|\mathbf{x}\|_X^{\mathbf{\Theta}}} \geq \max_{\mathbf{x}\in Z\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{R}_X^{-1}\mathbf{y}', \mathbf{x}\rangle_X| - \varepsilon \|\mathbf{y}'\|_{X'}\|\mathbf{x}\|_X}{\|\mathbf{x}\|_X^{\mathbf{\Theta}}} \\ &\geq \max_{\mathbf{x}\in Z\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{R}_X^{-1}\mathbf{y}', \mathbf{x}\rangle_X| - \varepsilon \|\mathbf{y}'\|_{X'}\|\mathbf{x}\|_X}{\sqrt{1+\varepsilon}\|\mathbf{x}\|_X} \\ &\geq \frac{1}{\sqrt{1+\varepsilon}}\left(\max_{\mathbf{x}\in Z\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{y}', \mathbf{x}\rangle|}{\|\mathbf{x}\|_X} - \varepsilon\|\mathbf{y}'\|_{X'}\right),\end{aligned}$$

which completes the proof. □

Proof of Proposition 2.3.7. Let us start with the case $\mathbb{K} = \mathbb{R}$. For the proof we shall follow standard steps (see, e.g., [157, Section 2.1]). Given a d-dimensional subspace $V \subseteq \mathbb{R}^n$, let $\mathcal{S} = \{\mathbf{x} \in V : \ \|\mathbf{x}\| = 1\}$ be the unit sphere of V. According to [30, Lemma 2.4], for any $\gamma > 0$ there exists a γ-net $\mathcal{N}$ of $\mathcal{S}$[6] satisfying $\#\mathcal{N} \leq (1+2/\gamma)^d$. For η such that $0 < \eta < 1/2$, let $\mathbf{\Theta} \in \mathbb{R}^{k\times n}$ be a rescaled Gaussian or Rademacher matrix with $k \geq 6\eta^{-2}(2d\log(1+2/\gamma) + \log(1/\delta))$. By [2, Lemmas 4.1 and 5.1] and an union bound argument we obtain for a fixed $\mathbf{x} \in V$

$$\mathbb{P}(|\|\mathbf{x}\|^2 - \|\mathbf{\Theta}\mathbf{x}\|^2| \leq \eta\|\mathbf{x}\|^2) \geq 1 - 2\exp(-k\eta^2/6).$$

Consequently, using a union bound for the probability of success, we have that

$$\left\{ \ |\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{\Theta}(\mathbf{x}+\mathbf{y})\|^2| \leq \eta\|\mathbf{x}+\mathbf{y}\|^2, \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{N} \right\},$$

[6] We have $\forall \mathbf{x} \in \mathcal{S}, \exists \mathbf{y} \in \mathcal{N}$ such that $\|\mathbf{x} - \mathbf{y}\| \leq \gamma$.

holds with probability at least $1-\delta$. Then we deduce that

$$\{\ |\langle \mathbf{x},\mathbf{y}\rangle - \langle \mathbf{\Theta x},\mathbf{\Theta y}\rangle| \leq \eta, \quad \forall \mathbf{x},\mathbf{y} \in \mathcal{N}\} \tag{2.52}$$

holds with probability at least $1-\delta$. Now, let $\mathbf{n}$ be some vector in $\mathcal{S}$. Assuming $\gamma < 1$, it can be proven by induction that $\mathbf{n} = \sum_{i\geq 0} \alpha_i \mathbf{n}_i$, where $\mathbf{n}_i \in \mathcal{N}$ and $0 \leq \alpha_i \leq \gamma^i$[7]. If (2.52) is satisfied, then

$$\begin{aligned}\|\mathbf{\Theta n}\|^2 &= \sum_{i,j\geq 0} \langle \mathbf{\Theta n}_i, \mathbf{\Theta n}_j\rangle \alpha_i\alpha_j \\ &\leq \sum_{i,j\geq 0} (\langle \mathbf{n}_i,\mathbf{n}_j\rangle \alpha_i\alpha_j + \eta\alpha_i\alpha_j) = 1+\eta(\sum_{i\geq 0}\alpha_i)^2 \leq 1+\frac{\eta}{(1-\gamma)^2},\end{aligned}$$

and similarly $\|\mathbf{\Theta n}\|^2 \geq 1-\frac{\eta}{(1-\gamma)^2}$. Therefore, if (2.52) is satisfied, we have

$$|1-\|\mathbf{\Theta n}\|^2| \leq \eta/(1-\gamma)^2. \tag{2.53}$$

For a given $\varepsilon \leq 0.5/(1-\gamma)^2$, let $\eta = (1-\gamma)^2\varepsilon$. Since (2.53) holds for an arbitrary vector $\mathbf{n} \in \mathcal{S}$, using the parallelogram identity, we easily obtain that

$$|\langle \mathbf{x},\mathbf{y}\rangle - \langle \mathbf{\Theta x},\mathbf{\Theta y}\rangle| \leq \varepsilon\|\mathbf{x}\|\|\mathbf{y}\| \tag{2.54}$$

holds for all $\mathbf{x},\mathbf{y} \in V$ if (2.52) is satisfied. We conclude that if $k \geq 6\varepsilon^{-2}(1-\gamma)^{-4}(2d\log(1+2/\gamma)+\log(1/\delta))$ then $\mathbf{\Theta}$ is a $\ell_2 \to \ell_2$ ε-subspace embedding for V with probability at least $1-\delta$. The lower bound for the number of rows of $\mathbf{\Theta}$ is obtained by taking $\gamma = \arg\min_{x\in(0,1)}(\log(1+2/x)/(1-x)^4) \approx 0.0656$.

The statement of the proposition for the case $\mathbb{K} = \mathbb{C}$ can be deduced from the fact that if $\mathbf{\Theta}$ is $(\varepsilon,\delta,2d)$ oblivious $\ell_2 \to \ell_2$ subspace embedding for $\mathbb{K} = \mathbb{R}$, then it is (ε,δ,d) oblivious $\ell_2 \to \ell_2$ subspace embedding for $\mathbb{K} = \mathbb{C}$. A detailed proof of this fact is provided in the supplementary material. To show this we first note that the real part and the imaginary part of any vector from a d-dimensional subspace $V^* \subseteq \mathbb{C}^n$ belong to a certain subspace $W \subseteq \mathbb{R}^n$ with $\dim(W) \leq 2d$. Further, one can show that if $\mathbf{\Theta}$ is $\ell_2 \to \ell_2$ ε-subspace embedding for W, then it is $\ell_2 \to \ell_2$ ε-subspace embedding for V^*. □

Proof of Proposition 2.3.9. Let $\mathbf{\Theta} \in \mathbb{R}^{k\times n}$ be a P-SRHT matrix, let V be an arbitrary d-dimensional subspace of $\mathbb{K}^n$, and let $\mathbf{V} \in \mathbb{K}^{n\times d}$ be a matrix whose columns form an orthonormal basis of V. Recall, $\mathbf{\Theta}$ is equal to the first n columns of matrix $\mathbf{\Theta}^* = k^{-1/2}(\mathbf{RH}_s\mathbf{D}) \in \mathbb{R}^{k\times s}$. Next we shall use the fact that for any orthonormal matrix $\mathbf{V}^* \in \mathbb{K}^{s\times d}$, all singular values of a matrix $\mathbf{\Theta}^*\mathbf{V}^*$ belong to the interval

[7] Indeed, $\exists \mathbf{n}_0 \in \mathcal{N}$ such that $\|\mathbf{n}-\mathbf{n}_0\| := \alpha_1 \leq \gamma$. Let $\alpha_0 = 1$. Then assuming that $\|\mathbf{n}-\sum_{i=0}^{m}\alpha_i\mathbf{n}_i\| := \alpha_{m+1} \leq \gamma^{m+1}$, $\exists \mathbf{n}_{m+1} \in \mathcal{N}$ such that $\|\frac{1}{\alpha_{m+1}}(\mathbf{n}-\sum_{i=0}^{m}\alpha_i\mathbf{n}_i)-\mathbf{n}_{m+1}\| \leq \gamma \implies \|\mathbf{n}-\sum_{i=0}^{m+1}\alpha_i\mathbf{n}_i\| \leq \alpha_{m+1}\gamma \leq \gamma^{m+2}$.

$[\sqrt{1-\varepsilon}, \sqrt{1+\varepsilon}]$ with probability at least $1-\delta$. This result is basically a restatement of [31, Lemma 4.1] and [146, Theorem 3.1] including the complex case and with improved constants. It can be shown to hold by mimicking the proof in [146] with a few additional algebraic operations. For a detailed proof of the statement, see the supplementary material.

By taking $\mathbf{V}^*$ with the first $n \times d$ block equal to $\mathbf{V}$ and zeros elsewhere, and using the fact that $\mathbf{\Theta V}$ and $\mathbf{\Theta}^* \mathbf{V}^*$ have the same singular values, we obtain that

$$||\|\mathbf{Vz}\|^2 - \|\mathbf{\Theta V z}\|^2| = |\mathbf{z}^{\mathrm{H}}(\mathbf{I} - \mathbf{V}^{\mathrm{H}}\mathbf{\Theta}^{\mathrm{H}}\mathbf{\Theta V})\mathbf{z}| \leq \varepsilon \|\mathbf{z}\|^2 = \varepsilon \|\mathbf{Vz}\|^2, \quad \forall \mathbf{z} \in \mathbb{K}^d \tag{2.55}$$

holds with probability at least $1-\delta$. Using the parallelogram identity, it can be easily proven that relation (2.55) implies

$$|\langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{\Theta x}, \mathbf{\Theta y} \rangle| \leq \varepsilon \|\mathbf{x}\| \|\mathbf{y}\|, \quad \forall \mathbf{x}, \mathbf{y} \in V.$$

We conclude that $\mathbf{\Theta}$ is a (ε, δ, d) oblivious $\ell_2 \to \ell_2$ subspace embedding. □

Proof of Proposition 2.3.11. Let V be any d-dimensional subspace of X and let $V^* := \{\mathbf{Qx} : \mathbf{x} \in V\}$. Since the following relations hold $\langle \cdot, \cdot \rangle_U = \langle \mathbf{Q}\cdot, \mathbf{Q}\cdot \rangle$ and $\langle \cdot, \cdot \rangle_U^{\mathbf{\Theta}} = \langle \mathbf{Q}\cdot, \mathbf{Q}\cdot \rangle_2^{\mathbf{\Omega}}$, we have that sketching matrix $\mathbf{\Theta}$ is an ε-embedding for V if and only if $\mathbf{\Omega}$ is an ε-embedding for V^*. It follows from the definition of $\mathbf{\Omega}$ that this matrix is an ε-embedding for V^* with probability at least $1-\delta$, which completes the proof.□

Proof of Proposition 2.4.1 (sketched Cea's lemma). The proof exactly follows the one of Proposition 2.2.2 with $\|\cdot\|_{U_r'}$ replaced by $\|\cdot\|_{U_r'}^{\mathbf{\Theta}}$. □

Proof of Proposition 2.4.2. According to Proposition 2.3.4, and by definition of $a_r(\mu)$, we have

$$\begin{aligned}
\alpha_r^{\mathbf{\Theta}}(\mu) = \min_{\mathbf{x} \in U_r \backslash \{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'}^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U} &\geq \frac{1}{\sqrt{1+\varepsilon}} \min_{\mathbf{x} \in U_r \backslash \{\mathbf{0}\}} \frac{(\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'} - \varepsilon \|\mathbf{A}(\mu)\mathbf{x}\|_{U'})}{\|\mathbf{x}\|_U} \\
&\geq \frac{1}{\sqrt{1+\varepsilon}} (1 - \varepsilon a_r(\mu)) \min_{\mathbf{x} \in U_r \backslash \{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'}}{\|\mathbf{x}\|_U}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\beta_r^{\mathbf{\Theta}}(\mu) &= \max_{\mathbf{x} \in (\mathrm{span}\{\mathbf{u}(\mu)\} + U_r) \backslash \{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'}^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U} \\
&\leq \frac{1}{\sqrt{1-\varepsilon}} \max_{\mathbf{x} \in (\mathrm{span}\{\mathbf{u}(\mu)\} + U_r) \backslash \{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'} + \varepsilon \|\mathbf{A}(\mu)\mathbf{x}\|_{U'}}{\|\mathbf{x}\|_U} \\
&\leq \frac{1}{\sqrt{1-\varepsilon}} \left(\max_{\mathbf{x} \in (\mathrm{span}\{\mathbf{u}(\mu)\} + U_r) \backslash \{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'}}{\|\mathbf{x}\|_U} + \varepsilon \max_{\mathbf{x} \in U \backslash \{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}}{\|\mathbf{x}\|_U} \right).
\end{aligned}$$

□

Proof of Proposition 2.4.3. Let $\mathbf{a} \in \mathbb{K}^r$ and $\mathbf{x} := \mathbf{U}_r\mathbf{a}$. Then

$$\begin{aligned}\frac{\|\mathbf{A}_r(\mu)\mathbf{a}\|}{\|\mathbf{a}\|} &= \max_{\mathbf{z}\in\mathbb{K}^r\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{z}, \mathbf{A}_r(\mu)\mathbf{a}\rangle|}{\|\mathbf{z}\|\|\mathbf{a}\|} = \max_{\mathbf{z}\in\mathbb{K}^r\backslash\{\mathbf{0}\}} \frac{|\mathbf{z}^{\mathrm{H}}\mathbf{U}_r^{\mathrm{H}}\mathbf{\Theta}^{\mathrm{H}}\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_r\mathbf{a}|}{\|\mathbf{z}\|\|\mathbf{a}\|} \\ &= \max_{\mathbf{y}\in U_r\backslash\{\mathbf{0}\}} \frac{|\mathbf{y}^{\mathrm{H}}\mathbf{\Theta}^{\mathrm{H}}\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{x}|}{\|\mathbf{y}\|_U^{\mathbf{\Theta}}\|\mathbf{x}\|_U^{\mathbf{\Theta}}} = \max_{\mathbf{y}\in U_r\backslash\{\mathbf{0}\}} \frac{|\langle \mathbf{y}, \mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{x}\rangle_U^{\mathbf{\Theta}}|}{\|\mathbf{y}\|_U^{\mathbf{\Theta}}\|\mathbf{x}\|_U^{\mathbf{\Theta}}} \\ &= \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'}^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U^{\mathbf{\Theta}}}.\end{aligned} \tag{2.56}$$

By definition,

$$\sqrt{1-\varepsilon}\|\mathbf{x}\|_U \leq \|\mathbf{x}\|_U^{\mathbf{\Theta}} \leq \sqrt{1+\varepsilon}\|\mathbf{x}\|_U. \tag{2.57}$$

Combining (2.56) and (2.57) we conclude that

$$\frac{1}{\sqrt{1+\varepsilon}}\frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'}^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U} \leq \frac{\|\mathbf{A}_r(\mu)\mathbf{a}\|}{\|\mathbf{a}\|} \leq \frac{1}{\sqrt{1-\varepsilon}}\frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U_r'}^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U}.$$

The statement of the proposition follows immediately from definitions of $\alpha_r^{\mathbf{\Theta}}(\mu)$ and $\beta_r^{\mathbf{\Theta}}(\mu)$. □

Proof of Proposition 2.4.4. The proposition directly follows from relations (2.10), (2.19), (2.20) and (2.31). □

Proof of Proposition 2.4.6. We have

$$\begin{aligned}|s^{\mathrm{pd}}(\mu) - s_r^{\mathrm{spd}}(\mu)| &= |\langle \mathbf{u}_r^{\mathrm{du}}(\mu), \mathbf{R}_U^{-1}\mathbf{r}(\mathbf{u}_r(\mu);\mu)\rangle_U - \langle \mathbf{u}_r^{\mathrm{du}}(\mu), \mathbf{R}_U^{-1}\mathbf{r}(\mathbf{u}_r(\mu);\mu)\rangle_U^{\mathbf{\Theta}}| \\ &\leq \varepsilon\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)||_{U'}\|\mathbf{u}_r^{\mathrm{du}}(\mu)\|_U \\ &\leq \varepsilon\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)||_{U'}\frac{\|\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{u}_r^{\mathrm{du}}(\mu)\|_{U'}}{\eta(\mu)} \\ &\leq \varepsilon\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)||_{U'}\frac{\|\mathbf{r}^{\mathrm{du}}(\mathbf{u}_r^{\mathrm{du}}(\mu);\mu)\|_{U'} + \|\mathbf{l}(\mu)\|_{U'}}{\eta(\mu)},\end{aligned} \tag{2.58}$$

and (2.34) follows by combining (2.58) with (2.15). □

Proof of Proposition 2.5.1. In total, there are at most $\binom{m}{r}$ r-dimensional subspaces that could be spanned from m snapshots. Therefore, by using the definition of $\mathbf{\Theta}$, the fact that $\dim(Y_r(\mu)) \leq 2r+1$ and a union bound for the probability of success, we deduce that $\mathbf{\Theta}$ is a $U \to \ell_2$ ε-subspace embedding for $Y_r(\mu)$, for fixed $\mu \in \mathcal{P}_{\mathrm{train}}$, with probability at least $1 - m^{-1}\delta$. The proposition then follows from another union bound. □

Proof of Proposition 2.5.4. We have,

$$\Delta^{\mathrm{POD}}(V) = \frac{1}{m}\|\mathbf{U}_m^{\mathbf{\Theta}} - \mathbf{\Theta}\mathbf{P}_V^{\mathbf{\Theta}}\mathbf{U}_m\|_F.$$

Moreover, the matrix $\mathbf{\Theta}\mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{U}_m$ is the rank-r truncated SVD approximation of $\mathbf{U}_m^{\mathbf{\Theta}}$. The statements of the proposition can be then derived from the standard properties of SVD. □

Proof of Theorem 2.5.5. Clearly, if $\mathbf{\Theta}$ is a $U \to \ell_2$ ε-subspace embedding for Y, then $\mathrm{rank}(\mathbf{U}_m^{\mathbf{\Theta}}) \geq r$. Therefore U_r is well-defined. Let $\{(\lambda_i, \mathbf{t}_i)\}_{i=1}^{l}$ and $\mathbf{T}_r$ be given by Definition 2.5.3. In general, $\mathbf{P}_{U_r}^{\mathbf{\Theta}}$ defined by (2.44) may not be unique. Let us further assume that $\mathbf{P}_{U_r}^{\mathbf{\Theta}}$ is provided for $\mathbf{x} \in U_m$ by $\mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{x} := \mathbf{U}_r\mathbf{U}_r^{\mathrm{H}}\mathbf{\Theta}^{\mathrm{H}}\mathbf{\Theta}\mathbf{x}$, where $\mathbf{U}_r = \mathbf{U}_m[\frac{1}{\sqrt{\lambda_1}}\mathbf{t}_1, ..., \frac{1}{\sqrt{\lambda_r}}\mathbf{t}_r]$. Observe that $\mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{U}_m = \mathbf{U}_m\mathbf{T}_r\mathbf{T}_r^{\mathrm{H}}$. For the first part of the theorem, we establish the following inequalities. Let $\mathbf{Q} \in \mathbb{K}^{n\times n}$ (e.g., adjoint of a Cholesky factor) be such that $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{R}_U$, then

$$\begin{aligned}
&\frac{1}{m}\sum_{i=1}^{m}\|(\mathbf{I}-\mathbf{P}_Y)(\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{u}(\mu^i))\|_U^2 = \frac{1}{m}\|\mathbf{Q}(\mathbf{I}-\mathbf{P}_Y)\mathbf{U}_m(\mathbf{I}-\mathbf{T}_r\mathbf{T}_r^{\mathrm{H}})\|_F^2 \\
&\leq \frac{1}{m}\|\mathbf{Q}(\mathbf{I}-\mathbf{P}_Y)\mathbf{U}_m\|_F^2\|\mathbf{I}-\mathbf{T}_r\mathbf{T}_r^{\mathrm{H}}\|^2 = \Delta_Y\|\mathbf{I}-\mathbf{T}_r\mathbf{T}_r^{\mathrm{H}}\|^2 \leq \Delta_Y,
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{m}\sum_{i=1}^{m}\left(\|(\mathbf{I}-\mathbf{P}_Y)(\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{u}(\mu^i))\|_U^{\mathbf{\Theta}}\right)^2 = \frac{1}{m}\|\mathbf{\Theta}(\mathbf{I}-\mathbf{P}_Y)\mathbf{U}_m(\mathbf{I}-\mathbf{T}_r\mathbf{T}_r^{\mathrm{H}})\|_F^2 \\
&\leq \frac{1}{m}\|\mathbf{\Theta}(\mathbf{I}-\mathbf{P}_Y)\mathbf{U}_m\|_F^2\|\mathbf{I}-\mathbf{T}_r\mathbf{T}_r^{\mathrm{H}}\|^2 \leq (1+\varepsilon)\Delta_Y\|\mathbf{I}-\mathbf{T}_r\mathbf{T}_r^{\mathrm{H}}\|^2 \leq (1+\varepsilon)\Delta_Y.
\end{aligned}$$

Now, we have

$$\begin{aligned}
&\frac{1}{m}\sum_{i=1}^{m}\|\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}\mathbf{u}(\mu^i)\|_U^2 \leq \frac{1}{m}\sum_{i=1}^{m}\|\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{u}(\mu^i)\|_U^2 \\
&= \frac{1}{m}\sum_{i=1}^{m}\left(\|\mathbf{P}_Y(\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{u}(\mu^i))\|_U^2 + \|(\mathbf{I}-\mathbf{P}_Y)(\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{u}(\mu^i))\|_U^2\right) \\
&\leq \frac{1}{m}\sum_{i=1}^{m}\|\mathbf{P}_Y(\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{u}(\mu^i))\|_U^2 + \Delta_Y \leq \frac{1}{m}\frac{1}{1-\varepsilon}\sum_{i=1}^{m}\left(\|\mathbf{P}_Y(\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{u}(\mu^i))\|_U^{\mathbf{\Theta}}\right)^2 + \Delta_Y \\
&\leq \frac{1}{1-\varepsilon}\frac{1}{m}\sum_{i=1}^{m}2\left(\left(\|\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{u}(\mu^i)\|_U^{\mathbf{\Theta}}\right)^2 + \left(\|(\mathbf{I}-\mathbf{P}_Y)(\mathbf{u}(\mu^i) - \mathbf{P}_{U_r}^{\mathbf{\Theta}}\mathbf{u}(\mu^i))\|_U^{\mathbf{\Theta}}\right)^2\right) + \Delta_Y \\
&\leq \frac{1}{1-\varepsilon}\frac{1}{m}\sum_{i=1}^{m}2\left(\|\mathbf{u}(\mu^i) - \mathbf{P}_{U_r^*}\mathbf{u}(\mu^i)\|_U^{\mathbf{\Theta}}\right)^2 + (\frac{2(1+\varepsilon)}{1-\varepsilon}+1)\Delta_Y \\
&\leq \frac{2(1+\varepsilon)}{1-\varepsilon}\frac{1}{m}\sum_{i=1}^{m}\|\mathbf{u}(\mu^i) - \mathbf{P}_{U_r^*}\mathbf{u}(\mu^i)\|_U^2 + (\frac{2(1+\varepsilon)}{1-\varepsilon}+1)\Delta_Y,
\end{aligned}$$

which is equivalent to (2.48).

The second part of the theorem can be proved as follows. Assume that $\mathbf{\Theta}$ is a $U \to \ell_2$ ε-subspace embedding for U_m, then

$$\begin{aligned}
&\frac{1}{m}\sum_{i=1}^{m}\|\mathbf{u}(\mu^i)-\mathbf{P}_{U_r}\mathbf{u}(\mu^i)\|_U^2 \leq \frac{1}{m}\sum_{i=1}^{m}\|\mathbf{u}(\mu^i)-\mathbf{P}^{\mathbf{\Theta}}_{U_r}\mathbf{u}(\mu^i)\|_U^2\\
&\leq \frac{1}{m}\frac{1}{1-\varepsilon}\sum_{i=1}^{m}\left(\|\mathbf{u}(\mu^i)-\mathbf{P}^{\mathbf{\Theta}}_{U_r}\mathbf{u}(\mu^i)\|_U^{\mathbf{\Theta}}\right)^2 \leq \frac{1}{m}\frac{1}{1-\varepsilon}\sum_{i=1}^{m}\left(\|\mathbf{u}(\mu^i)-\mathbf{P}_{U_r^*}\mathbf{u}(\mu^i)\|_U^{\mathbf{\Theta}}\right)^2\\
&\leq \frac{1}{m}\frac{1+\varepsilon}{1-\varepsilon}\sum_{i=1}^{m}\|\mathbf{u}(\mu^i)-\mathbf{P}_{U_r^*}\mathbf{u}(\mu^i)\|_U^2,
\end{aligned}$$

which completes the proof. □

2.9 Supplementary material

Here we provide detailed proofs of Remark 2.3.8 and of some statements used in the proofs of Propositions 2.3.7 and 2.3.9.

Supplementary material for the proof of Proposition 2.3.7

In the proof of Proposition 2.3.7 for the complex case (i.e., $\mathbb{K} = \mathbb{C}$) we used the following result.

Proposition 2.9.1. *Let* $\mathbf{\Theta}$ *be a real random matrix. If* $\mathbf{\Theta}$ *is* $(\varepsilon, \delta, 2d)$ *oblivious* $\ell_2 \to \ell_2$ *subspace embedding for subspaces of vectors in* $\mathbb{R}^n$*, then it is* (ε, δ, d) *oblivious* $\ell_2 \to \ell_2$ *subspace embedding for subspaces of vectors in* $\mathbb{C}^n$*.*

Proof of Proposition 2.9.1. Let $V \subset \mathbb{C}^n$ be a d-dimensional subspace with a basis $\{\mathbf{v}_i\}_{i=1}^d$. Let us introduce a real subspace $W = \text{span}(\{\text{Re}(\mathbf{v}_i)\}_{i=1}^d) + \text{span}(\{\text{Im}(\mathbf{v}_i)\}_{i=1}^d)$. Observe that

$$\text{Re}(\mathbf{v}) \textit{ and } \text{Im}(\mathbf{v}) \in W, \quad \forall \mathbf{v} \in V.$$

Consequently, if $\mathbf{\Theta}$ is an ε-embedding for W, then

$$| \|\text{Re}(\mathbf{v})\|^2 - \|\mathbf{\Theta}\text{Re}(\mathbf{v})\|^2 | \leq \varepsilon \|\text{Re}(\mathbf{v})\|^2, \quad \forall \mathbf{v} \in V, \tag{2.59a}$$

$$| \|\text{Im}(\mathbf{v})\|^2 - \|\mathbf{\Theta}\text{Im}(\mathbf{v})\|^2 | \leq \varepsilon \|\text{Im}(\mathbf{v})\|^2, \quad \forall \mathbf{v} \in V. \tag{2.59b}$$

Since $\mathbf{\Theta}$ is a real matrix, relations (2.59) imply

$$| \|\mathbf{v}\|^2 - \|\mathbf{\Theta}\mathbf{v}\|^2 | \leq \varepsilon \|\mathbf{v}\|^2, \quad \forall \mathbf{v} \in V. \tag{2.60}$$

By definition of $\mathbf{\Theta}$ and the fact that $\dim(W) \leq 2d$, it follows that $\mathbf{\Theta}$ is an ε-embedding for W with probability at least $1 - \delta$. From this we deduce that (2.60) holds with probability at least $1 - \delta$. It remains to show that (2.60) implies

$$| \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{\Theta}\mathbf{x}, \mathbf{\Theta}\mathbf{y} \rangle | \leq \varepsilon \|\mathbf{x}\| \|\mathbf{y}\|, \quad \forall \mathbf{x}, \mathbf{y} \in V. \tag{2.61}$$

Let $\mathbf{x}, \mathbf{y} \in V$ be any two vectors from V. Define $\mathbf{x}^* := \mathbf{x}/\|\mathbf{x}\|$, $\mathbf{y}^* := \mathbf{y}/\|\mathbf{y}\|$ and

$$\omega := \frac{\langle \mathbf{x}^*, \mathbf{y}^* \rangle - \langle \mathbf{\Theta}\mathbf{x}^*, \mathbf{\Theta}\mathbf{y}^* \rangle}{|\langle \mathbf{x}^*, \mathbf{y}^* \rangle - \langle \mathbf{\Theta}\mathbf{x}^*, \mathbf{\Theta}\mathbf{y}^* \rangle|}.$$

Observe that $|\omega| = 1$ and $\langle \mathbf{x}^*, \omega \mathbf{y}^* \rangle - \langle \mathbf{\Theta}\mathbf{x}^*, \omega \mathbf{\Theta}\mathbf{y}^* \rangle$ is a real number. Then, (2.60) and the parallelogram identity yield

$$\begin{aligned}4|\langle \mathbf{x}^*,\mathbf{y}^*\rangle - \langle \mathbf{\Theta}\mathbf{x}^*,\mathbf{\Theta}\mathbf{y}^*\rangle| &= |4\langle \mathbf{x}^*,\omega\mathbf{y}^*\rangle - 4\langle \mathbf{\Theta}\mathbf{x}^*,\omega\mathbf{\Theta}\mathbf{y}^*\rangle| \\ &= |\|\mathbf{x}^*+\omega\mathbf{y}^*\|^2 - \|\mathbf{x}^*-\omega\mathbf{y}^*\|^2 + 4\mathrm{Im}(\langle \mathbf{x}^*,\omega\mathbf{y}^*\rangle) \\ &- \left(\|\mathbf{\Theta}(\mathbf{x}^*+\omega\mathbf{y}^*)\|^2 - \|\mathbf{\Theta}(\mathbf{x}^*-\omega\mathbf{y}^*)\|^2 + 4\mathrm{Im}(\langle \mathbf{\Theta}\mathbf{x}^*,\omega\mathbf{\Theta}\mathbf{y}^*\rangle)\right)| \\ &= |\|\mathbf{x}^*+\omega\mathbf{y}^*\|^2 - \|\mathbf{\Theta}(\mathbf{x}^*+\omega\mathbf{y}^*)\|^2 - \left(\|\mathbf{x}^*-\omega\mathbf{y}^*\|^2 - (\|\mathbf{\Theta}(\mathbf{x}^*-\omega\mathbf{y}^*)\|^2\right) \\ &+ 4\mathrm{Im}(\langle \mathbf{x}^*,\omega\mathbf{y}^*\rangle - \langle \mathbf{\Theta}\mathbf{x}^*,\omega\mathbf{\Theta}\mathbf{y}^*\rangle)| \\ &\leq \varepsilon\|\mathbf{x}^*+\omega\mathbf{y}^*\|^2 + \varepsilon\|\mathbf{x}^*-\omega\mathbf{y}^*\|^2 = 4\varepsilon.\end{aligned}$$

The relation (2.61) follows immediately. □

Supplementary material for Remark 2.3.8

Recall that $\mathbf{\Theta} \in \mathbb{C}^{k\times n}$ is called a rescaled complex Gaussian matrix if

$$\mathbf{\Theta} := \frac{1}{\sqrt{2}}(\mathbf{\Theta}_{\mathrm{Re}} + j\mathbf{\Theta}_{\mathrm{Im}}), \tag{2.62}$$

where $j = \sqrt{-1}$, $\mathbf{\Theta}_{\mathrm{Re}}$, $\mathbf{\Theta}_{\mathrm{Im}}$ are two independent rescaled real Gaussian matrices (that have i.i.d. entries with mean 0 and variance k^{-1}). Let us now give a proof of the statement in Remark 2.3.8 (see proposition below).

Proposition 2.9.2. *A distribution of rescaled complex Gaussian matrices (defined by (2.62)) with $k \geq 3.94\varepsilon^{-2}(13.8d + \log(1/\delta))$ rows satisfies (ε,δ,d) oblivious $\ell_2 \to \ell_2$ subspace embedding property.*

The proof of this statement shall be obtained by following the proof of Proposition 2.3.7 updating the constants in some places. First, let us establish the following result.

Lemma 2.9.3. *A rescaled complex Gaussian matrix $\mathbf{\Theta}$ (defined by (2.62)) with $k \geq (\varepsilon^2/2 - \varepsilon^3/3)^{-1}\log(2/\delta)$ is a $(\varepsilon,\delta,1)$ oblivious $\ell_2 \to \ell_2$ subspace embedding.*

Proof of Lemma 2.9.3. Let $\mathbf{z} \in \mathbb{C}^n$ be an arbitrary unit vector. Define $\mathbf{x} := \begin{bmatrix}\mathrm{Re}(\mathbf{z}) \\ -\mathrm{Im}(\mathbf{z})\end{bmatrix}$, $\mathbf{y} := \begin{bmatrix}\mathrm{Im}(\mathbf{z}) \\ \mathrm{Re}(\mathbf{z})\end{bmatrix}$ and $\mathbf{\Theta}^* := [\mathbf{\Theta}_{\mathrm{Re}}, \mathbf{\Theta}_{\mathrm{Im}}]$. Observe that $\mathbf{\Theta}^*$ is a rescaled real Gaussian matrix, $\mathbf{x}$ and $\mathbf{y}$ are orthogonal unit vectors, and

$$\|\mathbf{\Theta}\mathbf{z}\|^2 = \frac{1}{2}\|\mathbf{\Theta}^*\mathbf{x}\|^2 + \frac{1}{2}\|\mathbf{\Theta}^*\mathbf{y}\|^2.$$

Since products of a Gaussian matrix with orthogonal unit vectors are independent Gaussian vectors, consequently, $k\mathbf{\Theta}^*\mathbf{x}$ and $k\mathbf{\Theta}^*\mathbf{y}$ are independent k-dimensional standard Gaussian vectors. We conclude that $2k\|\mathbf{\Theta z}\|^2$ has a chi-squared distribution with $2k$ degrees of freedom. Finally, the standard tail-bounds for chi-squared distribution ensure that

$$||\mathbf{\Theta z}\|^2 - 1| > \varepsilon,$$

holds with probability less than $\delta = 2\exp(-k(\varepsilon^2/2 - \varepsilon^3/3))$, which completes the proof. □

Proof of Proposition 2.9.2. We can use a similar proof as the one of Proposition 2.3.7 for the real case. Let $V \subset \mathbb{C}^n$ be a d-dimensional subspace and let $\mathcal{S} = \{\mathbf{x} \in W : \|\mathbf{x}\| = 1\}$ be the unit sphere of V. By the volume argument it follows that for any $\gamma > 0$ there exists a γ-net $\mathcal{N}$ of $\mathcal{S}$ satisfying $\#\mathcal{N} \leq (1+2/\gamma)^{2d}$. For η such that $0 < \eta < 1/2$, let $\mathbf{\Theta} \in \mathbb{C}^{k\times n}$ be a rescaled complex Gaussian matrix (defined by (2.62)) with $k \geq 3\eta^{-2}(4d\log(1+2/\gamma)+\log(1/\delta))$ rows. By Lemma 2.9.3 and a union bound argument, we have that

$$\left\{ ||\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{\Theta}(\mathbf{x}+\mathbf{y})\|^2| \leq \eta\|\mathbf{x}+\mathbf{y}\|^2, \quad \forall \mathbf{x},\mathbf{y} \in \mathcal{N}\right\}$$

holds with probability at least $1-\delta$. This implies that

$$\{ \ |\langle \mathbf{x},\mathbf{y}\rangle - \langle \mathbf{\Theta x}, \mathbf{\Theta y}\rangle| \leq \eta, \quad \forall \mathbf{x},\mathbf{y} \in \mathcal{N}\} \tag{2.63}$$

holds with probability at least $1-\delta$.

Assume that $\gamma < 1$. It can be shown that any vector $\mathbf{n} \in \mathcal{S}$ can be expressed as $\mathbf{n} = \sum_{i\geq 0}\alpha_i \mathbf{n}_i$, where $\mathbf{n}_i \in \mathcal{N}$ and α_i are real coefficients such that $0 \leq \alpha_i \leq \gamma^i$. The proof of this fact directly follows the one for the real case in the proof of Proposition 2.3.7. Then (2.63) implies that for all $\mathbf{n} \in \mathcal{S}$,

$$\begin{aligned}\|\mathbf{\Theta n}\|^2 &= \sum_{i,j\geq 0} \langle \mathbf{\Theta n}_i, \mathbf{\Theta n}_j\rangle \alpha_i\alpha_j \\ &\leq \sum_{i,j\geq 0} (\langle \mathbf{n}_i, \mathbf{n}_j\rangle \alpha_i\alpha_j + \eta\alpha_i\alpha_j) = 1 + \eta(\sum_{i\geq 0}\alpha_i)^2 \leq 1 + \frac{\eta}{(1-\gamma)^2},\end{aligned}$$

and, similarly, $\|\mathbf{\Theta n}\|^2 \geq 1 - \frac{\eta}{(1-\gamma)^2}$. Therefore, (2.63) implies

$$|1 - \|\mathbf{\Theta n}\|^2| \leq \eta/(1-\gamma)^2, \ \ \forall \mathbf{n} \in \mathcal{S}. \tag{2.64}$$

For any $\varepsilon \leq 0.5/(1-\gamma)^2$ let us choose $\eta = (1-\gamma)^2\varepsilon$. Now we use the argument from the proof of Proposition 2.9.1, which states that (2.60) yields (2.61). This result implies that, if (2.64) is satisfied, then $\mathbf{\Theta}$ is a $\ell_2 \to \ell_2$ ε-subspace embedding for V. We have that if $k \geq 3\varepsilon^{-2}(1-\gamma)^{-4}(4d\log(1+2/\gamma)+\log(1/\delta))$, then (2.63) (and as a consequence (2.64)) holds with probability at least $1-\delta$, which means that $\mathbf{\Theta}$ is (ε,δ,d) oblivious $\ell_2 \to \ell_2$ subspace embedding. As in Proposition 2.3.7 the lower bound for k is attained with $\gamma = 0.0656$. □

Supplementary material for the proof of Proposition 2.3.9

Let us now present a proof of [31, Lemma 4.3] and [146, Theorem 3.1] for the complex case and with improved constants (see Proposition 2.9.4), which is used in the proof of Proposition 2.3.9. Consider a SRHT matrix $\mathbf{\Theta}$ of size $k \times n$, with n being a power of 2. Recall that

$$\mathbf{\Theta} = k^{-1/2}(\mathbf{R}\mathbf{H}_n\mathbf{D}), \tag{2.65}$$

where $\mathbf{R}$ are the first k rows of a random permutation of rows of the identity matrix, $\mathbf{H}_n$ is a Hadamard matrix, and $\mathbf{D}$ is a random diagonal matrix with i.i.d. entries with Rademacher distribution (i.e., taking values ±1 with equal probabilities). To be consistent with the notations from [31, 146] let us also define a rescaled Hadamard matrix $\mathbf{H} := \frac{1}{\sqrt{n}}\mathbf{H}_n$ with orthonormal columns.

Proposition 2.9.4 (Complex version of Lemma 4.3 in [31], Theorem 3.1 in [146]). *Let $\mathbf{V} \in \mathbb{C}^{n\times d}$ be a matrix with orthonormal columns. Let $0 < \varepsilon < 1$ and $0 < \delta < 1$. Draw at random a matrix $\mathbf{\Theta}$ defined in (2.65) with*

$$k \geq 2(\varepsilon^2 - \varepsilon^3/3)^{-1}[\sqrt{d} + \sqrt{8\log(n/\delta)}]^2 \log(d/\delta).$$

Then with probability at least $1-3\delta$, the singular values of $\mathbf{\Theta V}$ belong to the interval $[\sqrt{1-\varepsilon}, \sqrt{1+\varepsilon}]$.

Proposition 2.9.4 can be derived from complex extensions of [146, Lemmas 3.3 and 3.4] presented below.

Lemma 2.9.5 (Lemma 3.3 in [146]). *Let $\mathbf{V} \in \mathbb{C}^{n\times d}$ be a matrix with orthonormal columns. Draw at random a diagonal matrix $\mathbf{D}$ in (2.65). The rows $\mathbf{w}_j^{\mathrm{T}}$ of $\mathbf{HDV}$ satisfy*

$$\mathbb{P}\left(\max_{j=1,\ldots,n} \|\mathbf{w}_j\| \leq \sqrt{\frac{d}{n}} + \sqrt{\frac{8\log(n/\delta)}{n}}\right) \geq 1-\delta.$$

Proof of Lemma 2.9.5. This lemma can be proven with exactly the same steps as in the proof of [146, Lemma 3.3]. We have $\mathbf{w}_j = \mathbf{V}^{\mathrm{T}}\mathrm{diag}(\mathbf{i})\mathbf{H}\mathbf{e}_j$ where $\mathbf{e}_j$ is the jth column of the identity matrix and $\mathbf{i}$ is a Rademacher vector. Define functions $f_j(\mathbf{x}) := \|\mathbf{V}^{\mathrm{T}}\mathrm{diag}(\mathbf{x})\mathbf{H}\mathbf{e}_j\|$. Observe that $f_j(\mathbf{x}) = \|\mathbf{V}^{\mathrm{T}}\mathbf{E}_j\mathbf{x}\|$, with $\mathbf{E}_j := \mathrm{diag}(\mathbf{H}\mathbf{e}_j)$ being a matrix with 2-norm $\|\mathbf{E}_j\| = \frac{1}{\sqrt{n}}$. We have,

$$\forall \mathbf{x}, \mathbf{y},\ |f_j(\mathbf{x}) - f_j(\mathbf{y})| \leq \|\mathbf{V}^{\mathrm{T}}\mathbf{E}_j(\mathbf{x}-\mathbf{y})\| \leq \|\mathbf{V}\|\|\mathbf{E}_j\|\|\mathbf{x}-\mathbf{y}\| = \frac{1}{\sqrt{n}}\|\mathbf{x}-\mathbf{y}\|.$$

Moreover, the functions $f_j(\mathbf{x})$ are convex, which allows to apply the Rademacher tail bound [146, Proposition 2.1]

$$\mathbb{P}(f_j(\mathbf{i}) \geq \mathbb{E}f_j(\mathbf{i}) + \frac{1}{\sqrt{n}}t) \leq \exp\left(-t^2/8\right),\ \forall t \geq 0, \tag{2.66}$$

with $\mathbf{i}$ being a Rademacher vector. Observe that $\mathbb{E} f_j(\mathbf{i}) \leq (\mathbb{E}(f_j(\mathbf{i}))^2)^{1/2} = \|\mathbf{E}_j \mathbf{V}\|_F \leq \|\mathbf{E}_j\| \|\mathbf{V}\|_F = \sqrt{\frac{d}{n}}$. The statement of the proposition follows by combining this relation with (2.66) with $t = \sqrt{8\log(n/\delta)}$, and by using the union bound argument. □

Lemma 2.9.6 (Lemma 3.4 in [146]). *Let $\mathbf{W} \in \mathbb{C}^{n\times d}$ have orthonormal columns. Let $0<\varepsilon<1$ and $0<\delta<1$. Let $\mathbf{w}_j^{\mathrm{T}}$ denote the rows of $\mathbf{W}$ and let $M := n \max_{j=1,\dots,n} \|\mathbf{w}_j\|^2$. Draw at random a permutation matrix $\mathbf{R}$ in (2.65) with*

$$k \geq 2(\varepsilon^2 - \varepsilon^3/3)^{-1} M \log(d/\delta).$$

Then with probability at least $1-2\delta$, all the singular values of $\sqrt{\frac{n}{k}}\mathbf{R}\mathbf{W}$ belong to the interval $[\sqrt{1-\varepsilon}, \sqrt{1+\varepsilon}]$.

To prove Lemma 2.9.6 we shall use the matrix Chernoff tail bounds from [146]. For any Hermitian matrix $\mathbf{X}$, let $\lambda_{\min}(\mathbf{X})$ and $\lambda_{\max}(\mathbf{X})$ denote the minimal and the maximal eigenvalues of $\mathbf{X}$.

Theorem 2.9.7 (Theorem 2.2 in [146]). *Consider a finite set $X \subseteq \mathbb{C}^{d\times d}$ of Hermitian positive semi-definite matrices. Define the constant $L := \max_{\mathbf{X}_j \in X} \lambda_{\max}(\mathbf{X}_j)$. Let $\{\mathbf{X}_i\}_{i=1}^k \subseteq X$ be a uniformly sampled, without replacement, random subset of X and $\mathbf{X} := \sum_{i=1}^k \mathbf{X}_i$. Then*

$$\mathbb{P}(\lambda_{\min}(\mathbf{X}) \leq (1-\varepsilon)\mu_{\min}) \leq d\left(\frac{e^{-\varepsilon}}{(1-\varepsilon)^{1-\varepsilon}}\right)^{\mu_{\min}/L},$$
$$\mathbb{P}(\lambda_{\max}(\mathbf{X}) \geq (1+\varepsilon)\mu_{\max}) \leq d\left(\frac{e^{\varepsilon}}{(1+\varepsilon)^{1+\varepsilon}}\right)^{\mu_{\max}/L},$$

where $\mu_{\min} = k\ \lambda_{\min}(\mathbb{E}\mathbf{X}_1)$ and $\mu_{\max} = k\ \lambda_{\max}(\mathbb{E}\mathbf{X}_1)$.

Proof of Theorem 2.9.7. The proof directly follows the one in [146], since all the ingredients used in the proof of [146, Theorem 2.2] (which are [147, Proposition 3.1, Lemma 3.4, Lemma 5.8] and the result of [78]) are formulated for (Hermitian) positive semi-definite matrices. □

Proof of Lemma 2.9.6. Define $X := \{\mathbf{w}_j \mathbf{w}_j^{\mathrm{H}}\}_{j=1}^n$. Consider the matrix

$$\mathbf{X} := (\mathbf{R}\mathbf{W})^{\mathrm{H}} \mathbf{R}\mathbf{W} = \sum_{j\in T} \mathbf{w}_j \mathbf{w}_j^{\mathrm{H}},$$

where T is a set, with $\#T = k$, of elements of $\{1,2,\dots,n\}$ drawn uniformly and without replacement. The matrix $\mathbf{X}$ can be written as

$$\mathbf{X} = \sum_{i=1}^k \mathbf{X}_i,$$

where $\{\mathbf{X}_i\}_{i=1}^{k}$ is a uniformly drawn, without replacement, random subset of X. We have $\mathbb{E}(\mathbf{X}_1) = \frac{1}{n}\mathbf{W}^{\mathrm{H}}\mathbf{W} = \frac{1}{n}\mathbf{I}$. Furthermore,

$$\lambda_{\max}(\mathbf{w}_j \mathbf{w}_j^{\mathrm{H}}) = \|\mathbf{w}_j\|^2 \leq \frac{M}{n},\ 1 \leq j \leq n.$$

By applying Theorem 2.9.7 and some algebraic operations, we obtain

$$\mathbb{P}(\lambda_{\min}(\mathbf{X}) \leq (1-\varepsilon)k/n) \leq d\left(\frac{e^{-\varepsilon}}{(1-\varepsilon)^{1-\varepsilon}}\right)^{k/M} \leq d\ e^{-(\varepsilon^2/2-\varepsilon^3/6)k/M} \leq \delta,$$

$$\mathbb{P}(\lambda_{\max}(\mathbf{X}) \geq (1+\varepsilon)k/n) \leq d\left(\frac{e^{\varepsilon}}{(1+\varepsilon)^{1+\varepsilon}}\right)^{k/M} \leq d\ e^{-(\varepsilon^2/2-\varepsilon^3/6)k/M} \leq \delta.$$

The statement of the lemma follows by a union bound argument. □

Proof of Proposition 2.9.4. Let $\mathbf{W} = \mathbf{HDV}$. Observe that $\mathbf{W}$ has orthonormal columns. The statement of the proposition follows from Lemma 2.9.5 with the tail bound from Lemma 2.9.6 and a union bound argument. □

Chapter 3

Random sketching for minimal residual methods and dictionary-based approximation

This chapter essentially constitutes the article [13] that was submitted for a publication in journal "Advances in Computational Mathematics". Following the framework from Chapter 2, we here construct a reduced model from a small, efficiently computable random object called a sketch of a reduced model, using minimal residual methods. We introduce a sketched version of the minimal residual based projection as well as a novel nonlinear approximation method, where for each parameter value, the solution is approximated by minimal residual projection onto a subspace spanned by several vectors picked from a dictionary of candidate basis vectors. It is shown that random sketching technique can improve not only efficiency but also numerical stability. A rigorous analysis of the conditions on the random sketch required to obtain a given accuracy is presented. These conditions may be ensured a priori with high probability by considering for the sketching matrix an oblivious embedding of sufficiently large size. Unlike with Galerkin methods, with minimal residual methods the quality of the sketching matrix can be characterized regardless of the operator's properties. Furthermore, a simple and reliable procedure for a posteriori verification of the quality of the sketch is provided. This approach can be used for certification of the approximation as well as for adaptive selection of the optimal size of the random sketching matrix. We also propose a two-step procedure for an efficient and stable estimation of an inner product between parameter-dependent vectors having affine decompositions with many (possibly expensive to maintain) terms. This procedure can be used for extraction of a quantity of interest (linear or quadratic functional of the solution), estimation of the primal-dual correction, etc.

Contents

3.1 Introduction

We consider large parameter-dependent systems of equations

$$\mathbf{A}(\mu)\mathbf{u}(\mu) = \mathbf{b}(\mu),\ \mu \in \mathcal{P}, \tag{3.1}$$

where $\mathbf{u}(\mu)$ is a solution vector, $\mathbf{A}(\mu)$ is a parameter-dependent matrix, $\mathbf{b}(\mu)$ is a parameter-dependent right hand side and $\mathcal{P}$ is a parameter set. Parameter-dependent problems are considered for many purposes such as design, control, optimization, uncertainty quantification or inverse problems.

Solving (3.1) for many parameter values can be computationally unfeasible. Moreover, for real-time applications, a quantity of interest ($\mathbf{u}(\mu)$ or a function of $\mathbf{u}(\mu)$) has to be estimated on the fly in highly limited computational time for a certain value of μ. Model order reduction (MOR) methods are developed for efficient approximation of the quantity of interest for each parameter value. They typically consist of two stages. In the first so-called offline stage a reduced model is constructed from the full order model. This stage usually involves expensive computations such as evaluations of $\mathbf{u}(\mu)$ for several parameter values, computing multiple high-dimensional matrix-vector and inner products, etc., but this stage is performed only once. Then, for each given parameter value, the precomputed reduced model is used for efficient approximation of the solution or an output quantity with a computational cost independent of the dimension of the initial system of equations (3.1). For a detailed presentation of the classical MOR methods such as Reduced Basis (RB) method and Proper Orthogonal Decomposition (POD) the reader can refer to [24]. In the present work the approximation of the solution shall be obtained with a minimal residual (minres) projection on a reduced (possibly parameter-dependent) subspace. The minres projection can be interpreted as a Petrov-Galerkin projection where the test space is chosen to minimize some norm of the residual [7, 38]. Major benefits over the classical Galerkin projection include an improved stability (quasi-optimality) for non-coercive problems and more effective residual-based error bounds of an approximation (see e.g. [38]). In addition, minres methods are better suited to random sketching as will be seen in the present chapter.

In recent years randomized linear algebra (RLA) became a popular approach in the fields such as data analysis, machine learning, signal processing, etc. [110, 149, 157]. This probabilistic approach for numerical linear algebra can yield a drastic computational cost reduction in terms of classical metrics of efficiency such as complexity (number of flops) and memory consumption. Moreover, it can be highly beneficial in extreme computational environments that are typical in contemporary scientific computing. For instance, RLA can be essential when data has to be analyzed only in one pass (e.g., when it is streamed from a server) or when it is distributed on multiple workstations with expensive communication costs.

Despite their indisputable success in fields closely related to MOR, the aforementioned techniques only recently started to be used for model order reduction. One of the earliest works considering RLA in the context of MOR is [160], where the authors proposed to use RLA for interpolation of (implicit) inverse of a parameter-dependent matrix. In [37] the RLA was used for approximating the range of a transfer operator and for computing a probabilistic bound for the approximation error. In [141] the authors developed a probabilistic error estimator, which can also be reformulated in the RLA framework.

As already shown in Chapter 2, random sketching can lead to drastic reduction of the computational costs of classical MOR methods. A random sketch of a reduced model is defined as a set of small random projections of the reduced basis vectors and the associated residuals. Its representation (i.e, affine decomposition[1]) can be efficiently precomputed in basically any computational architecture. The random projections should be chosen according to the metric of efficiency, e.g., number of flops, memory consumption, communication cost between distributed machines, scalability, etc. A rigorous analysis of the cost of obtaining a random sketch in different computational environments can be found in Section 2.4.4. When a sketch has been computed, the reduced model can be approximated without operating on large vectors but only on their small sketches typically requiring negligible computational costs. The approximation can be guaranteed to almost preserve the quality of the original reduced model with user-specified probability. The computational cost depends only logarithmically on the probability of failure, which can therefore be chosen very small, say 10^{-10}. In Chapter 2 it was shown how random sketching can be employed for an efficient estimation of the Galerkin projection, the computation of the norm of the residual for error estimation, and the computation of a primal-dual correction. Furthermore, new efficient sketched versions of greedy algorithms and Proper Orthogonal Decomposition were introduced for generation of reduced bases.

3.1.1 Contributions

The present work is a continuation of Chapter 2, where we adapt the random sketching technique to minimal residual methods, propose a dictionary-based approximation method and additionally discuss the questions of a posteriori certification of the sketch and efficient extraction of the quantity of interest from a solution of the reduced model. A detailed discussion on the major contributions of the chapter is provided below.

First a sketched version of the minres projection is proposed in Section 3.3, which is more efficient (in both offline and online stages) and numerically stable than the classical approach. The construction of the reduced order model with minres projection involves the evaluation of multiple inner products, which can become a burden for high-dimensional problems. Furthermore, the classical procedure

[1] Recall that a parameter-dependent quantity $\mathbf{v}(\mu)$ with values in a vector space V over $\mathbb{K}$ is said to admit an affine representation if $\mathbf{v}(\mu) = \sum \mathbf{v}_i \lambda_i(\mu)$ with $\lambda_i(\mu) \in \mathbb{K}$ and $\mathbf{v}_i \in V$.

(through orthogonalization of the basis) for ensuring stability of the reduced system of equations only guarantees that the condition number is bounded by the square of the condition number of $\mathbf{A}(\mu)$. Such a bound can be insufficient for applications with very or even moderately ill-conditioned operators. In addition, the online evaluation of the reduced system of equations from the precomputed affine expansions can also be very expensive and suffer from round-off errors. The random sketching technique can offer more efficient (in both offline and online stages) and numerically stable procedures for estimating solutions to the minimal residual problem. Here the reduced model is approximated from its efficiently computable random sketch with a negligible computational cost. The precomputation of a random sketch can require much lower complexity, memory consumption and communication cost than the computations involved in the classical offline stage. As shown in Section 3.3.2, with random sketching the online solution can be found by solving a small $\mathcal{O}(r)$ by r least-squares problem. The construction of this problem takes only $\mathcal{O}(r^2 m_A + r m_b)$ flops (compared to $\mathcal{O}(r^2 m_A^2 + r m_b^2)$ flops required for forming the classical minres reduced system of equations), where m_A and m_b are the numbers of terms in the affine expansions of $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$. Moreover, when the basis is orthogonalized one can guarantee a better stability of the reduced least-squares matrix (the condition number is bounded by the generalized condition number of $\mathbf{A}(\mu)$). In addition, the parameter-dependent reduced matrix can have an affine expansion with considerably less terms and therefore its evaluation is less sensitive to round-off errors. It is also shown that the size of the sketching matrix which is sufficient to preserve the quasi-optimality constants of minres projection can be characterized regardless the properties of the operator (e.g., the condition number). This feature proves the sketched minres projection to be more robust than the sketched Galerkin projection for which the preservation of the approximation's quality can degrade dramatically for ill-conditioned or non-coercive problems as was revealed in Chapter 2.

In Section 3.4 we introduce a novel nonlinear method for approximating the solution to (3.1), where for each value of the parameter, the solution is approximated by a minimal residual projection onto a subspace spanned by several vectors from a dictionary. From an algebraic point of view, this approach can be formulated as a parameter-dependent sparse least-squares problem. It is shown that the solution can be accurately estimated (with high probability) from a random sketch associated with the dictionary, which allows drastic reduction of the computational costs. A condition on the dimension of the random sketch required to obtain a given accuracy is provided. Again, the construction of a reduced model does not require operations on high-dimensional vectors but only on their sketches. In particular, in the offline stage we only need to maintain a sketch of the dictionary. In Section 3.4.4 we propose an efficient greedy-like procedure for the dictionary construction based on snapshots.

The dictionary-based approach is more natural than the classical *hp*-refinement method [68, 69] and it should always provide a better approximation (see Section 3.4.1). The potential of approximation with dictionaries for problems with a slow decay of the Kolmogorov r-widths of the solution manifold was revealed in [64, 97]. Although they improved classical approaches, the algorithms proposed in [64, 97] still involve in general heavy computations in both offline and online stages, and can suffer from round-off errors. If n and r are the dimensions of the full solution space and the (parameter-dependent) reduced approximation space, respectively, and K is the cardinality of the dictionary, then the offline complexity and the memory consumption associated with post-processing of the snapshots in [64, 97] are at least $\mathcal{O}(n(K^2m_A^2+m_b^2))$ and $\mathcal{O}(nK+K^2m_A^2+m_b^2)$, respectively. Furthermore, the online stage in [97] requires $\mathcal{O}((r^3+m_A^2r)K+m_b^2)$ flops and $\mathcal{O}(m_A^2K^2+m_b^2)$ bytes of memory. The high offline cost and the high cost of maintenance of the reduced model (which are proportional to K^2) limits the effectiveness of the method in many computational environments. Moreover, we see that the online complexity of the approach in [97] is proportional to Kr^3, which leads to high computational costs for moderate r. Random sketching can drastically improve efficiency and stability of the dictionary-based approximation. The offline complexity and memory requirements of the construction of a reduced model with random sketching (see Section 3.4.4) are $\mathcal{O}(n(Km_A+m_b)(\log r+\log\log K))$ and $\mathcal{O}(n+(Km_A+m_b)r\log K)$, respectively. We observe reduction (compared to [64, 97]) of the complexity and memory consumption of the offline stage by at least a factor of $\mathcal{O}(K)$. In its turn, the online stage with random sketching needs $\mathcal{O}((m_AK+m_b)r\log K+r^2K\log K)$ flops and $\mathcal{O}((m_AK+m_b)r\log K)$ bytes of memory, which are in $\mathcal{O}((r+m_A+\frac{m_b}{r})/\log K)$ and $\mathcal{O}((m_AK+m_b)/(r\log K))$ times less than the requirements in [97]. Note that in some places logarithmic terms were neglected. A more detailed analysis of the computational costs can be found in Sections 3.4.3 and 3.4.4.

The online stage usually proceeds with computation of the coordinates of an approximation $\mathbf{u}_r(\mu)$ of $\mathbf{u}(\mu)$ in a certain basis. Then the coordinates are used for the efficient evaluation of an estimation $s_r(\mu):=l(\mathbf{u}_r(\mu);\mu)$ of a quantity of interest from an affine decomposition of $l(\cdot,\mu)$. When the affine decomposition of $l(\cdot,\mu)$ is expensive to maintain and to operate with, precomputation of the affine decomposition of $s_r(\mu)$ can become too cumbersome. This is the case when $l(\cdot,\mu)$ contains numerous terms in its affine decomposition or when one considers too large, possibly distributed, basis for $\mathbf{u}_r(\mu)$. In Section 3.5 we provide a way to efficiently estimate $s_r(\mu)$. Our procedure is two-phased. First the solution $\mathbf{u}_r(\mu)$ is approximated by a projection $\mathbf{w}_p(\mu)$ on a new basis, which is cheap to operate with. The affine decomposition of an approximation $s_r(\mu)\approx l(\mathbf{w}_p(\mu);\mu)$ can now be efficiently precomputed. Then in the second step, the accuracy of $l(\mathbf{w}_p(\mu);\mu)$ is improved with a random correction computable from the sketches of the two bases with a negligible computational cost. Note that our approach can be employed for the efficient computation of the affine decomposition of primal-dual corrections and quadratic quantities of interest (see Remark 3.5.1).

As shown in Chapter 2 for Galerkin methods and in Sections 3.3.2 and 3.4.3 of the present chapter for minres methods, a sketch of a reduced model almost preserves the quality of approximation when the sketching matrix satisfies an ε-embedding property. Such a matrix may be generated randomly by considering an oblivious subspace embedding of sufficiently large size. The number of rows for the oblivious embedding may be selected with the theoretical bounds provided in Chapter 2. However, it was revealed that these bounds are pessimistic or even impractical (e.g., for adaptive algorithms or POD). In practice, one can consider embeddings of much smaller sizes and still obtain accurate reduced order models. Moreover, for some random matrices, theoretical bounds may not be available although there might exist strong empirical evidence that these matrices should yield outstanding performances (e.g., matrices constructed with random Givens rotations as in [131]). When no a priori guarantee on the accuracy of the given sketching matrix is available or when conditions on the size of the sketch based on a priori analysis are too pessimistic, one can provide a posteriori guarantee. An easy and robust procedure for a posteriori verification of the quality of a sketch of a reduced model is provided in Section 3.6. The methodology can also be used for deriving a criterion for adaptive selection of the size of the random sketching matrix to yield an accurate estimation of the reduced model with high probability.

The outline of the chapter is as follows. In Section 3.2 we introduce the problem setting and recall the main ingredients of the framework developed in Chapter 2. The minimal residual method considering a projection on a single low-dimensional subspace is presented in Section 3.3. We present a standard minres projection in a discrete form followed by its efficient approximation with random sketching. Section 3.4 presents a novel dictionary-based minimal residual method using random sketching. A two-phased procedure for efficient and stable extraction of the output quantity of interest from the reduced model's solution is proposed in Section 3.5. A posteriori verification of the quality of a sketch and few scenarios where such a procedure can be used are provided in Section 3.6. The methodology is validated numerically on two benchmark problems in Section 3.7.

3.2 Preliminaries

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and let $U := \mathbb{K}^n$ and $U' := \mathbb{K}^n$ represent the solution space and its dual, respectively. The solution $\mathbf{u}(\mu)$ is an element from U, $\mathbf{A}(\mu)$ is a linear operator from U to U', the right hand side $\mathbf{b}(\mu)$ and the extractor of the quantity of interest $\mathbf{l}(\mu)$ are elements of U'.

Spaces U and U' are equipped with inner products $\langle \cdot, \cdot \rangle_U := \langle \mathbf{R}_U \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle_{U'} := \langle \cdot, \mathbf{R}_U^{-1} \cdot \rangle$, where $\langle \cdot, \cdot \rangle$ is the canonical inner product on $\mathbb{K}^n$ and $\mathbf{R}_U : U \to U'$ is some symmetric (for $\mathbb{K} = \mathbb{R}$), or Hermitian (for $\mathbb{K} = \mathbb{C}$), positive definite operator. We denote by $\| \cdot \|$ the canonical norm on $\mathbb{K}^n$. Finally, for a matrix $\mathbf{M}$ we denote by $\mathbf{M}^{\mathrm{H}}$ its (Hermitian) transpose.

3.2.1 Random sketching

A framework for using random sketching (see [88, 157]) in the context of MOR was introduced in Chapter 2. The sketching technique is seen as a modification of the inner product in a given subspace (or a collection of subspaces). The modified inner product is an estimation of the original one and is much easier and more efficient to operate with. Next, we briefly recall the basic preliminaries from Chapter 2.

Let V be a subspace of U. The dual of V is identified with a subspace $V' := \{\mathbf{R}_U\mathbf{y} : \mathbf{y} \in V\}$ of U'. For a matrix $\mathbf{\Theta} \in \mathbb{K}^{k\times n}$ with $k \leq n$ we define the following semi-inner products on U:

$$\langle\cdot,\cdot\rangle_U^{\mathbf{\Theta}} := \langle\mathbf{\Theta}\cdot,\mathbf{\Theta}\cdot\rangle, \text{ and } \langle\cdot,\cdot\rangle_{U'}^{\mathbf{\Theta}} := \langle\mathbf{\Theta}\mathbf{R}_U^{-1}\cdot,\mathbf{R}_U^{-1}\cdot\rangle, \tag{3.2}$$

and we let $\|\cdot\|_U^{\mathbf{\Theta}}$ and $\|\cdot\|_{U'}^{\mathbf{\Theta}}$ denote the associated semi-norms.

Remark 3.2.1. *The extension of the methodology to the case where $\langle\cdot,\cdot\rangle_U$ is not definite, i.e., $\mathbf{R}_U$ is positive semi-definite, is straightforward. Let us assume that $\langle\cdot,\cdot\rangle_U$ is an inner product on a subspace $W \subseteq U$ of interest. Then, it follows that $W' := \{\mathbf{R}_U\mathbf{x} : \mathbf{x} \in W\}$ can be equipped with $\langle\cdot,\cdot\rangle_{U'} := \langle\cdot,\mathbf{R}_U^{\dagger}\cdot\rangle$, where $\mathbf{R}_U^{\dagger}$ is a pseudo-inverse of $\mathbf{R}_U$. Such products $\langle\cdot,\cdot\rangle_U$ and $\langle\cdot,\cdot\rangle_{U'}$ can be approximated by*

$$\langle\cdot,\cdot\rangle_U^{\mathbf{\Theta}} := \langle\mathbf{\Theta}\cdot,\mathbf{\Theta}\cdot\rangle, \text{ and } \langle\cdot,\cdot\rangle_{U'}^{\mathbf{\Theta}} := \langle\mathbf{\Theta}\mathbf{R}_U^{\dagger}\cdot,\mathbf{\Theta}\mathbf{R}_U^{\dagger}\cdot\rangle. \tag{3.3}$$

This will be useful for the estimation of a (semi-)inner product between parameter-dependent vectors in Section 3.5 (see Remark 3.5.2).

Definition 3.2.2. *A matrix $\mathbf{\Theta}$ is called a $U \to \ell_2$ ε-subspace embedding (or simply an ε-embedding) for V, if it satisfies*

$$\forall \mathbf{x},\mathbf{y} \in V, \ \left|\langle\mathbf{x},\mathbf{y}\rangle_U - \langle\mathbf{x},\mathbf{y}\rangle_U^{\mathbf{\Theta}}\right| \leq \varepsilon\|\mathbf{x}\|_U\|\mathbf{y}\|_U. \tag{3.4}$$

Here ε-embeddings shall be constructed as realizations of random matrices that are built in an oblivious way without any a priori knowledge of V.

Definition 3.2.3. *A random matrix $\mathbf{\Theta}$ is called a (ε,δ,d) oblivious $U \to \ell_2$ subspace embedding if it is an ε-embedding for an arbitrary d-dimensional subspace $V \subset U$ with probability at least $1-\delta$.*

Oblivious $\ell_2 \to \ell_2$ subspace embeddings (defined by Definition 3.2.2 with $\langle\cdot,\cdot\rangle_U := \langle\cdot,\cdot\rangle_2$) include the rescaled Gaussian distribution, the rescaled Rademacher distribution, the Subsampled Randomized Hadamard Transform (SRHT), the Subsampled Randomized Fourier Transform (SRFT), CountSketch matrix, SRFT combined with sequences of random Givens rotations, and others [12, 88, 131, 157]. In this work we shall rely on the rescaled Gaussian distribution and SRHT.

An oblivious $U \to \ell_2$ subspace embedding for a general inner product $\langle \cdot, \cdot \rangle_U$ can be constructed as

$$\mathbf{\Theta} = \mathbf{\Omega} \mathbf{Q}, \tag{3.5}$$

where $\mathbf{\Omega}$ is a $\ell_2 \to \ell_2$ subspace embedding and $\mathbf{Q} \in \mathbb{K}^{s \times n}$ is an easily computable (possibly rectangular) matrix such that $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{R}_U$ (see Remark 2.2.7).

It follows that an $U \to \ell_2$ ε-subspace embedding for V can be obtained with high probability as a realization of an oblivious subspace embedding of sufficiently large size.

The a priori estimates for the required size of $\mathbf{\Theta}$ are usually pessimistic for practical use. Moreover, a good performance of certain random embeddings (e.g., matrices with sequences of random Givens rotations) was validated only empirically [88]. Therefore, in Section 3.6 we provide a simple and reliable a posteriori procedure for characterizing the quality of an embedding for each given subspace. Such a procedure can be used for the adaptive selection of the number of rows for $\mathbf{\Theta}$ or for certifying the quality of the sketched reduced order model.

3.2.2 A sketch of a reduced model

Here the output of a reduced order model is efficiently estimated from its random sketch. The $\mathbf{\Theta}$-sketch of a reduced model associated with a subspace U_r is defined as

$$\left\{ \left\{ \mathbf{\Theta}\mathbf{x}, \mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{r}(\mathbf{x};\mu) \right\} : \quad \mathbf{x} \in U_r \right\}, \tag{3.6}$$

where $\mathbf{r}(\mathbf{x};\mu) := \mathbf{b}(\mu) - \mathbf{A}(\mu)\mathbf{x}$. Let $\mathbf{U}_r \in \mathbb{K}^{n \times r}$ be a matrix whose columns form a basis of U_r. Then each element of (3.6) can be characterized from the coordinates of $\mathbf{x}$ associated with $\mathbf{U}_r$, i.e., a vector $\mathbf{a}_r \in \mathbb{K}^r$ such that $\mathbf{x} = \mathbf{U}_r\mathbf{a}_r$, and the following quantities

$$\mathbf{U}_r^{\mathbf{\Theta}} := \mathbf{\Theta}\mathbf{U}_r, \ \mathbf{V}_r^{\mathbf{\Theta}}(\mu) := \mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_r \text{ and } \ \mathbf{b}^{\mathbf{\Theta}}(\mu) := \mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{b}(\mu). \tag{3.7}$$

Clearly $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ have affine expansions containing at most as many terms as the ones of $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$, respectively. The matrix $\mathbf{U}_r^{\mathbf{\Theta}}$ and the affine expansions of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ are referred to as the $\mathbf{\Theta}$-sketch of $\mathbf{U}_r$ (a representation of the $\mathbf{\Theta}$-sketch of a reduced model associated with U_r). With a good choice of an oblivious embedding, a $\mathbf{\Theta}$-sketch of $\mathbf{U}_r$ can be efficiently precomputed in any computational environment (see Section 2.4.4). Thereafter, an approximation of a reduced order model can be obtained with a negligible computational cost. Note that in Chapter 2 the affine expansion of $\mathbf{l}_r(\mu)^{\mathrm{H}} := \mathbf{U}_r^{\mathrm{H}}\mathbf{l}(\mu)$, where $\mathbf{l}(\mu) \in U'$ is an extractor of the linear quantity of interest, are also considered as a part of the $\mathbf{\Theta}$-sketch of $\mathbf{U}_r$. In the present chapter, however, we consider a more general scenario where the computation of the affine expansion of $\mathbf{l}_r(\mu)$ or its online evaluation can be too expensive (e.g., when $\mathbf{l}_r(\mu)$ has too many terms in the affine expansion) and has to be avoided.

Therefore, instead of computing the output quantity associated with the solution of the reduced model, we shall approximate it using

$$\mathbf{l}^{\mathbf{\Theta}}(\mu) := \mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{l}(\mu)$$

along with a few additional efficiently computable quantities (see Section 3.5 for more details). This procedure can also allow an efficient approximation of quadratic quantities of interest and primal-dual corrections.

3.3 Minimal residual projection

In this section we first present the standard minimal residual projection in a form that allows an easy introduction of random sketching. Then we introduce the sketched version of the minimal residual projection and provide conditions to guarantee its quality.

3.3.1 Standard minimal residual projection

Let $U_r \subset U$ be a subspace of U (typically obtained with a greedy algorithm or approximate POD). The minres approximation $\mathbf{u}_r(\mu)$ of $\mathbf{u}(\mu) \in U$ can be defined by

$$\mathbf{u}_r(\mu) = \arg\min_{\mathbf{w}\in U_r} \|\mathbf{r}(\mathbf{w};\mu)\|_{U'}. \tag{3.8}$$

For linear problems it is equivalently characterized by the following (Petrov-)Galerkin orthogonality condition:

$$\langle \mathbf{r}(\mathbf{u}_r(\mu);\mu), \mathbf{w}\rangle = 0, \ \forall \mathbf{w} \in V_r(\mu), \tag{3.9}$$

where $V_r(\mu) := \{\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{x} : \mathbf{x} \in U_r\}$.

If the operator $\mathbf{A}(\mu)$ is invertible then (3.8) is well-posed. In order to characterize the quality of the projection $\mathbf{u}_r(\mu)$ we define the following parameter-dependent constants

$$\zeta_r(\mu) := \min_{\mathbf{x}\in(\mathrm{span}\{\mathbf{u}(\mu)\}+U_r)\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}}{\|\mathbf{x}\|_U}, \tag{3.10a}$$

$$\iota_r(\mu) := \max_{\mathbf{x}\in(\mathrm{span}\{\mathbf{u}(\mu)\}+U_r)\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}}{\|\mathbf{x}\|_U}. \tag{3.10b}$$

Let $\mathbf{P}_W : U \to W$ denote the orthogonal projection from U on a subspace $W \subset U$, defined for $\mathbf{x} \in U$ by

$$\mathbf{P}_W\mathbf{x} = \arg\min_{\mathbf{w}\in W} \|\mathbf{x}-\mathbf{w}\|_U.$$

Proposition 3.3.1. *If $\mathbf{u}_r(\mu)$ satisfies* (3.8) *and $\zeta_r(\mu) > 0$, then*

$$\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \leq \frac{\iota_r(\mu)}{\zeta_r(\mu)} \|\mathbf{u}(\mu) - \mathbf{P}_{U_r}\mathbf{u}(\mu)\|_U. \tag{3.11}$$

Proof. See appendix. □

The constants $\zeta_r(\mu)$ and $\iota_r(\mu)$ can be bounded by the minimal and maximal singular values of $\mathbf{A}(\mu)$:

$$\alpha(\mu) := \min_{\mathbf{x} \in U \backslash \{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}}{\|\mathbf{x}\|_U} \leq \zeta_r(\mu), \tag{3.12a}$$

$$\beta(\mu) := \max_{\mathbf{x} \in U \backslash \{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}}{\|\mathbf{x}\|_U} \geq \iota_r(\mu). \tag{3.12b}$$

Bounds of $\alpha(\mu)$ and $\beta(\mu)$ can be obtained theoretically [82] or numerically with the successive constraint method [93].

For each μ, the vector $\mathbf{a}_r(\mu) \in \mathbb{K}^r$ such that $\mathbf{u}_r(\mu) = \mathbf{U}_r\mathbf{a}_r(\mu)$ satisfies (3.9) can be obtained by solving the following reduced system of equations:

$$\mathbf{A}_r(\mu)\mathbf{a}_r(\mu) = \mathbf{b}_r(\mu), \tag{3.13}$$

where $\mathbf{A}_r(\mu) = \mathbf{U}_r^{\mathrm{H}}\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_r \in \mathbb{K}^{r \times r}$ and $\mathbf{b}_r(\mu) = \mathbf{U}_r^{\mathrm{H}}\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{R}_U^{-1}\mathbf{b}(\mu) \in \mathbb{K}^r$. The numerical stability of (3.13) can be ensured through orthogonalization of $\mathbf{U}_r$ similarly as for the classical Galerkin projection. Such orthogonalization yields the following bound for the condition number of $\mathbf{A}_r(\mu)$:

$$\kappa(\mathbf{A}_r(\mu)) := \|\mathbf{A}_r(\mu)\| \|\mathbf{A}_r(\mu)^{-1}\| \leq \left(\frac{\iota_r(\mu)}{\zeta_r(\mu)}\right)^2 \leq \left(\frac{\beta(\mu)}{\alpha(\mu)}\right)^2. \tag{3.14}$$

This bound can be insufficient for problems with matrix $\mathbf{A}(\mu)$ having a high or even moderate condition number.

The random sketching technique can be used to improve the efficiency and numerical stability of the minimal residual projection, as shown below.

3.3.2 Sketched minimal residual projection

Let $\mathbf{\Theta} \in \mathbb{K}^{k \times n}$ be a certain $U \to \ell_2$ subspace embedding. The sketched minres projection can be defined by (3.8) with the dual norm $\|\cdot\|_{U'}$ replaced by its estimation $\|\cdot\|_{U'}^{\mathbf{\Theta}}$, which results in an approximation

$$\mathbf{u}_r(\mu) = \arg\min_{\mathbf{w} \in U_r} \|\mathbf{r}(\mathbf{w}; \mu)\|_{U'}^{\mathbf{\Theta}}. \tag{3.15}$$

The quasi-optimality of such a projection can be controlled in exactly the same manner as the quasi-optimality of the original minres projection. By defining the constants

$$\zeta_r^{\mathbf{\Theta}}(\mu) := \min_{\mathbf{x} \in (\operatorname{span}\{\mathbf{u}(\mu)\}+U_r)\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|^{\mathbf{\Theta}}_{U'}}{\|\mathbf{x}\|_U}, \tag{3.16a}$$

$$\iota_r^{\mathbf{\Theta}}(\mu) := \max_{\mathbf{x} \in (\operatorname{span}\{\mathbf{u}(\mu)\}+U_r)\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|^{\mathbf{\Theta}}_{U'}}{\|\mathbf{x}\|_U}, \tag{3.16b}$$

we obtain the following result.

Proposition 3.3.2. *If $\mathbf{u}_r(\mu)$ satisfies (3.15) and $\zeta_r^{\mathbf{\Theta}}(\mu) > 0$, then*

$$\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \leq \frac{\iota_r^{\mathbf{\Theta}}(\mu)}{\zeta_r^{\mathbf{\Theta}}(\mu)} \|\mathbf{u}(\mu) - \mathbf{P}_{U_r}\mathbf{u}(\mu)\|_U. \tag{3.17}$$

Proof. See appendix. □

It follows that if $\zeta_r^{\mathbf{\Theta}}(\mu)$ and $\iota_r^{\mathbf{\Theta}}(\mu)$ are almost equal to $\zeta_r(\mu)$ and $\iota_r(\mu)$, respectively, then the quasi-optimality of the original minres projection (3.8) shall be almost preserved by its sketched version (3.15). These properties of $\iota_r^{\mathbf{\Theta}}(\mu)$ and $\zeta_r^{\mathbf{\Theta}}(\mu)$ can be guaranteed under some conditions on $\mathbf{\Theta}$ (see Proposition 3.3.3).

Proposition 3.3.3. *Define the subspace*

$$R_r(U_r;\mu) := \operatorname{span}\{\mathbf{R}_U^{-1}\mathbf{r}(\mathbf{x};\mu) : \mathbf{x} \in U_r\}. \tag{3.18}$$

If $\mathbf{\Theta}$ is a $U \to \ell_2$ ε-subspace embedding for $R_r(U_r;\mu)$, then

$$\sqrt{1-\varepsilon}\ \zeta_r(\mu) \leq \zeta_r^{\mathbf{\Theta}}(\mu) \leq \sqrt{1+\varepsilon}\ \zeta_r(\mu), \text{ and } \sqrt{1-\varepsilon}\ \iota_r(\mu) \leq \iota_r^{\mathbf{\Theta}}(\mu) \leq \sqrt{1+\varepsilon}\ \iota_r(\mu). \tag{3.19}$$

Proof. See appendix. □

An embedding $\mathbf{\Theta}$ satisfying an $U \to \ell_2$ ε-subspace embedding property for the subspace $R_r(U_r;\mu)$ defined in (3.18), for all $\mu \in \mathcal{P}$ simultaneously, with high probability, may be generated from an oblivious embedding of sufficiently large size. Note that $\dim(R_r(U_r;\mu)) \leq r+1$. The number of rows k of the oblivious embedding may be selected a priori using the bounds provided in Chapter 2, along with a union bound for the probability of success or the fact that $\bigcup_{\mu\in\mathcal{P}} R_r(U_r;\mu)$ is contained in a low-dimensional space. Alternatively, a better value for k can be chosen with a posteriori procedure explained in Section 3.6. Note that if (3.19) is satisfied then the quasi-optimality constants of the minres projection are guaranteed to be preserved up to a small factor depending only on the value of ε. Since $\mathbf{\Theta}$ is here constructed in an oblivious way, the accuracy of random sketching for minres projection can be controlled regardless of the properties of $\mathbf{A}(\mu)$ (e.g., coercivity, condition number, etc.). Recall that in Chapter 2 it was revealed that the preservation of the quasi-optimality constants of the classical Galerkin projection by its sketched version is sensitive to the operator's properties. More specifically, random sketching can worsen quasi-optimality constants dramatically for non-coercive or ill-conditioned problems. Therefore, due to its remarkable advantages, the sketched minres projection should be preferred to the sketched Galerkin projection.

Remark 3.3.4. *Random sketching is not the only way to construct $\mathbf{\Theta}$ which satisfies the condition in Proposition 3.3.3 for all $\mu \in \mathcal{P}$. Such a sketching matrix can also be obtained deterministically through approximation of the manifold $R_r^*(U_r) = \{\mathbf{x} \in R_r(U_r, \mu) : \|\mathbf{x}\|_U = 1, \ \mu \in \mathcal{P}\}$. This approximation can be performed using POD or greedy algorithms. There are two main advantages of random sketching over the deterministic approaches. First, random sketching allows drastic reduction of the computational costs in the offline stage. The second advantage is the oblivious construction of $\mathbf{\Theta}$ without the knowledge of U_r, which can be particularly important when U_r is constructed adaptively (e.g., with a greedy algorithm). Note that the condition in Proposition 3.3.3 can be satisfied (for not too small ε, say $\varepsilon = 0.1$) with high probability by using an oblivious embedding with $\mathcal{O}(r)$ rows, which is close to the minimal possible value $k = r$. Therefore, the construction of $\mathbf{\Theta}$ with random sketching in general should be preferred over the deterministic construction.*

The vector of coordinates $\mathbf{a}_r(\mu) \in \mathbb{K}^r$ in the basis $\mathbf{U}_r$ of the sketched projection $\mathbf{u}_r(\mu)$ defined by (3.15) may be obtained in a classical way, i.e., by considering a parameter-dependent reduced system of equations similar to (3.13). As for the classical approach, this may lead to numerical instabilities during either the online evaluation of the reduced system from the affine expansions or its solution. A remedy is to directly consider

$$\mathbf{a}_r(\mu) = \arg\min_{\mathbf{x} \in \mathbb{K}^r} \|\mathbf{A}(\mu)\mathbf{U}_r\mathbf{x} - \mathbf{b}(\mu)\|_{U'}^{\mathbf{\Theta}} = \arg\min_{\mathbf{x} \in \mathbb{K}^r} \|\mathbf{V}_r^{\mathbf{\Theta}}(\mu)\mathbf{x} - \mathbf{b}^{\mathbf{\Theta}}(\mu)\|_2. \tag{3.20}$$

Since the sketched matrix $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ and vector $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ are of rather small sizes, the minimization problem (3.20) may be efficiently formed (from the precomputed affine expansions) and then solved (e.g., using QR factorization or SVD) in the online stage.

Proposition 3.3.5. *If $\mathbf{\Theta}$ is an ε-embedding for U_r, and $\mathbf{U}_r$ is orthogonal with respect $\langle \cdot, \cdot \rangle_U^{\mathbf{\Theta}}$ then the condition number of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ is bounded by $\sqrt{\frac{1+\varepsilon}{1-\varepsilon}} \frac{\iota_r^{\mathbf{\Theta}}}{\zeta_r^{\mathbf{\Theta}}}$.*

Proof. See appendix. □

It follows from Proposition 3.3.5 (along with Proposition 3.3.3) that considering (3.20) can provide better numerical stability than solving reduced systems of equations with standard methods. Furthermore, since affine expansions of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ have less terms than affine expansions of $\mathbf{A}_r(\mu)$ and $\mathbf{b}_r(\mu)$ in (3.13), their online assembling should also be much more stable.

The online efficiency can be further improved with a procedure similar to the one depicted in Section 2.4.3. Consider the following oblivious $U \to \ell_2$ subspace embedding

$$\mathbf{\Phi} = \mathbf{\Gamma}\mathbf{\Theta},$$

where $\mathbf{\Gamma} \in \mathbb{K}^{k' \times n}$, $k' < k$, is a small $(\varepsilon', \delta', r+1)$ oblivious $\ell_2 \to \ell_2$ subspace embedding. For a given value of the parameter, the solution to (3.20) can be accurately estimated by

$$\mathbf{u}_r(\mu) \approx \arg\min_{\mathbf{w} \in U_r} \|\mathbf{r}(\mathbf{w};\mu)\|_{U'}^{\mathbf{\Phi}}, \tag{3.21}$$

with a probability of at least $1-\delta'$. Note that by Section 2.3.1, $k' = \mathcal{O}(r)$ (in practice, with a small constant, say $k' = 3r$) is enough to provide an accurate estimation of (3.15) with high probability. For online efficiency, we can use a fixed $\mathbf{\Theta}$ such that (3.15) is guaranteed to provide an accurate approximation (see Proposition 3.3.3) for all $\mu \in \mathcal{P}$ simultaneously, but consider different realizations of a smaller matrix $\mathbf{\Gamma}$ for each particular test set $\mathcal{P}_{\text{test}}$ composed of several parameter values. In this way, in the offline stage a $\mathbf{\Theta}$-sketch of $\mathbf{U}_r$ can be precomputed and maintained for the online computations. Thereafter, for the given test set $\mathcal{P}_{\text{test}}$ (with the corresponding new realization of $\mathbf{\Gamma}$) the affine expansions of small matrices $\mathbf{V}^{\mathbf{\Phi}}(\mu) := \mathbf{\Gamma}\mathbf{V}^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Phi}}(\mu) := \mathbf{\Gamma}\mathbf{b}^{\mathbf{\Theta}}(\mu)$ can be efficiently precomputed from the $\mathbf{\Theta}$-sketch in the "intermediate" online stage. And finally, for each $\mu \in \mathcal{P}_{\text{test}}$, the vector of coordinates of $\mathbf{u}_r(\mu)$ can be obtained by evaluating $\mathbf{V}^{\mathbf{\Phi}}(\mu)$ and $\mathbf{b}^{\mathbf{\Phi}}(\mu)$ from just precomputed affine expansions, and solving

$$\mathbf{a}_r(\mu) = \arg\min_{\mathbf{x} \in \mathbb{K}^r} \|\mathbf{V}_r^{\mathbf{\Phi}}(\mu)\mathbf{x} - \mathbf{b}^{\mathbf{\Phi}}(\mu)\|_2 \tag{3.22}$$

with a standard method such as QR factorization or the SVD.

3.4 Dictionary-based minimal residual method

Classical RB method becomes ineffective for parameter-dependent problems for which the solution manifold $\mathcal{M} := \{\mathbf{u}(\mu) : \mu \in \mathcal{P}\}$ cannot be well approximated by a single low-dimensional subspace, i.e., its Kolmogorov r-width does not decay rapidly. One can extend the classical RB method by considering a reduced subspace $U_r(\mu)$ depending on a parameter μ. One way to obtain $U_r(\mu)$ is to use a hp-refinement method as in [68, 69], which consists in partitioning the parameter set $\mathcal{P}$ into subsets $\{\mathcal{P}_i\}_{i=1}^{M}$ and in associating to each subset $\mathcal{P}_i$ a subspace $U_r^i \subset U$ of dimension at most r, therefore resulting in $U_r(\mu) = U_r^i$ if $\mu \in \mathcal{P}_i$, $1 \leq i \leq M$. More formally, the *hp*-refinement method aims to approximate $\mathcal{M}$ with a library $\mathcal{L}_r := \{U_r^i : 1 \leq i \leq M\}$ of low-dimensional subspaces. For efficiency, the number of subspaces in $\mathcal{L}_r$ has to be moderate (no more than $\mathcal{O}(r^\nu)$ for some small ν, say $\nu = 2$ or 3, which should be dictated by the particular computational architecture). A nonlinear Kolmogorov r-width of $\mathcal{M}$ with a library of M subspaces can be defined as in [144] by

$$d_r(\mathcal{M}; M) = \inf_{\#\mathcal{L}_r = M} \sup_{\mathbf{u} \in \mathcal{M}} \min_{W_r \in \mathcal{L}_r} \|\mathbf{u} - \mathbf{P}_{W_r}\mathbf{u}\|_U, \tag{3.23}$$

where the infimum is taken over all libraries of M subspaces. Clearly, the approximation $\mathbf{P}_{U_r(\mu)}\mathbf{u}(\mu)$ over a parameter-dependent subspace $U_r(\mu)$ associated with a partitioning of $\mathcal{P}$ into M subdomains satisfies

$$d_r(\mathcal{M}; M) \leq \max_{\mu \in \mathcal{P}} \|\mathbf{u}(\mu) - \mathbf{P}_{U_r(\mu)}\mathbf{u}(\mu)\|_U. \tag{3.24}$$

Therefore, for the *hp*-refinement method to be effective, the solution manifold is required to be well approximable in terms of the measure $d_r(\mathcal{M}; M)$.

The *hp*-refinement method may present serious drawbacks: it can be highly sensitive to the parametrization, it can require a large number of subdomains in $\mathcal{P}$ (especially for high-dimensional parameter domains) and it can require computing too many solution samples. These drawbacks can be partially reduced by various modifications of the *hp*-refinement method [8, 109], but not circumvented.

We here propose a dictionary-based method, which can be seen as an alternative to a partitioning of $\mathcal{P}$ for defining $U_r(\mu)$, and argue why this method is more natural and can be applied to a larger class of problems.

3.4.1 Dictionary-based approximation

For each value μ of the parameter, the basis vectors for $U_r(\mu)$ are selected from a certain dictionary $\mathcal{D}_K$ of K candidate vectors in U, $K \geq r$. For efficiency of the algorithms in the particular computational environment, the value for K has to be chosen as $\mathcal{O}(r^\nu)$ with a small ν similarly as the number of subdomains M for the *hp*-refinement method. Let $\mathcal{L}_r(\mathcal{D}_K)$ denote the library of all subspaces spanned by r vectors from $\mathcal{D}_K$. A dictionary-based r-width is defined as

$$\sigma_r(\mathcal{M}; K) = \inf_{\#\mathcal{D}_K = K} \sup_{\mathbf{u} \in \mathcal{M}} \min_{W_r \in \mathcal{L}_r(\mathcal{D}_K)} \|\mathbf{u} - \mathbf{P}_{W_r}\mathbf{u}\|_U, \tag{3.25}$$

where the infimum is taken over all subsets $\mathcal{D}_K$ of U with cardinality $\#\mathcal{D}_K = K$. A dictionary $\mathcal{D}_K$ can be efficiently constructed offline with an adaptive greedy procedure (see Section 3.4.4).

In general, the performance of the method can be characterized through the approximability of the solution manifold $\mathcal{M}$ in terms of the r-width, and quasi-optimality of the considered $U_r(\mu)$ compared to the best approximation. The dictionary-based approximation can be beneficial over the refinement methods in either of these aspects, which is explained below.

It can be easily shown that

$$\sigma_r(\mathcal{M};K) \leq d_r(\mathcal{M};M), \text{ for } K \geq rM.$$

Therefore, if a solution manifold can be well approximated with a partitioning of the parameter domain into M subdomains each associated with a subspace of dimension r, then it should also be well approximated with a dictionary of size $K = rM$, which implies a similar computational cost. The converse statement, however, is not true. A dictionary with K vectors can generate a library with up to $\binom{K}{r}$ subspaces so that

$$d_r(\mathcal{M}; \tbinom{K}{r}) \leq \sigma_r(\mathcal{M};K).$$

Consequently, to obtain a decay of $d_r(\mathcal{M};M)$ with r similar to the decay of $\sigma_r(\mathcal{M};r^\nu)$, we can be required to use M which depends exponentially on r.

The great potential of the dictionary-based approximation can be justified by important properties of the dictionary-based r-width given in Proposition 3.4.1 and Corollary 3.4.2.

Proposition 3.4.1. *Let $\mathcal{M}$ be obtained by the superposition of parameter-dependent vectors:*

$$\mathcal{M} = \{\sum_{i=1}^{l} \mathbf{u}^{(i)}(\mu) : \mu \in \mathcal{P}\}, \tag{3.26}$$

where $\mathbf{u}^{(i)}(\mu) \in U, \quad i = 1, \dots, l$. Then, we have

$$\sigma_r(\mathcal{M};K) \leq \sum_{i=1}^{l} \sigma_{r_i}(\mathcal{M}^{(i)};K_i), \tag{3.27}$$

with $r = \sum_{i=1}^{l} r_i, \ K = \sum_{i=1}^{l} K_i$ and

$$\mathcal{M}^{(i)} = \{\mathbf{u}^{(i)}(\mu) : \mu \in \mathcal{P}\}. \tag{3.28}$$

Proof. See appendix. □

Corollary 3.4.2 (Approximability of a superposition of solutions). *Consider several solution manifolds $\mathcal{M}^{(i)}$ defined by (3.28), $1 \leq i \leq l$, and the resulting manifold $\mathcal{M}$ defined by (3.26). Let c, C, α, β and γ be some constants. The following properties hold.*

(i) If $\sigma_r(\mathcal{M}^{(i)};cr^\nu) \leq Cr^{-\alpha}$, then $\sigma_r(\mathcal{M};cl^{1-\nu}r^\nu) \leq Cl^{1+\alpha}r^{-\alpha}$,

(ii) If $\sigma_r(\mathcal{M}^{(i)}; cr^\nu) \leq Ce^{-\gamma r^\beta}$, *then* $\sigma_r(\mathcal{M}; cl^{1-\nu} r^\nu) \leq Cle^{-\gamma l^{-\beta} r^\beta}$.

From Proposition 3.4.1 and Corollary 3.4.2 it follows that the approximability of the solution manifold in terms of the dictionary-based r-width is preserved under the superposition operation. In other words, if the dictionary-based r-widths of manifolds $\mathcal{M}^{(i)}$ have a certain decay with r (e.g., exponential or algebraic), by using dictionaries containing $K = \mathcal{O}(r^\nu)$ vectors, then the type of decay is preserved by their superposition (with the same rate for the algebraic decay). This property can be crucial for problems where the solution is a superposition of several contributions (possibly unknown), which is a quite typical situation. For instance, we have such a situation for PDEs with multiple transport phenomena. A similar property as (3.27) also holds for the classical linear Kolmogorov r-width $d_r(\mathcal{M})$. Namely, we have

$$d_r(\mathcal{M}) \leq \sum_{i=1}^{l} d_{r_i}(\mathcal{M}^{(i)}), \tag{3.29}$$

with $r = \sum_{i=1}^l r_i$. This relation follows immediately from Proposition 3.4.1 and the fact that $d_r(\mathcal{M}) = \sigma_r(\mathcal{M}; 1)$. For the nonlinear Kolmogorov r-width (3.24), however, the relation

$$d_r(\mathcal{M}, M) \leq \sum_{i=1}^{l} d_{r_i}(\mathcal{M}^{(i)}, M^{(i)}), \tag{3.30}$$

where $r = \sum_{i=1}^l r_i$, holds under the condition that $M \geq \prod_{i=1}^l M^{(i)}$. In general, the preservation of the type of decay with r of $d_r(\mathcal{M}, M)$, by using libraries with $M = \mathcal{O}(r^\nu)$ terms, may not be guaranteed. It can require libraries with much larger numbers of r-dimensional subspaces than $\mathcal{O}(r^\nu)$, namely $M = \mathcal{O}(r^{l\nu})$ subspaces.

Another advantage of the dictionary-based method is its weak sensitivity to the parametrization of the manifold $\mathcal{M}$, in contrast to the *hp*-refinement method, for which a bad choice of parametrization can result in approximations with too many local reduced subspaces. Indeed, the solution map $\mu \to \mathbf{u}(\mu)$ is often expected to have certain properties (e.g., symmetries or anisotropies) that yield the existence of a better parametrization of $\mathcal{M}$ than the one proposed by the user. Finding a good parametrization of the solution manifold may require a deep intrusive analysis of the problem, and is therefore usually an unfeasible task. On the other hand, our dictionary-based methodology provides a reduced approximation subspace for each vector from $\mathcal{M}$ regardless of the chosen parametrization.

3.4.2 Sparse minimal residual approximation

Here we assume to be given a dictionary $\mathcal{D}_K$ of K vectors in U. Ideally, for each μ, $\mathbf{u}(\mu)$ should be approximated by orthogonal projection onto a subspace $W_r(\mu)$ that minimizes

$$\|\mathbf{u}(\mu) - \mathbf{P}_{W_r(\mu)} \mathbf{u}(\mu)\|_U \tag{3.31}$$

over the library $\mathcal{L}_r(\mathcal{D}_K)$. The selection of the optimal subspace requires operating with the exact solution $\mathbf{u}(\mu)$ which is prohibited. Therefore, the reduced approximation space $U_r(\mu)$ and the associated approximate solution $\mathbf{u}_r(\mu) \in U_r(\mu)$ are defined such that

$$U_r(\mu) \in \arg \min_{W_r \in \mathcal{L}_r(\mathcal{D}_K)} \min_{\mathbf{w} \in W_r} \|\mathbf{r}(\mathbf{w};\mu)\|_{U'}, \quad \mathbf{u}_r(\mu) = \arg \min_{\mathbf{w} \in U_r(\mu)} \|\mathbf{r}(\mathbf{w};\mu)\|_{U'}. \tag{3.32}$$

The solution $\mathbf{u}_r(\mu)$ from (3.32) shall be referred to as sparse minres approximation (relatively to the dictionary $\mathcal{D}_K$). The quasi-optimality of this approximation can be characterized with the following parameter-dependent constants:

$$\zeta_{r,K}(\mu) := \min_{W_r \in \mathcal{L}_r(\mathcal{D}_K)} \min_{\mathbf{x} \in (\mathrm{span}\{\mathbf{u}(\mu)\} + W_r) \backslash \{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}}{\|\mathbf{x}\|_U}, \tag{3.33a}$$

$$\iota_{r,K}(\mu) := \max_{W_r \in \mathcal{L}_r(\mathcal{D}_K)} \max_{\mathbf{x} \in (\mathrm{span}\{\mathbf{u}(\mu)\} + W_r) \backslash \{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}}{\|\mathbf{x}\|_U}. \tag{3.33b}$$

In general, one can bound $\zeta_{r,K}(\mu)$ and $\iota_{r,K}(\mu)$ by the minimal and the maximal singular values $\alpha(\mu)$ and $\beta(\mu)$ of $\mathbf{A}(\mu)$. Observe also that for $K = r$ (i.e., when the library $\mathcal{L}_r(\mathcal{D}_K) = \{U_r\}$ has a single subspace) we have $\zeta_{r,K}(\mu) = \zeta_r(\mu)$ and $\iota_{r,K}(\mu) = \iota_r(\mu)$.

Proposition 3.4.3. *Let $\mathbf{u}_r(\mu)$ be the solution of (3.32) and $\zeta_{r,K}(\mu) > 0$, then*

$$\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \leq \frac{\iota_{r,K}(\mu)}{\zeta_{r,K}(\mu)} \min_{W_r \in \mathcal{L}_r(\mathcal{D}_K)} \|\mathbf{u}(\mu) - \mathbf{P}_{W_r}\mathbf{u}(\mu)\|_U. \tag{3.34}$$

Proof. See appendix. □

Let $\mathbf{U}_K \in \mathbb{K}^{n \times K}$ be a matrix whose columns are the vectors in the dictionary $\mathcal{D}_K$ and $\mathbf{a}_{r,K}(\mu) \in \mathbb{K}^K$, with $\|\mathbf{a}_{r,K}(\mu)\|_0 \leq r$, be the r-sparse vector of coordinates of $\mathbf{u}_r(\mu)$ in the dictionary, i.e., $\mathbf{u}_r(\mu) = \mathbf{U}_K \mathbf{a}_{r,K}(\mu)$. The vector of coordinates associated with the solution $\mathbf{u}_r(\mu)$ of (3.32) is the solution to the following parameter-dependent sparse least-squares problem:

$$\min_{\mathbf{z} \in \mathbb{K}^K} \|\mathbf{A}(\mu)\mathbf{U}_K\mathbf{z} - \mathbf{b}(\mu)\|_{U'}, \text{ subject to } \|\mathbf{z}\|_0 \leq r. \tag{3.35}$$

For each $\mu \in \mathcal{P}$ an approximate solution to problem (3.35) can be obtained with a standard greedy algorithm depicted in Algorithm 4. It selects the nonzero entries of $\mathbf{a}_{r,K}(\mu)$ one by one to minimize the residual. The algorithm corresponds to either the orthogonal greedy (also called Orthogonal Matching Pursuit in signal processing community [149]) or stepwise projection algorithm (see [61]) depending on whether the (optional) Step 8 (which is the orthogonalization of $\{\mathbf{v}_j(\mu)\}_{j=1}^K$ with respect to $V_i(\mu)$)

is considered. It should be noted that performing Step 8 can be of great importance due to possible high mutual coherence of the dictionary $\{\mathbf{v}_j(\mu)\}_{j=1}^{K}$. Algorithm 4 is provided in a conceptual form. A more sophisticated procedure can be derived to improve the online efficiency (e.g., considering precomputed affine expansions of $\mathbf{A}_K(\mu) := \mathbf{U}_K^{\mathrm{H}}\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_K \in \mathbb{K}^{K\times K}$ and $\mathbf{b}_K(\mu) = \mathbf{U}_K^{\mathrm{H}}\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{R}_U^{-1}\mathbf{b}(\mu) \in \mathbb{K}^{K}$, updating the residual using a Gram-Schmidt procedure, etc). Algorithm 4, even when efficiently implemented, can still require heavy computations in both the offline and online stages, and be numerically unstable. One of the contributions of this chapter is a drastic improvement of its efficiency and stability by random sketching, thus making the use of dictionary-based model reduction feasible in practice.

Algorithm 4 Orthogonal greedy algorithm

```
Given: $\mu$, $\mathbf{U}_K = [\mathbf{w}_j]_{j=1}^{K}$, $\mathbf{A}(\mu)$, $\mathbf{b}(\mu)$, $\tau$, $r$.
Output: index set $\Lambda_r(\mu)$, the coordinates $\mathbf{a}_r(\mu)$ of $\mathbf{u}_r(\mu)$ on the basis $\{\mathbf{w}_j\}_{j\in\Lambda_r(\mu)}$.
1. Set $i := 0$, $U_0(\mu) = \{\mathbf{0}\}$, $\mathbf{u}_0(\mu) = \mathbf{0}$, $\Lambda_0(\mu) = \emptyset$, $\widetilde{\Delta}_0(\mu) = \infty$.
2. Set $[\mathbf{v}_1(\mu), ..., \mathbf{v}_K(\mu)] := \mathbf{A}(\mu)\mathbf{U}_K$ and normalize the vectors $\mathbf{v}_j(\mu)$, $1 \le j \le K$.
while $\widetilde{\Delta}_i(\mu) \ge \tau$ and $i < r$ do
    3. Set $i := i+1$.
    4. Find the index $p_i \in \{1,\dots,K\}$ which maximizes $|\langle \mathbf{v}_{p_i}(\mu), \mathbf{r}(\mathbf{u}_{i-1}(\mu);\mu)\rangle_{U'}|$.
    5. Set $\Lambda_i(\mu) := \Lambda_{i-1}(\mu) \cup \{p_i\}$.
    6. Solve (3.13) with a reduced matrix $\mathbf{U}_i(\mu) = [\mathbf{w}_j]_{j\in\Lambda_i(\mu)}$ and obtain
       the coordinates $\mathbf{a}_i(\mu)$.
    7. Compute error bound $\widetilde{\Delta}_i(\mu)$ of $\mathbf{u}_i(\mu) = \mathbf{U}_i(\mu)\mathbf{a}_i(\mu)$.
    8. (Optional) Set $\mathbf{v}_j(\mu) := \mathbf{v}_j(\mu) - \mathbf{P}_{V_i(\mu)}\mathbf{v}_j(\mu)$, where $\mathbf{P}_{V_i(\mu)}$ is the orthogonal
       projector on $V_i(\mu) := \mathrm{span}(\{\mathbf{v}_p(\mu)\}_{p\in\Lambda_i(\mu)})$, and normalize $\mathbf{v}_j(\mu)$, $1 \le j \le K$.
end while
```

3.4.3 Sketched sparse minimal residual approximation

Let $\mathbf{\Theta} \in \mathbb{K}^{k\times n}$ be a certain $U \to \ell_2$ subspace embedding. The sparse minres approximation defined by (3.32), associated with dictionary $\mathcal{D}_K$, can be estimated by the solution $\mathbf{u}_r(\mu)$ of the following minimization problem

$$U_r(\mu) \in \arg\min_{W_r\in\mathcal{L}_r(\mathcal{D}_K)} \min_{\mathbf{w}\in W_r} \|\mathbf{r}(\mathbf{w};\mu)\|_{U'}^{\mathbf{\Theta}}, \quad \mathbf{u}_r(\mu) = \arg\min_{\mathbf{w}\in U_r(\mu)} \|\mathbf{r}(\mathbf{w};\mu)\|_{U'}^{\mathbf{\Theta}}. \tag{3.36}$$

In order to characterize the quasi-optimality of the sketched sparse minres approximation defined by (3.36) we introduce the following parameter-dependent values

$$\zeta^{\mathbf{\Theta}}_{r,K}(\mu) := \min_{W_r \in \mathcal{L}_r(\mathcal{D}_K)} \min_{\mathbf{x} \in (\mathrm{span}\{\mathbf{u}(\mu)\}+W_r)\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|^{\mathbf{\Theta}}_{U'}}{\|\mathbf{x}\|_U}, \tag{3.37a}$$

$$\iota^{\mathbf{\Theta}}_{r,K}(\mu) := \max_{W_r \in \mathcal{L}_r(\mathcal{D}_K)} \max_{\mathbf{x} \in (\mathrm{span}\{\mathbf{u}(\mu)\}+W_r)\backslash\{\mathbf{0}\}} \frac{\|\mathbf{A}(\mu)\mathbf{x}\|^{\mathbf{\Theta}}_{U'}}{\|\mathbf{x}\|_U}. \tag{3.37b}$$

Observe that choosing $K = r$ yields $\zeta^{\mathbf{\Theta}}_{r,K}(\mu) = \zeta^{\mathbf{\Theta}}_{r}(\mu)$ and $\iota^{\mathbf{\Theta}}_{r,K}(\mu) = \iota^{\mathbf{\Theta}}_{r}(\mu)$.

Proposition 3.4.4. *If $\mathbf{u}_r(\mu)$ satisfies (3.36) and $\zeta^{\mathbf{\Theta}}_{r,K}(\mu) > 0$, then*

$$\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \leq \frac{\iota^{\mathbf{\Theta}}_{r,K}(\mu)}{\zeta^{\mathbf{\Theta}}_{r,K}(\mu)} \min_{W_r \in \mathcal{L}_r(\mathcal{D}_K)} \|\mathbf{u}(\mu) - \mathbf{P}_{W_r}\mathbf{u}(\mu)\|_U, \tag{3.38}$$

Proof. See appendix. □

It follows from Proposition 3.4.4 that the quasi-optimality of the sketched sparse minres approximation can be controlled by bounding the constants $\zeta^{\mathbf{\Theta}}_{r,K}(\mu)$ and $\iota^{\mathbf{\Theta}}_{r,K}(\mu)$.

Proposition 3.4.5. *If $\mathbf{\Theta}$ is a $U \to \ell_2$ ε-embedding for every subspace $R_r(W_r;\mu)$, defined by (3.18), with $W_r \in \mathcal{L}_r(\mathcal{D}_K)$, then*

$$\sqrt{1-\varepsilon}\zeta_{r,K}(\mu) \leq \zeta^{\mathbf{\Theta}}_{r,K}(\mu) \leq \sqrt{1+\varepsilon}\zeta_{r,K}(\mu), \tag{3.39a}$$

and

$$\sqrt{1-\varepsilon}\iota_{r,K}(\mu) \leq \iota^{\mathbf{\Theta}}_{r,K}(\mu) \leq \sqrt{1+\varepsilon}\iota_{r,K}(\mu). \tag{3.39b}$$

Proof. See appendix. □

By Definition 3.2.3 and the union bound for the probability of success, if $\mathbf{\Theta}$ is a $(\varepsilon, \binom{K}{r}^{-1}\delta, r+1)$ oblivious $U \to \ell_2$ subspace embedding, then $\mathbf{\Theta}$ satisfies the assumption of Proposition 3.4.5 with probability of at least $1-\delta$. The sufficient number of rows for $\mathbf{\Theta}$ may be chosen a priori with the bounds provided in Chapter 2 or adaptively with a procedure from Section 3.6. For the Gaussian embeddings the a priori bounds are logarithmic in K and n, and proportional to r. For P-SRHT they are also logarithmic in K and n, but proportional to r^2 (although in practice P-SRHT performs equally well as the Gaussian distribution). Moreover, if $\mathcal{P}$ is a finite set, an oblivious embedding $\mathbf{\Theta}$ which satisfies the hypothesis of Proposition 3.4.5 for all $\mu \in \mathcal{P}$, simultaneously, may be chosen using the above considerations and a union bound for the probability of success. Alternatively, for an infinite set $\mathcal{P}$, $\mathbf{\Theta}$ can be chosen as an ε-embedding for a collection of low-dimensional subspaces $R_r^*(W_r)$ (which can be obtained from the affine expansions of $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$) each containing $\bigcup_{\mu\in\mathcal{P}} R_r(\mu;W_r)$ and associated with a subspace W_r of $\mathcal{L}_r(\mathcal{D}_K)$. Such an embedding can be again generated in an oblivious way by considering Definition 3.2.3 and a union bound for the probability of success.

From an algebraic point of view, the optimization problem (3.36) can be formulated as the following sparse least-squares problem:

$$\min_{\substack{\mathbf{z}\in\mathbb{K}^K\\ \|\mathbf{z}\|_0\leq r}} \|\mathbf{A}(\mu)\mathbf{U}_K\mathbf{z}-\mathbf{b}(\mu)\|_{U'}^{\mathbf{\Theta}} = \min_{\substack{\mathbf{z}\in\mathbb{K}^K\\ \|\mathbf{z}\|_0\leq r}} \|\mathbf{V}_K^{\mathbf{\Theta}}(\mu)\mathbf{z}-\mathbf{b}^{\mathbf{\Theta}}(\mu)\|_2, \tag{3.40}$$

where $\mathbf{V}_K^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ are the components (3.7) of the $\mathbf{\Theta}$-sketch of $\mathbf{U}_K$ (a matrix whose columns are the vectors in $\mathcal{D}_K$). The solution $\mathbf{a}_{r,K}(\mu)$ of (3.40) is the r-sparse vector of the coordinates of $\mathbf{u}_r(\mu)$. We observe that (3.40) is simply an approximation of a small vector $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ with a dictionary composed from column vectors of $\mathbf{V}_K^{\mathbf{\Theta}}(\mu)$. Therefore, unlike the original sparse least-squares problem (3.35), the solution to its sketched version (3.40) can be efficiently approximated with standard tools in the online stage. For instance, we can use Algorithm 4 replacing $\langle\cdot,\cdot\rangle_{U'}$ with $\langle\cdot,\cdot\rangle_{U'}^{\mathbf{\Theta}}$. Clearly, in Algorithm 4 the inner products $\langle\cdot,\cdot\rangle_{U'}^{\mathbf{\Theta}}$ should be efficiently evaluated from $\mathbf{V}_K^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$. For this a $\mathbf{\Theta}$-sketch of $\mathbf{U}_K$ can be precomputed in the offline stage and then used for online evaluation of $\mathbf{V}_K^{\mathbf{\Theta}}(\mu)$ and $\mathbf{b}^{\mathbf{\Theta}}(\mu)$ for each value of the parameter.

Let us now characterize the algebraic stability (i.e., sensitivity to round-off errors) of the (approximate) solution of (3.40). The solution of (3.40) is essentially obtained from the following least-squares problem

$$\min_{\mathbf{x}\in\mathbb{K}^r} \|\mathbf{V}_r^{\mathbf{\Theta}}(\mu)\mathbf{x}-\mathbf{b}^{\mathbf{\Theta}}(\mu)\|_2, \tag{3.41}$$

where $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ is a matrix whose column vectors are (adaptively) selected from the columns of $\mathbf{V}_K^{\mathbf{\Theta}}(\mu)$. The algebraic stability of this problem can be measured by the condition number of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$. The minimal and the maximal singular values of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ can be bounded using the parameter-dependent coefficients $\iota_{r,K}^{\mathbf{\Theta}}(\mu)$, $\zeta_{r,K}^{\mathbf{\Theta}}(\mu)$ and the so-called restricted isometry property (RIP) constants associated with the dictionary $\mathcal{D}_K$, which are defined by

$$\Sigma_{r,K}^{\min} := \min_{\substack{\mathbf{z}\in\mathbb{K}^K\\ \|\mathbf{z}\|_0\leq r}} \frac{\|\mathbf{U}_K\mathbf{z}\|_U}{\|\mathbf{z}\|}, \quad \Sigma_{r,K}^{\max} := \max_{\substack{\mathbf{z}\in\mathbb{K}^K\\ \|\mathbf{z}\|_0\leq r}} \frac{\|\mathbf{U}_K\mathbf{z}\|_U}{\|\mathbf{z}\|}. \tag{3.42}$$

Proposition 3.4.6. *The minimal singular value of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ in (3.41) is bounded below by $\zeta_{r,K}^{\mathbf{\Theta}}(\mu)\Sigma_{r,K}^{\min}$, while the maximal singular value of $\mathbf{V}_r^{\mathbf{\Theta}}(\mu)$ is bounded above by $\iota_{r,K}^{\mathbf{\Theta}}(\mu)\Sigma_{r,K}^{\max}$.*

Proof. See appendix. □

The RIP constants quantify the linear dependency of the dictionary vectors. For instance, it is easy to see that for a dictionary composed of orthogonal unit vectors we have $\Sigma^{\min}_{r,K} = \Sigma^{\max}_{r,K} = 1$. From Proposition 3.4.6, one can deduce the maximal level of degeneracy of $\mathcal{D}_K$ for which the sparse optimization problem (3.40) remains sufficiently stable.

Remark 3.4.7. *In general, our approach is more stable than the algorithms from [64, 97]. These algorithms basically proceed with the solution of the reduced system of equations $\mathbf{A}_r(\mu)\mathbf{a}_r(\mu) = \mathbf{b}_r(\mu)$, where $\mathbf{A}_r(\mu) = \mathbf{U}_r(\mu)^{\mathrm{H}}\mathbf{A}(\mu)\mathbf{U}_r(\mu)$, with $\mathbf{U}_r(\mu)$ being a matrix whose column vectors are selected from the column vectors of $\mathbf{U}_K$. In this case, the bounds for the minimal and the maximal singular values of $\mathbf{A}_r(\mu)$ are proportional to the squares of the minimal and the maximal singular values of $\mathbf{U}_r(\mu)$, which implies a quadratic dependency on the RIP constants $\Sigma^{\min}_{r,K}$ and $\Sigma^{\max}_{r,K}$. On the other hand, with (sketched) minres methods the dependency of the singular values of the reduced matrix $\mathbf{V}^{\mathbf{\Theta}}(\mu)$ on $\Sigma^{\min}_{r,K}$ and $\Sigma^{\max}_{r,K}$ is only linear (see Proposition 3.4.6). Consequently, our methodology provides an improvement of not only efficiency but also numerical stability for problems with high linear dependency of dictionary vectors.*

Similarly to the sketched minres projection, a better online efficiency can be obtained by introducing

$$\mathbf{\Phi} = \mathbf{\Gamma}\mathbf{\Theta},$$

where $\mathbf{\Gamma} \in \mathbb{K}^{k' \times n}$, $k' < k$, is a small $(\varepsilon', \binom{K}{r}^{-1}\delta', r+1)$ oblivious $\ell_2 \to \ell_2$ subspace embedding, and approximating the solution to (3.36) by

$$U_r(\mu) \in \arg \min_{W_r \in \mathcal{L}_r(\mathcal{D}_K)} \min_{\mathbf{w} \in W_r} \|\mathbf{r}(\mathbf{w};\mu)\|^{\mathbf{\Phi}}_{U'}, \quad \mathbf{u}_r(\mu) = \arg \min_{\mathbf{w} \in U_r(\mu)} \|\mathbf{r}(\mathbf{w};\mu)\|^{\mathbf{\Phi}}_{U'}. \tag{3.43}$$

It follows that the accuracy (and the stability) of the solution of (3.43) is almost the same as the one of (3.36) with probability at least $1-\delta'$. In an algebraic setting, (3.43) can be expressed as

$$\min_{\substack{\mathbf{z} \in \mathbb{K}^K \\ \|\mathbf{z}\|_0 \le r}} \|\mathbf{V}^{\mathbf{\Phi}}_K(\mu)\mathbf{z} - \mathbf{b}^{\mathbf{\Phi}}(\mu)\|_2, \tag{3.44}$$

whose solution $\mathbf{a}_r(\mu)$ is a r-sparse vector of coordinates of $\mathbf{u}_r(\mu)$. An approximate solution to such a problem can be computed with Algorithm 4 by replacing $\langle \cdot, \cdot \rangle_{U'}$ with $\langle \cdot, \cdot \rangle^{\mathbf{\Phi}}_{U'}$. An efficient procedure for evaluating the coordinates of a sketched dictionary-based approximation on a test set $\mathcal{P}_{\text{test}}$ from the $\mathbf{\Theta}$-sketch of $\mathbf{U}_K$ is provided in Algorithm 5. Algorithm 5 uses a residual-based sketched error estimator from Chapter 2 defined by

$$\widetilde{\Delta}_i(\mu) = \Delta^{\mathbf{\Phi}}(\mathbf{u}_i(\mu);\mu) = \frac{\|\mathbf{r}(\mathbf{u}_i(\mu);\mu)\|^{\mathbf{\Phi}}_{U'}}{\eta(\mu)}, \tag{3.45}$$

where $\eta(\mu)$ is a computable lower bound of the minimal singular value of $\mathbf{A}(\mu)$. Let us underline the importance of performing Step 8 (orthogonalization of the dictionary vectors with respect to the previously selected basis vectors), for problems with "degenerate" dictionaries (with high mutual coherence). It should be noted that at Steps 7 and 8 we use a Gram-Schmidt procedure for orthogonalization because of its simplicity and efficiency, whereas a modified Gram-Schmidt algorithm could provide better accuracy. It is also important to note that Algorithm 5 satisfies a basic consistency property in the sense that it exactly recovers the vectors from the dictionary with high probability.

If $\mathcal{P}$ is a finite set, then the theoretical bounds for Gaussian matrices and the empirical experience for P-SRHT state that choosing $k = \mathcal{O}(r \log K + \log \delta + \log(\#\mathcal{P}))$ and $k' = \mathcal{O}(r \log K + \log \delta + \log(\#\mathcal{P}_{\text{test}}))$ in Algorithm 5 yield a quasi-optimal solution to (3.31) for all $\mu \in \mathcal{P}_{\text{test}}$ with probability at least $1-\delta$. Let us neglect the logarithmic summands. Assuming $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$ admit affine representations with m_A and m_b terms, it follows that the online complexity and memory consumption of Algorithm 5 is only $\mathcal{O}((m_A K + m_b) r \log K + r^2 K \log K) \#\mathcal{P}_{\text{test}}$ and $\mathcal{O}((m_A K + m_b) r \log K)$, respectively. The quasi-optimality for infinite $\mathcal{P}$ can be ensured with high probability by increasing k to $\mathcal{O}(r^* \log K + \log \delta)$, where r^* is the maximal dimension of subspaces $R_r^*(W_r)$ containing $\bigcup_{\mu \in \mathcal{P}} R_r(\mu; W_r)$ with $W_r \in \mathcal{L}_r(\mathcal{D}_K)$. This shall increase the memory consumption by a factor of r^*/r but should have a negligible effect (especially for large $\mathcal{P}_{\text{test}}$) on the complexity, which is mainly characterized by the size of $\mathbf{\Phi}$. Note that for parameter-separable problems we have $r^* \leq m_A r + m_b$.

3.4.4 Dictionary generation

The simplest way is to choose the dictionary as a set of solution samples (snapshots) associated with a training set $\mathcal{P}_{\text{train}}$, i.e.,

$$\mathcal{D}_K = \{\mathbf{u}(\mu) : \mu \in \mathcal{P}_{\text{train}}\}. \tag{3.46}$$

Let us recall that we are interested in computing a $\mathbf{\Theta}$-sketch of $\mathbf{U}_K$ (matrix whose columns form $\mathcal{D}_K$) rather than the full matrix. In certain computational environments, a $\mathbf{\Theta}$-sketch of $\mathbf{U}_K$ can be computed very efficiently. For instance, each snapshot may be computed and sketched on a separate distributed machine. Thereafter small sketches can be efficiently transfered to the master workstation for constructing the reduced order model.

A better dictionary may be computed with the greedy procedure presented in Algorithm 6, recursively enriching the dictionary with a snapshot at the parameter value associated with the maximal error at the previous iteration. The value for r (the dimension of the parameter-dependent reduced subspace $U_r(\mu)$) should be chosen according to the particular computational architecture. Since the provisional online solver (identified with Algorithm 5) guarantees exact recovery of snapshots belonging

Algorithm 5 Efficient/stable sketched orthogonal greedy algorithm

```
Given: $\mathbf{\Theta}$-sketch of $\mathbf{U}_K = [\mathbf{w}_j]_{j=1}^K$, $\mathcal{P}_{\text{test}}$, $\tau$.
Output: index set $\Lambda_r(\mu)$, the coordinates $\mathbf{a}_r(\mu)$ of $\mathbf{u}_r(\mu)$ on basis $\{\mathbf{w}_j\}_{j\in\Lambda_r(\mu)}$,
        and the error indicator $\Delta^{\mathbf{\Phi}}(\mathbf{u}_r(\mu);\mu)$ for each $\mu \in \mathcal{P}_{\text{test}}$.
1. Generate $\mathbf{\Gamma}$ and evaluate the affine factors of $\mathbf{V}_K^{\mathbf{\Phi}}(\mu) := \mathbf{\Gamma}\mathbf{V}_K^{\mathbf{\Theta}}(\mu)$
   and $\mathbf{b}^{\mathbf{\Phi}}(\mu) := \mathbf{\Gamma}\mathbf{b}^{\mathbf{\Theta}}(\mu)$.
for $\mu \in \mathcal{P}_{\text{test}}$ do
    2. Evaluate $[\mathbf{v}_1^{\mathbf{\Phi}}(\mu),\dots,\mathbf{v}_K^{\mathbf{\Phi}}(\mu)] := \mathbf{V}_K^{\mathbf{\Phi}}(\mu)$ and $\mathbf{b}^{\mathbf{\Phi}}(\mu)$ from the affine expansions
       and normalize $\mathbf{v}_j^{\mathbf{\Phi}}(\mu), 1 \leq j \leq K$.
    3. Set $i = 0$, obtain $\eta(\mu)$ in (3.45), set $\Lambda_0(\mu) = \emptyset$, $\mathbf{r}_0^{\mathbf{\Phi}}(\mu) := \mathbf{b}^{\mathbf{\Phi}}(\mu)$ and
       $\Delta^{\mathbf{\Phi}}(\mu) := \|\mathbf{b}^{\mathbf{\Phi}}(\mu)\|/\eta(\mu)$.
    while $\Delta^{\mathbf{\Phi}}(\mathbf{u}_i(\mu);\mu) \geq \tau$ and $i \leq r$ do
        4. Set $i := i+1$.
        5. Find the index $p_i$ which maximizes $|\mathbf{v}_{p_i}^{\mathbf{\Phi}}(\mu)^{\mathrm{H}}\mathbf{r}_{i-1}^{\mathbf{\Phi}}(\mu)|$.
           Set $\Lambda_i(\mu) := \Lambda_{i-1}(\mu) \cup \{p_i\}$.
        6. Set $\mathbf{v}_{p_i}^{\mathbf{\Phi}}(\mu) := \mathbf{v}_{p_i}^{\mathbf{\Phi}}(\mu) - \sum_{j=1}^{i-1}\mathbf{v}_{p_j}^{\mathbf{\Phi}}(\mu)[\mathbf{v}_{p_j}^{\mathbf{\Phi}}(\mu)^{\mathrm{H}}\mathbf{v}_{p_i}^{\mathbf{\Phi}}(\mu)]$ and normalize it.
        7. Compute $\mathbf{r}_i^{\mathbf{\Phi}}(\mu) := \mathbf{r}_{i-1}^{\mathbf{\Phi}}(\mu) - \mathbf{v}_{p_i}(\mu)[\mathbf{v}_{p_i}^{\mathbf{\Phi}}(\mu)^{\mathrm{H}}\mathbf{r}_{i-1}^{\mathbf{\Phi}}(\mu)]$ and
           $\Delta^{\mathbf{\Phi}}(\mathbf{u}_i(\mu);\mu) = \|\mathbf{r}_i^{\mathbf{\Phi}}(\mu)\|/\eta(\mu)$.
        8. (Optional) Set $\mathbf{v}_j^{\mathbf{\Phi}}(\mu) = \mathbf{v}_j^{\mathbf{\Phi}}(\mu) - \mathbf{v}_{p_i}^{\mathbf{\Phi}}(\mu)[\mathbf{v}_{p_i}^{\mathbf{\Phi}}(\mu)^{\mathrm{H}}\mathbf{v}_j^{\mathbf{\Phi}}(\mu)]$
           and normalize it, $1 \leq j \leq K$.
    end while
    9. Solve (3.22) choosing $r := i$, and the columns $p_1, p_2, \dots, p_i$ of $\mathbf{V}_K^{\mathbf{\Phi}}(\mu)$ as
       the columns for $\mathbf{V}_r^{\mathbf{\Phi}}(\mu)$ and obtain solution $\mathbf{a}_r(\mu)$.
end for
```

to $\mathcal{D}_i$, Algorithm 6 is consistent. It has to be noted that the first r iterations of the proposed greedy algorithm for the dictionary generation coincide with the first r iterations of the standard greedy algorithm for the reduced basis generation.

Algorithm 6 Greedy algorithm for dictionary generation

```
**Given:** $\mathcal{P}_{\text{train}}$, $\mathbf{A}(\mu)$, $\mathbf{b}(\mu)$, $\mathbf{l}(\mu)$, $\mathbf{\Theta}$, $\tau$, $r$.
**Output**: $\mathbf{\Theta}$-sketch of $\mathbf{U}_K$.
1. Set $i = 0$, $\mathcal{D}_0 = \emptyset$, obtain $\eta(\mu)$ in (3.45), set $\Delta^{\mathbf{\Phi}}(\mu) = \|\mathbf{b}(\mu)\|^{\mathbf{\Phi}}_{U'}/\eta(\mu)$
   and pick $\mu^1 \in \mathcal{P}_{\text{train}}$.
**while** $\max_{\mu \in \mathcal{P}\text{train}} \Delta^{\mathbf{\Phi}}(\mathbf{u}_r(\mu);\mu) > \tau$ **do**
    2. Set $i = i+1$.
    3. Evaluate $\mathbf{u}(\mu^i)$ and set $\mathcal{D}_i := \mathcal{D}_{i-1} \cup \{\mathbf{u}(\mu^i)\}$
    4. Update the $\mathbf{\Theta}$-sketch of $\mathbf{U}_i$ (matrix composed from the vectors in $\mathcal{D}_i$).
    5. Use Algorithm 5 (if $i < r$, choosing $r := i$) with $\mathcal{P}_{\text{test}}$ replaced by $\mathcal{P}_{\text{train}}$ to
       solve (3.43) for all $\mu \in \mathcal{P}_{\text{train}}$.
    6. Find $\mu^{i+1} := \arg\max_{\mu \in \mathcal{P}\text{train}} \Delta^{\mathbf{\Phi}}(\mathbf{u}_r(\mu);\mu)$.
**end while**
```

By Proposition 3.4.5, a good quality of a $\mathbf{\Theta}$-sketch for the sketched sparse minres approximation associated with dictionary $\mathcal{D}_K$ on $\mathcal{P}_{\text{train}}$ can be guaranteed if $\mathbf{\Theta}$ is an ε-embedding for every subspace $R_r(W_r;\mu)$, defined by (3.18), with $W_r \in \mathcal{L}_r(\mathcal{D}_K)$ and $\mu \in \mathcal{P}_{\text{train}}$. This condition can be enforced a priori for all possible outcomes of Algorithm 6 by choosing $\mathbf{\Theta}$ such that it is an ε-embedding for every subspace $R_r(W_r;\mu)$ with $W_r \in \mathcal{L}_r(\{\mathbf{u}(\mu) : \mu \in \mathcal{P}_{\text{train}}\})$ and $\mu \in \mathcal{P}_{\text{train}}$. An embedding $\mathbf{\Theta}$ satisfying this property with probability at least $1-\delta$ can be obtained as a realization of a $(\varepsilon, (\#P_{\text{train}})^{-1}\binom{\#\mathcal{P}_{\text{train}}}{r}^{-1}\delta, r+1)$ oblivious $U \to \ell_2$ subspace embedding. The computational cost of Algorithm 6 is dominated by the calculation of the snapshots $\mathbf{u}(\mu^i)$ and their $\mathbf{\Theta}$-sketches. As was argued in Chapter 2, the computation of the snapshots can have only a minor impact on the overall cost of an algorithm. For the classical sequential or limited-memory computational architectures, each snapshot should require a log-linear complexity and memory consumption, while for parallel and distributed computing the routines for computing the snapshots should be well-parallelizable and require low communication between cores. Moreover, for the computation of the snapshots one may use a highly-optimized commercial solver or a powerful server. The $\mathbf{\Theta}$-sketch of the snapshots may be computed extremely efficiently in basically any computational architecture Section 2.4.4. With P-SRHT, sketching of K snapshots requires only $\mathcal{O}(n(Km_A + m_b)\log k)$ flops, and the maintenance of the sketch requires $\mathcal{O}((Km_A + m_b)k)$ bytes of memory. By using similar arguments as in Section 3.4.3 it can be shown that $k = \mathcal{O}(r \log K)$ (or $k = \mathcal{O}(r^* \log K)$) is enough to yield with high probability an accurate approximation of the dictionary-based reduced model. With this value of k, the required number of flops for the computation and the amount of memory for the storage of a $\mathbf{\Theta}$-sketch becomes $\mathcal{O}(n(Km_A+m_b)(\log r+\log\log K))$, and $\mathcal{O}((Km_A+m_b)r\log K)$, respectively.

3.5 Post-processing the reduced model's solution

So far we presented a methodology for efficient computation of an approximate solution $\mathbf{u}_r(\mu)$, or to be more precise, its coordinates in a certain basis, which can be the classical reduced basis for a fixed approximation space, or the dictionary vectors for dictionary-based approximation presented in Section 3.4. The approximate solution $\mathbf{u}_r(\mu)$, however, is usually not what one should consider as the output. In fact, the amount of allowed online computations is highly limited and should be independent of the dimension of the full order model. Therefore outputting $\mathcal{O}(n)$ bytes of data as $\mathbf{u}_r(\mu)$ should be avoided when $\mathbf{u}(\mu)$ is not the quantity of interest.

Further, we shall consider an approximation with a single subspace U_r noting that the presented approach can also be used for post-processing the dictionary-based approximation from Section 3.4 (by taking U_r as the subspace spanned by the dictionary vectors). Let $\mathbf{U}_r$ be a matrix whose column vectors form a basis for U_r and let $\mathbf{a}_r(\mu)$ be the coordinates of $\mathbf{u}_r(\mu)$ in this basis. A general quantity of interest $s(\mu) := l(\mathbf{u}(\mu);\mu)$ can be approximated by $s_r(\mu) := l(\mathbf{u}_r(\mu);\mu)$. Further, let us assume a linear case where $l(\mathbf{u}(\mu);\mu) := \langle \mathbf{l}(\mu), \mathbf{u}(\mu)\rangle$ with $\mathbf{l}(\mu) \in U'$ being the extractor of the quantity of interest. Then

$$s_r(\mu) = \langle \mathbf{l}(\mu), \mathbf{u}_r(\mu)\rangle = \mathbf{l}_r(\mu)^{\mathrm{H}}\mathbf{a}_r(\mu), \tag{3.47}$$

where $\mathbf{l}_r(\mu) := \mathbf{U}_r^{\mathrm{H}}\mathbf{l}(\mu)$.

Remark 3.5.1. *In general, our approach can be used for estimating an inner product between arbitrary parameter-dependent vectors. The possible applications include efficient estimation of the primal-dual correction and an extension to quadratic quantities of interest. In particular, the estimation of the primal-dual correction can be obtained by replacing $\mathbf{l}(\mu)$ by $\mathbf{r}(\mathbf{u}_r(\mu);\mu)$ and $\mathbf{u}_r(\mu)$ by $\mathbf{v}_r(\mu)$ in (3.47), where $\mathbf{v}_r(\mu) \in U$ is a reduced basis (or dictionary-based) approximate solution to the adjoint problem. A quadratic output quantity of interest has the form $l(\mathbf{u}_r(\mu);\mu) := \langle \mathbf{L}(\mu)\mathbf{u}_r(\mu) + \mathbf{l}(\mu), \mathbf{u}_r(\mu)\rangle$, where $\mathbf{L}(\mu) : U \to U'$ and $\mathbf{l}(\mu) \in U'$. Such $l(\mathbf{u}_r(\mu);\mu)$ can be readily derived from (3.47) by replacing $\mathbf{l}(\mu)$ with $\mathbf{L}(\mu)\mathbf{u}_r(\mu) + \mathbf{l}(\mu)$.*

The affine factors of $\mathbf{l}_r(\mu)$ should be first precomputed in the offline stage and then used for online evaluation of $\mathbf{l}_r(\mu)$ for each parameter value with a computational cost independent of the dimension of the original problem. The offline computations required for evaluating the affine factors of $\mathbf{l}_r(\mu)$, however, can still be too expensive or even unfeasible to perform. Such a scenario may arise when using a high-dimensional approximation space (or a dictionary), when the extractor $\mathbf{l}(\mu)$ has many (possibly expensive to maintain) affine terms, or when working in an extreme computational environment, e.g., with data streamed or distributed among multiple workstations. In addition, evaluating $\mathbf{l}_r(\mu)$ from the affine expansion as well as evaluating $\mathbf{l}_r(\mu)^{\mathrm{H}}\mathbf{a}_r(\mu)$ itself can be subject to round-off errors (especially when $\mathbf{U}_r$ is ill-conditioned and may not be orthogonalized). Further, we shall provide a (probabilistic) way for estimating $s_r(\mu)$ with a reduced computational cost and better numerical stability. As the core we take the idea from Section 2.4.3 proposed as a workaround to expensive offline computations for the evaluation of the primal-dual correction.

Remark 3.5.2. *The spaces U and U' are equipped with inner products $\langle\cdot,\cdot\rangle_U$ and $\langle\cdot,\cdot\rangle_{U'}$ (defined by matrix $\mathbf{R}_U$), which are used for controlling the accuracy of the approximate solution $\mathbf{u}_r(\mu)$. In general, $\mathbf{R}_U$ is chosen according to both the operator $\mathbf{A}(\mu)$ and the extractor $\mathbf{l}(\mu)$ of the quantity of interest. The goal of this section, however, is only the estimation of the quantity of interest from the given $\mathbf{u}_r(\mu)$. Consequently, for many problems it can be more pertinent to use here a different $\mathbf{R}_U$ than the one employed for obtaining and characterizing $\mathbf{u}_r(\mu)$. The choice for $\mathbf{R}_U$ should be done according to $\mathbf{l}(\mu)$ (independently of $\mathbf{A}(\mu)$). For instance, for discretized parametric PDEs, if $\mathbf{l}(\mu)$ represents an integral of the solution field over the spatial domain then it is natural to choose $\langle\cdot,\cdot\rangle_U$ corresponding to the L_2 inner product. Moreover, $\langle\cdot,\cdot\rangle_U$ is required to be an inner product only on a certain subspace of interest, which means that $\mathbf{R}_U$ may be a positive semi-definite matrix. This consideration can be particularly helpful when the quantity of interest depends only on the restriction of the solution field to a certain subdomain. In such a case, $\langle\cdot,\cdot\rangle_U$ can be chosen to correspond with an inner product between restrictions of functions to this subdomain. The extension of random sketching for estimation of semi-inner products is straightforward (see Remark 3.2.1).*

3.5.1 Approximation of the quantity of interest

An efficiently computable and accurate estimation of $s_r(\mu)$ can be obtained in two phases. In the first phase, the manifold $\mathcal{M}_r := \{\mathbf{u}_r(\mu) : \mu \in \mathcal{P}\}$ is (accurately enough) approximated with a subspace $W_p := \mathrm{span}(\mathbf{W}_p) \subset U$, which is spanned by an efficient to multiply (i.e., sparse or low-dimensional) matrix $\mathbf{W}_p$. This matrix can be selected a priori or obtained depending on $\mathcal{M}_r$. In Section 3.5.2 we shall provide some strategies for choosing or computing the columns for $\mathbf{W}_p$. The appropriate strategy should be selected depending on the particular problem and computational architecture. Further, the solution vector $\mathbf{u}_r(\mu)$ is approximated by its orthogonal projection $\mathbf{w}_p(\mu) := \mathbf{W}_p\mathbf{c}_p(\mu)$ on W_p. The coordinates $\mathbf{c}_p(\mu)$ can be obtained from $\mathbf{a}_r(\mu)$ by

$$\mathbf{c}_p(\mu) = \mathbf{H}_p\mathbf{a}_r(\mu), \tag{3.48}$$

where $\mathbf{H}_p := [\mathbf{W}_p^{\mathrm{H}}\mathbf{R}_U\mathbf{W}_p]^{-1}\mathbf{W}_p^{\mathrm{H}}\mathbf{R}_U\mathbf{U}_r$. Note that since $\mathbf{W}_p$ is efficient to multiply by, the matrix $\mathbf{H}_p$ can be efficiently precomputed in the offline stage. We arrive at the following estimation of $s_r(\mu)$:

$$s_r(\mu) \approx \langle \mathbf{l}(\mu), \mathbf{w}_p(\mu) \rangle = \mathbf{l}_r^\star(\mu)^{\mathrm{H}} \mathbf{a}_r(\mu), \tag{3.49}$$

where $\mathbf{l}_r^\star(\mu)^{\mathrm{H}} := \mathbf{l}(\mu)^{\mathrm{H}} \mathbf{W}_p \mathbf{H}_p$. Unlike $\mathbf{l}_r(\mu)$, the affine factors of $\mathbf{l}_r^\star(\mu)$ can now be efficiently precomputed thanks to the structure of $\mathbf{W}_p$.

In the second phase of the algorithm, the precision of (3.49) is improved with a sketched (random) correction associated with an $U \to \ell_2$ subspace embedding $\mathbf{\Theta}$:

$$\begin{aligned} s_r(\mu) &= \langle \mathbf{l}(\mu), \mathbf{w}_p(\mu) \rangle + \langle \mathbf{l}(\mu), \mathbf{u}_r(\mu) - \mathbf{w}_p(\mu) \rangle \\ &\approx \langle \mathbf{l}(\mu), \mathbf{w}_p(\mu) \rangle + \langle \mathbf{R}_U^{-1} \mathbf{l}(\mu), \mathbf{u}_r(\mu) - \mathbf{w}_p(\mu) \rangle_U^{\mathbf{\Theta}} =: s_r^\star(\mu). \end{aligned} \tag{3.50}$$

In practice, $s_r^\star(\mu)$ can be efficiently evaluated using the following relation:

$$s_r^\star(\mu) = [\mathbf{l}_K^\star(\mu)^{\mathrm{H}} + {}_\Delta\mathbf{l}_K^\star(\mu)^{\mathrm{H}}] \mathbf{a}_r(\mu), \tag{3.51}$$

where the affine terms of ${}_\Delta\mathbf{l}_r^\star(\mu)^{\mathrm{H}} := \mathbf{l}^{\mathbf{\Theta}}(\mu)^{\mathrm{H}} (\mathbf{U}_r^{\mathbf{\Theta}} - \mathbf{W}_p^{\mathbf{\Theta}} \mathbf{H}_p)$ can be precomputed from the $\mathbf{\Theta}$-sketch of $\mathbf{U}_r$, a sketched matrix $\mathbf{W}_p^{\mathbf{\Theta}} := \mathbf{\Theta} \mathbf{W}_p$ and the matrix $\mathbf{H}_p$ with a negligible computational cost.

Proposition 3.5.3. *If $\mathbf{\Theta}$ is an $(\varepsilon, \delta, 1)$ oblivious $U \to \ell_2$ subspace embedding,*

$$|s_r(\mu) - s_r^\star(\mu)| \leq \varepsilon \|\mathbf{l}(\mu)\|_{U'} \|\mathbf{u}_r(\mu) - \mathbf{w}_p(\mu)\|_U \tag{3.52}$$

holds for a fixed parameter $\mu \in \mathcal{P}$ with probability at least $1 - 2\delta$.

Proof. See appendix. □

Proposition 3.5.4. *Let $L \subset U$ denote a subspace containing $\{\mathbf{R}_U^{-1} \mathbf{l}(\mu) : \mu \in \mathcal{P}\}$. Let*

$$\mathcal{Y} := \{Y_r + L + W_p : Y_r \in \mathcal{L}_r(\mathcal{D}_K)\}.$$

If $\mathbf{\Theta}$ is an ε-embedding for every subspace in $\mathcal{Y}$, then (3.52) holds for all $\mu \in \mathcal{P}$.

Proof. See appendix. □

It follows that the accuracy of $s_r^\star(\mu)$ can be controlled through the quality of W_p for approximating $\mathcal{M}_r$, the quality of $\mathbf{\Theta}$ as an $U \to \ell_2$ ε-embedding, or both. Note that choosing $\mathbf{\Theta}$ as a null matrix (i.e., an ε-embedding for U with $\varepsilon = 1$) leads to a single first-phase approximation (3.49), while letting $W_p := \{\mathbf{0}\}$ corresponds to a single sketched (second-phase) approximation. Such particular choices for $\mathbf{\Theta}$ or W_p can be pertinent when the subspace W_p is highly accurate so that there is practically no benefit to use a sketched correction or, the other way around, when the computational environment or the problem does not permit a sufficiently accurate approximation of $\mathcal{M}_r$ with W_p, therefore making the use of a non-zero $\mathbf{w}_p(\mu)$ unjustified.

Remark 3.5.5. *When interpreting random sketching as a Monte Carlo method for the estimation of the inner product $\langle \mathbf{l}(\mu), \mathbf{u}_r(\mu) \rangle$, the proposed approach can be interpreted as a control variate method where $\mathbf{w}_p(\mu)$ plays the role of the control variate. A multileveled Monte Carlo method with different control variates should further improve the efficiency of post-processing.*

3.5.2 Construction of reduced subspaces

Further we address the problem of computing the basis vectors for W_p. In general, the strategy for obtaining W_p has to be chosen according to the problem's structure and the constraints due to the computational environment.

A simple way, used in Chapter 2, is to choose W_p as the span of samples of $\mathbf{u}(\mu)$ either chosen randomly or during the first few iterations of the reduced basis (or dictionary) generation with a greedy algorithm. Such W_p, however, may be too costly to operate with. Then we propose more sophisticated constructions of W_p.

Approximate Proper Orthogonal Decomposition

A subspace W_p can be obtained by an (approximate) POD of the reduced vectors evaluated on a training set $\mathcal{P}_{\text{train}} \subseteq \mathcal{P}$. Here, randomized linear algebra can be again employed for improving efficiency. The computational cost of the proposed POD procedure shall mainly consist of the solution of $m = \#\mathcal{P}_{\text{train}}$ reduced problems and the multiplication of $\mathbf{U}_r$ by $p = \dim(W_p) \ll r$ small vectors. Unlike the classical POD, our methodology does not require computation or maintenance of the full solution's samples and therefore allows large training sets.

Let $L_m = \{\mathbf{a}_r(\mu^i)\}_{i=1}^m$ be a training sample of the coordinates of $\mathbf{u}_r(\mu)$ in a basis $\mathbf{U}_r$. We look for a POD subspace W_r associated with the snapshot matrix

$$\mathbf{W}_m := [\mathbf{u}_r(\mu^1), \mathbf{u}_r(\mu^2), \ldots, \mathbf{u}_r(\mu^m)] = \mathbf{U}_r \mathbf{L}_m,$$

where $\mathbf{L}_m$ is a matrix whose columns are the elements from L_m.

An accurate estimation of POD can be efficiently computed via the sketched method of snapshots introduced in Section 2.5.2. More specifically, a quasi-optimal (with high probability) POD basis can be calculated as

$$\mathbf{W}_p := \mathbf{U}_r \mathbf{T}_p^*, \tag{3.53}$$

where

$$\mathbf{T}_p^* := \mathbf{L}_m[\mathbf{t}_1, \ldots, \mathbf{t}_p],$$

with $\mathbf{t}_1, \ldots, \mathbf{t}_p$ being the p dominant singular vectors of $\mathbf{U}_r^{\mathbf{\Theta}} \mathbf{L}_m$. Note that the matrix $\mathbf{T}_p^*$ can be efficiently obtained with a computational cost independent of the dimension of the full order model. The dominant cost is the multiplication of $\mathbf{U}_r$ by $\mathbf{T}_p^*$, which is also expected to be inexpensive since $\mathbf{T}_p^*$ has a small number of columns. Guarantees for the quasi-optimality of W_p can be readily derived from Theorem 2.5.5.

Sketched greedy algorithm

A greedy search over the training set $\{\mathbf{u}_r(\mu) : \mu \in \mathcal{P}_{\text{train}}\}$ of approximate solutions is another way to construct W_p. At the ith iteration, $\mathbf{W}_i$ is enriched with a vector

$\mathbf{u}_r(\mu^{i+1})$ with the largest distance to W_i over the training set. Note that in this case the resulting matrix $\mathbf{W}_p$ has the form (3.53), where $\mathbf{T}^*_p = [\mathbf{a}_r(\mu^1), \ldots, \mathbf{a}_r(\mu^p)]$. The efficiency of the greedy selection can be improved by employing random sketching technique. At each iteration, the distance to W_i can be measured with the sketched norm $\| \cdot \|^{\mathbf{\Theta}}_U$, which can be computed from sketches $\mathbf{\Theta}\mathbf{u}_r(\mu) = \mathbf{U}^{\mathbf{\Theta}}_r \mathbf{a}_r(\mu)$ of the approximate solutions with no need to operate with large matrix $\mathbf{U}_r$ but only its sketch. This allows efficient computation of the quasi-optimal interpolation points $\mu^1, \ldots, \mu^p$ and the associated matrix $\mathbf{T}^*_p$. Note that for numerical stability an orthogonalization of $\mathbf{W}_i$ with respect to $\langle \cdot, \cdot \rangle^{\mathbf{\Theta}}_U$ can be performed, that can be done by modifying $\mathbf{T}^*_i$ so that $\mathbf{U}^{\mathbf{\Theta}}_r \mathbf{T}^*_i$ is an orthogonal matrix. Such $\mathbf{T}^*_i$ can be obtained with standard QR factorization. When $\mathbf{T}^*_p$ has been computed, the matrix $\mathbf{W}_p$ can be calculated by multiplying $\mathbf{U}_r$ with $\mathbf{T}^*_p$. The quasi-optimality of $\mu^1, \ldots, \mu^p$ and approximate orthogonality of $\mathbf{W}_p$ is guaranteed if $\mathbf{\Theta}$ is an ε-embedding for all subspaces from the set $\{W_p + \mathrm{span}(\mathbf{u}_r(\mu^i))\}^m_{i=1}$. This property of $\mathbf{\Theta}$ can be guaranteed a priori by considering $(\varepsilon, \binom{m}{p}^{-1}\delta, p+1)$ oblivious $U \to \ell_2$ subspace embeddings, or certified a posteriori with the procedure explained in Section 3.6.

Coarse approximation

Let us notice that the online cost of evaluating $s^\star_r(\mu)$ does not depend on the dimension p of W_p. Consequently, if W_p is spanned by structured (e.g., sparse) basis vectors then a rather high dimension is allowed (possibly larger than r).

For classical numerical methods for PDEs, the resolution of the mesh (or grid) is usually chosen to guarantee both an approximability of the solution manifold by the approximation space and the stability. For many problems the latter factor is dominant and one choses the mesh primary to it. This is a typical situation for wave problems, advection-diffusion-reaction problems and many others. For these problems, the resolution of the mesh can be much higher than needed for the estimation of the quantity of interest from the given solution field. In these cases, the quantity of interest can be efficiently yet accurately approximated using a coarse-grid interpolation of the solution.

Suppose that each vector $\mathbf{u} \in U$ represents a function $u : \Omega \to \mathbb{K}$ in a finite-dimensional approximation space spanned by basis functions $\{\psi_i(x)\}^n_{i=1}$ associated with a fine mesh of Ω. The function $u(x)$ can be approximated by a projection on a coarse-grid approximation space spanned by basis functions $\{\phi_i(x)\}^p_{i=1}$. For simplicity assume that each $\phi_i(x) \in \mathrm{span}\{\psi_j(x)\}^n_{j=1}$. Then the ith basis vector for W_p can be obtained simply by evaluating the coordinates of $\phi_i(x)$ in the basis $\{\psi_j(x)\}^n_{j=1}$. Note that for the classical finite element approximation, each basis function has a local support and the resulting matrix $\mathbf{W}_p$ is sparse.

3.6 A posteriori certification of the sketch and solution

Here we provide a simple, yet efficient procedure for a posteriori verification of the quality of a sketching matrix and describe a few scenarios where such a procedure can be employed. The proposed a posteriori certification of the sketched reduced model and its solution is probabilistic. It does not require operating with high-dimensional vectors but only with their small sketches. The quality of a certificate shall be characterized by two user specified parameters: $0 < \delta^* < 1$ for the probability of success and $0 < \varepsilon^* < 1$ for the tightness of the computed error bounds.

3.6.1 Verification of an ε-embedding for a given subspace

Let $\mathbf{\Theta}$ be a $U \to \ell_2$ subspace embedding and $V \subset U$ be a subspace of U (chosen depending on the reduced model, e.g., $V := R_r(U_r;\mu)$ in (3.18)). Recall that the quality of $\mathbf{\Theta}$ can be measured by the accuracy of $\langle \cdot, \cdot \rangle_U^{\mathbf{\Theta}}$ as an approximation of $\langle \cdot, \cdot \rangle_U$ for vectors in V.

We propose to verify the accuracy of $\langle \cdot, \cdot \rangle_U^{\mathbf{\Theta}}$ simply by comparing it to an inner product $\langle \cdot, \cdot \rangle_U^{\mathbf{\Theta}^*}$ associated with a new random embedding $\mathbf{\Theta}^* \in \mathbb{K}^{k^* \times n}$, where $\mathbf{\Theta}^*$ is chosen such that the concentration inequality

$$\mathbb{P}\left(\left| \|\mathbf{x}\|_U^2 - (\|\mathbf{x}\|_U^{\mathbf{\Theta}^*})^2 \right| \leq \varepsilon^* \|\mathbf{x}\|_U^2 \right) \geq 1 - \delta^* \tag{3.54}$$

is satisfied for all vectors $\mathbf{x} \in V$. One way to ensure (3.54) is to choose $\mathbf{\Theta}^*$ as an $(\varepsilon^*, \delta^*, 1)$ oblivious $U \to \ell_2$ subspace embedding. A condition on the number of rows for the oblivious embedding can be either obtained theoretically (see Chapter 2) or chosen from the practical experience, which should be the case for embeddings constructed with P-SRHT matrices (recall that they have worse theoretical guarantees than the Gaussian matrices but perform equally well in practice). Alternatively, $\mathbf{\Theta}^*$ can be built by random sampling of the rows of a larger ε-embedding for V. This approach can be far more efficient than generating $\mathbf{\Theta}^*$ as an oblivious embedding (see Remark 3.6.1) or even essential for some computational architectures. Another requirement for $\mathbf{\Theta}^*$ is that it is generated independently from $\mathbf{\Theta}$. Therefore, in the algorithms we suggest to consider $\mathbf{\Theta}^*$ only for the certification of the solution and nothing else.

Remark 3.6.1. *In some scenarios it can be beneficial to construct $\mathbf{\Theta}$ and $\mathbf{\Theta}^*$ by sampling their rows from a fixed realization of a larger oblivious embedding $\hat{\mathbf{\Theta}}$, which is guaranteed a priori to be an ε-embedding for V with high probability. More precisely, $\mathbf{\Theta}$ and $\mathbf{\Theta}^*$ can be defined as*

$$\mathbf{\Theta} := \mathbf{\Gamma}\hat{\mathbf{\Theta}}, \quad \mathbf{\Theta}^* := \mathbf{\Gamma}^*\hat{\mathbf{\Theta}}, \tag{3.55}$$

where $\mathbf{\Gamma}$ and $\mathbf{\Gamma}^*$ are random independent sampling (or Gaussian, or P-SRHT) matrices. In this way, a $\hat{\mathbf{\Theta}}$-sketch of a reduced order model can be first precomputed and then used for efficient evaluation/update of the sketches associated with $\mathbf{\Theta}$ and $\mathbf{\Theta}^*$. This approach can be essential for the adaptive selection of the optimal size for $\mathbf{\Theta}$ in a limited-memory environment where only one pass (or a few passes) over the reduced basis vectors is allowed and therefore there is no chance to recompute a sketch associated with an oblivious embedding at each iteration. It may also reduce the complexity (number of flops) of an algorithm (especially when $\mathbf{\Theta}$ is constructed with P-SRHT matrices) by not requiring to recompute high-dimensional matrix-vector products multiple times.

Let $\mathbf{V}$ denote a matrix whose columns form a basis of V. Define the sketches $\mathbf{V}^{\mathbf{\Theta}} := \mathbf{\Theta}\mathbf{V}$ and $\mathbf{V}^{\mathbf{\Theta}^*} := \mathbf{\Theta}^*\mathbf{V}$. Note that $\mathbf{V}^{\mathbf{\Theta}}$ and $\mathbf{V}^{\mathbf{\Theta}^*}$ contain as columns low-dimensional vectors and therefore are cheap to maintain and to operate with (unlike the matrix $\mathbf{V}$).

We start with the certification of the inner product between two fixed vectors from V (see Proposition 3.6.2).

Proposition 3.6.2. *For any two vectors $\mathbf{x}, \mathbf{y} \in V$, we have that*

$$\begin{aligned} |\langle \mathbf{x}, \mathbf{y}\rangle_U^{\mathbf{\Theta}^*} - \langle \mathbf{x}, \mathbf{y}\rangle_U^{\mathbf{\Theta}}| - \frac{\varepsilon^*}{1-\varepsilon^*}\|\mathbf{x}\|_U^{\mathbf{\Theta}^*}\|\mathbf{y}\|_U^{\mathbf{\Theta}^*} &\leq |\langle \mathbf{x}, \mathbf{y}\rangle_U - \langle \mathbf{x}, \mathbf{y}\rangle_U^{\mathbf{\Theta}}| \\ &\leq |\langle \mathbf{x}, \mathbf{y}\rangle_U^{\mathbf{\Theta}^*} - \langle \mathbf{x}, \mathbf{y}\rangle_U^{\mathbf{\Theta}}| + \frac{\varepsilon^*}{1-\varepsilon^*}\|\mathbf{x}\|_U^{\mathbf{\Theta}^*}\|\mathbf{y}\|_U^{\mathbf{\Theta}^*} \end{aligned} \tag{3.56}$$

holds with probability at least $1-4\delta^*$.

Proof. See appendix. □

The error bounds in Proposition 3.6.2 can be computed from the sketches of $\mathbf{x}$ and $\mathbf{y}$, which may be efficiently evaluated from $\mathbf{V}^{\mathbf{\Theta}}$ and $\mathbf{V}^{\mathbf{\Theta}^*}$ and the coordinates of $\mathbf{x}$ and $\mathbf{y}$ associated with $\mathbf{V}$, with no operations on high-dimensional vectors. A certification for several pairs of vectors should be obtained using a union bound for the probability of success. By replacing $\mathbf{x}$ by $\mathbf{R}_U^{-1}\mathbf{x}'$ and $\mathbf{y}$ by $\mathbf{R}_U^{-1}\mathbf{y}'$ in Proposition 3.6.2 and using definition (3.2) one can derive a certification of the dual inner product $\langle \cdot, \cdot \rangle_{U'}^{\mathbf{\Theta}}$ for vectors $\mathbf{x}'$ and $\mathbf{y}'$ in $V' := \{\mathbf{R}_U\mathbf{x} : \mathbf{x} \in V\}$.

In general, the quality of an approximation with a $\mathbf{\Theta}$-sketch of a reduced model should be characterized by the accuracy of $\langle \cdot, \cdot \rangle_U^{\mathbf{\Theta}}$ for the whole subspace V. Let ω be the minimal value for ε such that $\mathbf{\Theta}$ satisfies an ε-embedding property for V. Now, we address the problem of computing an a posteriori upper bound $\bar{\omega}$ for ω from the sketches $\mathbf{V}^{\mathbf{\Theta}}$ and $\mathbf{V}^{\mathbf{\Theta}^*}$ and we provide conditions to ensure quasi-optimality of $\bar{\omega}$.

Proposition 3.6.3. *For a fixed realization of* $\mathbf{\Theta}^*$*, let us define*

$$\bar{\omega} := \max\left\{1-(1-\varepsilon^*)\min_{\mathbf{x}\in V/\{\mathbf{0}\}}\left(\frac{\|\mathbf{x}\|_U^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U^{\mathbf{\Theta}^*}}\right)^2, (1+\varepsilon^*)\max_{\mathbf{x}\in V/\{\mathbf{0}\}}\left(\frac{\|\mathbf{x}\|_U^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U^{\mathbf{\Theta}^*}}\right)^2-1\right\}. \tag{3.57}$$

If $\bar{\omega}<1$*, then* $\mathbf{\Theta}$ *is guaranteed to be a* $U\to\ell_2$ $\bar{\omega}$*-subspace embedding for* V *with probability at least* $1-\delta^*$*.*

Proof. See appendix. □

It follows that if $\bar{\omega}<1$ then it is an upper bound for ω with high probability. Assume that $\mathbf{V}^{\mathbf{\Theta}}$ and $\mathbf{V}^{\mathbf{\Theta}^*}$ have full ranks. Let $\mathbf{T}^*$ be the matrix such that $\mathbf{V}^{\mathbf{\Theta}^*}\mathbf{T}^*$ is orthogonal (with respect to ℓ_2-inner product). Such a matrix can be computed with a QR factorization. Then $\bar{\omega}$ defined in (3.57) can be obtained from the following relation

$$\bar{\omega} = \max\left\{1-(1-\varepsilon^*)\sigma_{\min}^2, (1+\varepsilon^*)\sigma_{\max}^2-1\right\}, \tag{3.58}$$

where $\sigma_{\min}$ and $\sigma_{\max}$ are the minimal and the maximal singular values of the small matrix $\mathbf{V}^{\mathbf{\Theta}}\mathbf{T}^*$.

We have that $\bar{\omega}\geq\varepsilon^*$. The value for ε^* may be selected an order of magnitude less than ω with no considerable impact on the computational cost, therefore in practice the effect of ε^* on $\bar{\omega}$ can be considered to be negligible. Proposition 3.6.3 implies that $\bar{\omega}$ is an upper bound of ω with high probability. A guarantee of effectivity of $\bar{\omega}$ (i.e., its closeness to ω), however, has not been yet provided. To do so we shall need a stronger assumption on $\mathbf{\Theta}^*$ than (3.54).

Proposition 3.6.4. *If the realization of* $\mathbf{\Theta}^*$ *is a* $U\to\ell_2$ ω^**-subspace embedding for* V*, then* $\bar{\omega}$ *(defined by* (3.57)*) satisfies*

$$\bar{\omega} \leq \frac{1+\varepsilon^*}{1-\omega^*}(1+\omega)-1. \tag{3.59}$$

Proof. See appendix. □

If $\mathbf{\Theta}^*$ is a $(\omega^*,\gamma^*,\dim(V))$ oblivious $U\to\ell_2$ subspace embedding, then the condition on $\mathbf{\Theta}^*$ in Proposition 3.6.4 is satisfied with probability at least $1-\gamma^*$ (for some user-specified value γ^*). Therefore, a matrix $\mathbf{\Theta}^*$ of moderate size should yield a very good upper bound $\bar{\omega}$ of ω. Moreover, if $\mathbf{\Theta}$ and $\mathbf{\Theta}^*$ are drawn from the same distribution, then $\mathbf{\Theta}^*$ can be expected to be an ω^*-embedding for V with $\omega^*=\mathcal{O}(\omega)$ with high probability. Combining this consideration with Proposition 3.6.4 we deduce that a sharp upper bound should be obtained for some $k^*\leq k$. Therefore, in the algorithms one may readily consider $k^*:=k$. If pertinent, a better value for k^* can be selected adaptively, at each iteration increasing k^* by a constant factor until the desired tolerance or a stagnation of $\bar{\omega}$ is reached.

3.6.2 Certification of a sketch of a reduced model and its solution

The results of Propositions 3.6.2 and 3.6.3 can be employed for certification of a sketch of a reduced model and its solution. They can also be used for adaptive selection of the number of rows of a random sketching matrix to yield an accurate approximation of the reduced model. Thereafter we discuss several practical applications of the methodology described above.

Approximate solution

Let $\mathbf{u}_r(\mu) \in U$ be an approximation of $\mathbf{u}(\mu)$. The accuracy of $\mathbf{u}_r(\mu)$ can be measured with the residual error $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}$, which can be efficiently estimated by

$$\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'} \approx \|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}}_{U'}.$$

The certification of such estimation can be derived from Proposition 3.6.2 choosing $\mathbf{x} = \mathbf{y} := \mathbf{R}_U^{-1}\mathbf{r}(\mathbf{u}_r(\mu);\mu)$ and using definition (3.2) of $\|\cdot\|^{\mathbf{\Theta}}_{U'}$.

For applications, which involve computation of snapshots over the training set (e.g., approximate POD or greedy algorithm with the exact error indicator), one should be able to efficiently precompute the sketches of $\mathbf{u}(\mu)$. Then the error $\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U$ can be efficiently estimated by

$$\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \approx \|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|^{\mathbf{\Theta}}_U.$$

Such an estimation can be certified with Proposition 3.6.2 choosing $\mathbf{x} = \mathbf{y} := \mathbf{u}(\mu) - \mathbf{u}_r(\mu)$.

Output quantity

In Section 3.5 we provided a way for estimating the output quantity $s_r(\mu) = \langle \mathbf{l}(\mu), \mathbf{u}_r(\mu)\rangle$. More specifically $s_r(\mu)$ can be efficiently estimated by $s_r^\star(\mu)$ defined in (3.51). We have

$$|s_r(\mu) - s_r^\star(\mu)| = |\langle \mathbf{R}_U^{-1}\mathbf{l}(\mu), \mathbf{u}_r(\mu) - \mathbf{w}_p(\mu)\rangle_U - \langle \mathbf{R}_U^{-1}\mathbf{l}(\mu), \mathbf{u}_r(\mu) - \mathbf{w}_p(\mu)\rangle^{\mathbf{\Theta}}_U|,$$

therefore the quality of $s_r^\star(\mu)$ may be certified by Proposition 3.6.2 with $\mathbf{x} = \mathbf{R}_U^{-1}\mathbf{l}(\mu)$, $\mathbf{y} = \mathbf{u}_r(\mu) - \mathbf{w}_p(\mu)$.

Minimal residual projection

By Proposition 3.3.3, the quality of the $\mathbf{\Theta}$-sketch of a subspace U_r for approximating the minres projection for a given parameter value can be characterized by the lowest value ω for ε such that $\mathbf{\Theta}$ satisfies the ε-embedding property for subspace

$V := R_r(U_r; \mu)$, defined in (3.18). The upper bound $\bar{\omega}$ of such ω can be efficiently computed using (3.57). The verification of the $\mathbf{\Theta}$-sketch for all parameter values in $\mathcal{P}$, simultaneously, can be performed by considering a subspace V in (3.57), which contains $\bigcup_{\mu \in \mathcal{P}} R_r(U_r; \mu)$.

Dictionary-based approximation

For each parameter value, the quality of $\mathbf{\Theta}$ for dictionary-based approximation defined in (3.43) can be characterized by the quality of the sketched minres projection associated with a subspace $U_r(\mu)$, which can be verified by computing $\bar{\omega}$ in (3.57) associated with $V := R_r(U_r(\mu); \mu)$.

Adaptive selection of the size for a random sketching matrix

When no a priori bound for the size of $\mathbf{\Theta}$ sufficient to yield an accurate sketch of a reduced model is available, or when the bounds are pessimistic, the sketching matrix should be selected adaptively. At each iteration, if the certificate indicates a poor quality of a $\mathbf{\Theta}$-sketch for approximating the solution (or the error) on $\mathcal{P}_{\text{train}} \subseteq \mathcal{P}$, one can improve the accuracy of the sketch by adding extra rows to $\mathbf{\Theta}$. In the analysis, the embedding $\mathbf{\Theta}^*$ used for certification was assumed to be independent of $\mathbf{\Theta}$, consequently a new realization of $\mathbf{\Theta}^*$ should be sampled after each decision to improve $\mathbf{\Theta}$ has been made. To save computational costs, the previous realizations of $\mathbf{\Theta}^*$ and the associated $\mathbf{\Theta}^*$-sketches can be readily reused as parts of the updates for $\mathbf{\Theta}$ and the $\mathbf{\Theta}$-sketch.

We finish with a practical application of Propositions 3.6.3 and 3.6.4. Consider a situation where one is given a class of random embeddings (e.g., oblivious subspace embeddings mentioned in Section 3.2.1 or the embeddings constructed with random sampling of rows of an ε-embedding as in Remark 3.6.1) and one is interested in generating an ε-embedding $\mathbf{\Theta}$ (or rather computing the associated sketch), with a user-specified accuracy $\varepsilon \leq \tau$, for V (e.g., a subspace containing $\cup_{\mu \in \mathcal{P}} R_r(U_r; \mu)$) with nearly optimal number of rows. Moreover, we consider that no bound of the size of matrices to yield an ε-embedding is available or that the bound is pessimistic. It is only known that matrices with more than k_0 rows satisfy (3.54). This condition could be derived theoretically (as for Gaussian matrices) or deduced from practical experience (for P-SRHT). The matrix $\mathbf{\Theta}$ can be readily generated adaptively using $\bar{\omega}$ defined by (3.57) as an error indicator (see Algorithm 7). It directly follows by a union bound argument that $\mathbf{\Theta}$ generated in Algorithm 7 is an ε-embedding for V, with $\varepsilon \leq \tau$, with probability at least $1 - t\delta^*$, where t is the number of iterations taken by the algorithm.

To improve the efficiency, at each iteration of Algorithm 7 we could select the number of rows for $\mathbf{\Theta}^*$ adaptively instead of choosing it equal to k. In addition, the embeddings from previous iterations can be considered as parts of $\mathbf{\Theta}$ at further

Algorithm 7 Adaptive selection of the number of rows for $\mathbf{\Theta}$

```
Given: $k_0$, $\mathbf{V}$, $\tau > \varepsilon^*$.
Output: $\mathbf{V}^{\mathbf{\Theta}}$, $\mathbf{V}^{\mathbf{\Theta}^*}$
1. Set $k = k_0$ and $\bar{\omega} = \infty$.
while $\bar{\omega} > \tau$ do
    2. Generate $\mathbf{\Theta}$ and $\mathbf{\Theta}^*$ with $k$ rows and evaluate $\mathbf{V}^{\mathbf{\Theta}} := \mathbf{\Theta}\mathbf{V}$ and $\mathbf{V}^{\mathbf{\Theta}^*} := \mathbf{\Theta}^*\mathbf{V}$.
    3. Use (3.58) to compute $\bar{\omega}$.
    4. Increase $k$ by a constant factor.
end while
```

iterations.

3.7 Numerical experiments

This section is devoted to experimental validation of the methodology as well as realization of its great practical potential. The numerical tests were carried out on two benchmark problems that are difficult to tackle with the standard projection-based MOR methods due to a high computational cost and issues with numerical stability of the computation (or minimization) of the residual norm, or bad approximability of the solution manifold with a low-dimensional space.

In all the experiments we used oblivious $U \to \ell_2$ embeddings of the form

$$\mathbf{\Theta} := \mathbf{\Omega}\mathbf{Q},$$

where $\mathbf{Q}$ was taken as the (sparse) transposed Cholesky factor of $\mathbf{R}_U$ (or another metric matrix as in Remark 3.5.2) and $\mathbf{\Omega}$ as a P-SRHT matrix. The random embedding $\mathbf{\Gamma}$ used for the online efficiency was also taken as P-SRHT. Moreover, for simplicity in all the experiments the coefficient $\eta(\mu)$ for the error estimation was chosen as 1.

The experiments were executed on an Intel® Core™ i7-7700HQ 2.8GHz CPU, with 16.0GB RAM memory using Matlab® R2017b.

3.7.1 Acoustic invisibility cloak

The first numerical example is inspired by the development of invisibility cloaking [46, 48]. It consists in an acoustic wave scattering in 2D with a perfect scatterer covered in an invisibility cloak composed of layers of homogeneous isotropic materials. The geometry of the problem is depicted in Figure 3.1a. The cloak consists of 32 layers of equal thickness 1.5625 cm each constructed with 4 sublayers of equal thickness of alternating materials: mercury (a heavy liquid) followed by a light liquid. The properties (density and bulk modulus) of the light liquids are chosen to minimize the visibility of the scatterer for the frequency band [7.5, 8.5] kHz. The associated boundary value problem with the first order absorbing boundary conditions is the following

$$\begin{cases} \nabla \cdot (\rho^{-1} \nabla u) + \rho^{-1} \kappa^2 u &= 0, & \text{in } \Omega \\ (j\kappa - \frac{1}{2R_\Omega}) u + \frac{\partial u}{\partial \boldsymbol{n}} &= (j\kappa - \frac{1}{2R_\Omega}) u^{in} + \frac{\partial u^{in}}{\partial \boldsymbol{n}}, & \text{on } \Gamma \\ \frac{\partial u}{\partial \boldsymbol{n}} &= 0, & \text{on } \Gamma_s, \end{cases} \tag{3.60}$$

where $j = \sqrt{-1}$, $u = u^{in} + u^{sc}$ is the total pressure, with $u^{in} = \exp(-j\kappa(y-4))$ Pa·m being the pressure of the incident plane wave and u^{sc} being the pressure of the scattered wave, ρ is the material's density, $\kappa = \frac{2\pi f}{c}$ is the wave number, $c = \sqrt{\frac{b}{\rho}}$ is the speed of sound and b is the bulk modulus. The background material is chosen as water having density $\rho = \rho_0 = 997$ kg/m^3 and bulk modulus $b = b_0 = 2.23$ GPa. For the frequency $f = 8$ kHz the associated wave number of the background is $\kappa = \kappa_0 = 33.6$ m^{-1}. The i-th layer of the cloak (enumerated starting from the outermost layer) is composed of 4 alternating layers of mercury with density $\rho = \rho_m = 13500$ kg/m^3 and bulk modulus $b = b_m = 28$ GPa and a light liquid with density $\rho = \rho_i$ and bulk modulus $b = b_i$ given in Table 3.1. The light liquids from Table 3.1 can in practice be obtained, for instance, with the pentamode mechanical metamaterials [27, 96].

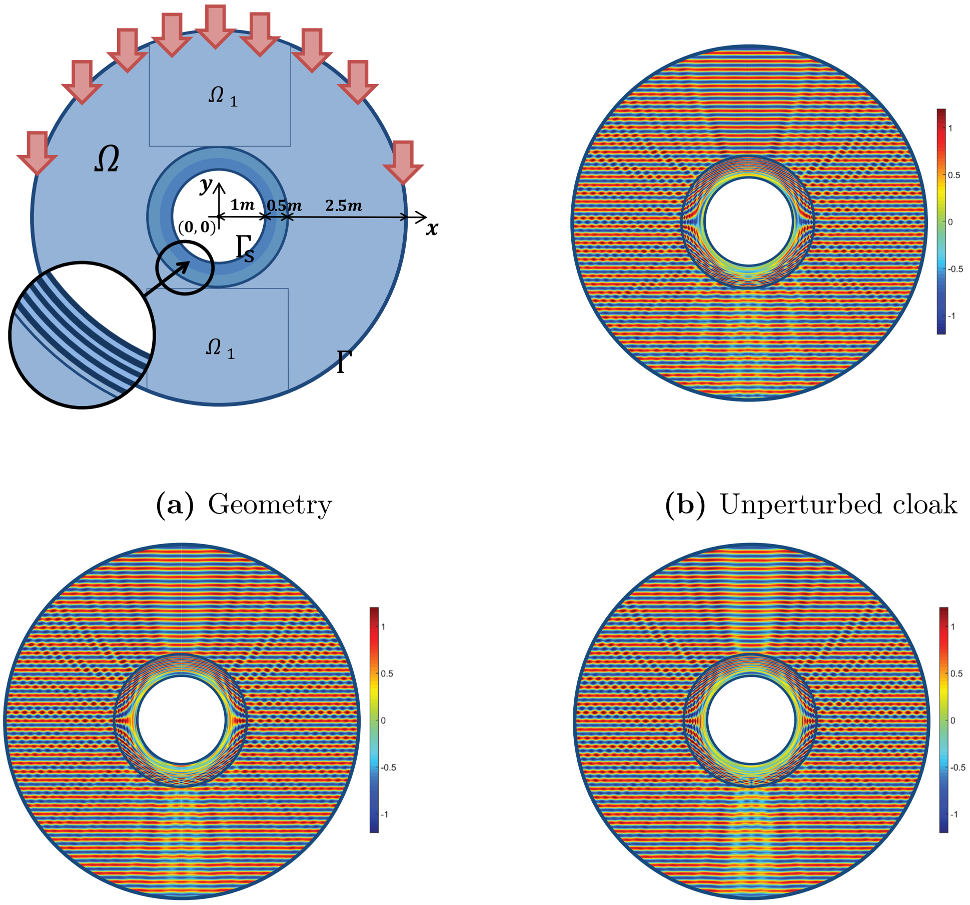
Table 3.1: The properties, density in kg/m^3 and bulk modulus in GPa, of the light liquids in the cloak.

i	ρ_i	b_i	i	ρ_i	b_i	i	ρ_i	b_i	i	ρ_i	b_i
1	231	0.483	9	179	0.73	17	56.1	0.65	25	9	1.34
2	121	0.328	10	166	0.78	18	59.6	0.687	26	9	2.49
3	162	0.454	11	150	0.745	19	40.8	0.597	27	9	2.5
4	253	0.736	12	140	0.802	20	32.1	0.682	28	9	2.5
5	259	0.767	13	135	0.786	21	22.5	0.521	29	9	0.58
6	189	0.707	14	111	0.798	22	15.3	0.6	30	9.5	1.91
7	246	0.796	15	107	0.8	23	10	0.552	31	9.31	0.709
8	178	0.739	16	78	0.656	24	9	1.076	32	9	2.44

The last 10 layers contain liquids with small densities that can be subject to imperfections during the manufacturing process. Moreover, the external conditions (such as temperature and pressure) may also affect the material's properties. We then consider a characterization of the impact of small perturbations of the density and the bulk modulus of the light liquids in the last 10 layers on the quality of the cloak in the frequency regime [7.8, 8.2] kHz. Assuming that the density and the bulk modulus may vary by 2.5%, the corresponding parameter set is

$$\mathcal{P} = \underset{23 \le i \le 32}{\times} [0.975\rho_i, 1.025\rho_i] \underset{23 \le i \le 32}{\times} [0.975 b_i, 1.025 b_i] \times [7.8 \text{ kHz}, 8.2 \text{ kHz}].$$

Note that in this case $\mathcal{P} \subset \mathbb{R}^{21}$.



(a) Geometry

(b) Unperturbed cloak

(c) Random snapshot

(d) Random snapshot

Figure 3.1: (a) Geometry of the invisibility cloak benchmark. (b) The real component of u in $\mathrm{Pa}\cdot\mathrm{m}$ for the parameter value $\mu \in \mathcal{P}$ corresponding to Table 3.1 and frequency $f = 8$ kHz. (c)-(d) The real component of u in $\mathrm{Pa}\cdot\mathrm{m}$ for two random samples from $\mathcal{P}$ with $f = 7.8$ kHz.

The quantity of interest is chosen to be the following

$$s(\mu) = l(u(\mu);\mu) = \|u(\mu) - u^{in}(\mu)\|^2_{L_2(\Omega_1)}/b_0 = \|u^{sc}(\mu)\|^2_{L_2(\Omega_1)}/b_0,$$

which represents the (rescaled, time-averaged) acoustic energy of the scattered wave concealed in the region Ω_1 (see Figure 3.1a). For the considered parameter set $s(\mu)$ is ranging from around $0.02A_s$ to around $0.1A_s$, where $A_s = \|u^{in}\|^2_{L_2(\Omega_1)}/b_0 = 7.2\mathrm{J}\cdot m \cdot \mathrm{Pa}/b_0$ at frequency 8 kHz.

The problem is symmetric with respect to the $x = 0$ axis, therefore only half of the domain has to be considered for discretization. For the discretization, we used piecewise quadratic approximation on a mesh of triangular (finite) elements. The mesh was chosen such that there were at least 20 degrees of freedom per wavelength, which is a standard choice for Helmholtz problems with a moderate wave number. It yielded approximately 400000 complex degrees of freedom for the discretization. Figures 3.1b to 3.1d depict the solutions $u(\mu)$ for different parameter values with quantities of interest $s(u(\mu)) = 0.033A_s, 0.044A_s$ and $0.065A_s$, respectively.

It is revealed that for this problem, considering the classical H^1 inner product for the solution space leads to dramatic instabilities of the projection-based MOR methods. To improve the stability, the inner product is chosen corresponding to the specific structure of the operator in (3.60). The solution space U is here equipped with the following inner product

$$\langle \mathbf{v}, \mathbf{w} \rangle_U := \langle \rho_s^{-1} \kappa_s^2 v, w \rangle_{L_2} + \langle \rho_s^{-1} \nabla v, \nabla w \rangle_{L_2}, \quad \mathbf{v}, \mathbf{w} \in U, \tag{3.61}$$

where v and w are the functions identified with $\mathbf{v}$ and $\mathbf{w}$, respectively, and ρ_s and κ_s are the density and the wave number associated with the unperturbed cloak (i. e., with properties from Table 3.1) at frequency 8 kHz.

The operator for this benchmark is directly given in an affine form with 23 terms. Furthermore, for online efficiency we used EIM to obtain an approximate affine representation of $u^{in}(\mu)$ (or rather a vector $\mathbf{u}^{in}(\mu)$ from U representing a discrete approximation of $u^{in}(\mu)$) and the right hand side vector with 50 affine terms (with error close to machine precision). The approximation space U_r of dimension $r = 150$ was constructed with a greedy algorithm (based on sketched minres projection) performed on a training set of 50000 uniform random samples in $\mathcal{P}$. The test set $\mathcal{P}_{test} \subset \mathcal{P}$ was taken as 1000 uniform random samples in $\mathcal{P}$.

Minimal residual projection. Let us first address the validation of the sketched minres projection from Section 3.3.2. For this we computed sketched (and standard) minres projections $\mathbf{u}_r(\mu)$ of $\mathbf{u}(\mu)$ onto U_r for each $\mu \in \mathcal{P}_{\text{test}}$ with sketching matrix $\mathbf{\Theta}$ of varying sizes. The error of approximation is here characterized by $\Delta_{\mathcal{P}} := \max_{\mu \in \mathcal{P}_{\text{test}}} \|\mathbf{r}(\mathbf{u}_r(\mu); \mu)\|_{U'} / \|\mathbf{b}(\mu_s)\|_{U'}$ and $e_{\mathcal{P}} := \max_{\mu \in \mathcal{P}_{\text{test}}} \|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U / \|\mathbf{u}^{in}\|_U$, where $\mathbf{b}(\mu_s)$ is the right hand side vector associated with the unperturbed cloak and the frequency $f = 8$ kHz (see Figures 3.2a and 3.2c). Furthermore, in Figure 3.2e we provide the characterization of the maximal error in the quantity of interest $e^s_{\mathcal{P}} := \max_{\mu \in \mathcal{P}_{\text{test}}} |s(\mu) - s_r(\mu)| / A_s$. For each size of $\mathbf{\Theta}$, 20 samples of the sketching matrix were considered to analyze the statistical properties of $e_{\mathcal{P}}$, $\Delta_{\mathcal{P}}$ and $e^s_{\mathcal{P}}$.

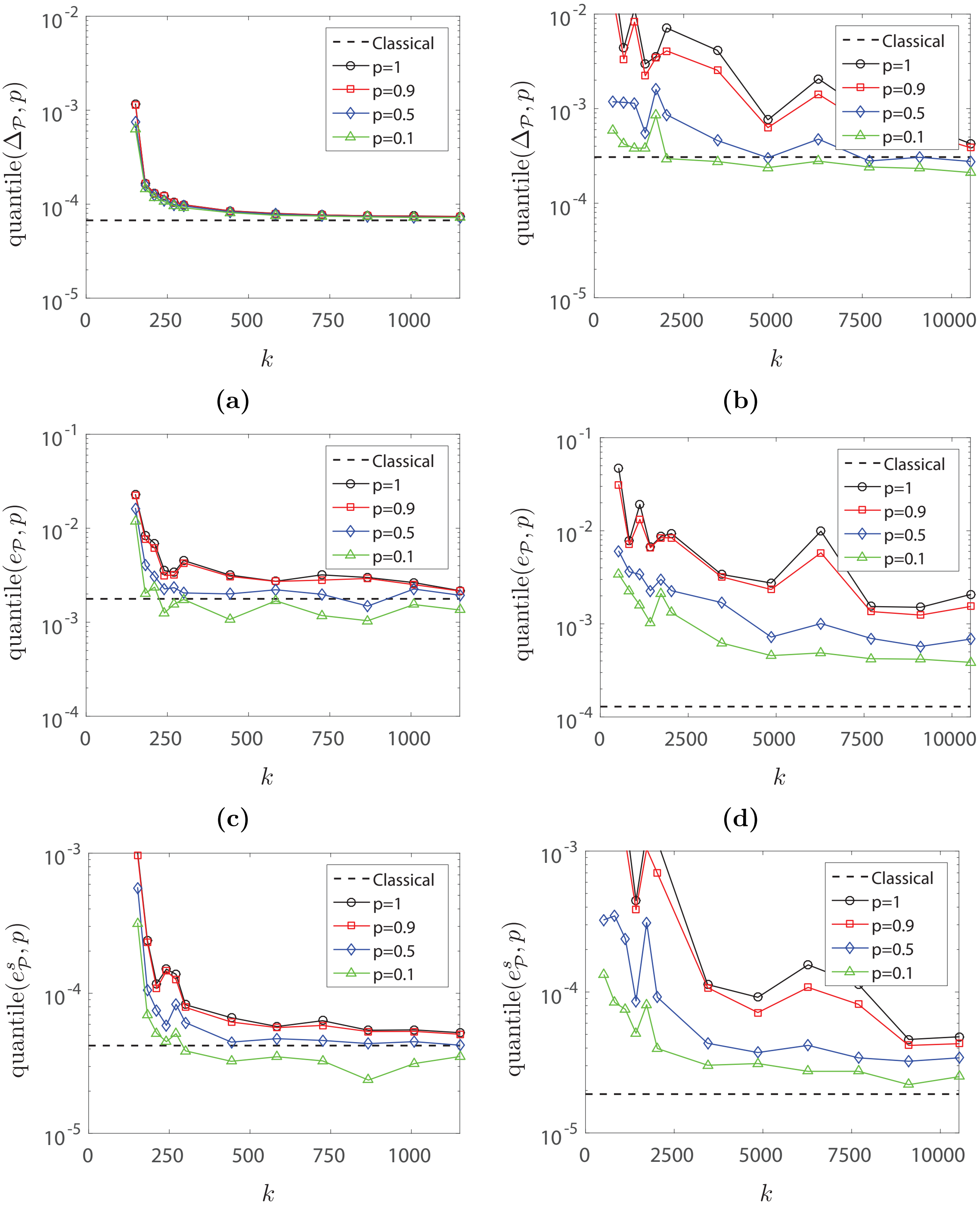
For comparison, along with the minimal residual projections we also computed the sketched (and classical) Galerkin projection introduced in Chapter 2. Figures 3.2b, 3.2d and 3.2f depict the errors $\Delta_{\mathcal{P}}$, $e_{\mathcal{P}}$, $e^s_{\mathcal{P}}$ of a sketched (and classical) Galerkin projection using $\mathbf{\Theta}$ of different sizes. Again, for each k we used 20 samples of $\mathbf{\Theta}$ to characterize the statistical properties of the error. We see that the classical Galerkin projection is more accurate in the exact norm and the quantity of interest than the standard minres projection. On the other hand, it is revealed that the minres projection is far better suited to random sketching.

From Figure 3.2 one can clearly report the (essential) preservation of the quality of the classical minres projection for $k \geq 500$. Note that for the minres projection a small deviation of $e_{\mathcal{P}}$ and $e^s_{\mathcal{P}}$ is observed. These errors are higher or lower than the standard values with (roughly) equal probability (for $k \geq 500$). In contrast to the minres projection, the quality of the Galerkin projection is not preserved even for large k up to 10000. This can be explained by the fact that the approximation of the Galerkin projection with random sketching is highly sensitive to the properties of the operator, which here is non-coercive and has a high condition number (for some parameter values), while the (essential) preservation of the accuracy of the standard minres projection by its sketched version is guaranteed regardless of the operator's properties. One can clearly see that the sketched minres projection using $\mathbf{\Theta}$ with just $k = 500$ rows yields better approximation (in terms of the maximal observed error) of the solution than the sketched Galerkin projection with $k = 5000$, even though the standard minres projection is less accurate than the Galerkin one.

As was discussed, random sketching improves not only efficiency but also has an advantage of making the reduced model less sensitive to round-off errors thanks to direct minimization of the (sketched) residual norm and not its square. Figure 3.3 depicts the maximal condition number $\kappa_{\mathcal{P}}$ over $\mathcal{P}_{\text{test}}$ of the reduced matrix $\mathbf{V}^{\mathbf{\Theta}}_r(\mu) := \mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_r$ associated with the sketched minres projection using reduced basis matrix $\mathbf{U}_r$ with (approximately) unit-orthogonal columns with respect to $\langle\cdot,\cdot\rangle_U$, for varying sizes of $\mathbf{\Theta}$. We also provide the maximal condition number of the reduced system of equations associated with the classical minres projection. It is observed that indeed random sketching yields an improvement of numerical stability by a square root.

Approximation of the output quantity. The next experiment was performed for a fixed approximation $\mathbf{u}_r(\mu)$ obtained with sketched minres projection using $\mathbf{\Theta}$ with 1000 rows. For such $\mathbf{u}_r(\mu)$, the approximate extraction of the quantity $s_r(\mu) = s(\mathbf{u}_r(\mu);\mu)$ from $\mathbf{u}_r(\mu)$ (represented by coordinates in reduced basis) was considered with the efficient procedure from Section 3.5.

The post-processing procedure was performed by choosing $\mathbf{l}(\mu) = \mathbf{u}_r(\mu) := \mathbf{u}_r(\mu) - \mathbf{u}^{in}(\mu)$ in (3.47). Furthermore, for better accuracy the solution space was here equipped with (semi-)inner product $\langle\cdot,\cdot\rangle_U := \langle\cdot,\cdot\rangle_{L_2(\Omega_1)}$ that is different from the inner product (3.61) considered for ensuring quasi-optimality of the minres projection and error estimation (see Remark 3.5.2). For such choices of $\mathbf{l}(\mu), \mathbf{u}_r(\mu)$ and $\langle\cdot,\cdot\rangle_U$, we employed a greedy search with error indicator $\|\mathbf{u}_r(\mu) - \mathbf{w}_p(\mu)\|^{\mathbf{\Theta}}_U$ over training set of 50000 uniform samples in $\mathcal{P}$ to find W_p. Then $s_r(\mu)$ was efficiently approximated by $s^\star_r(\mu)$ given in (3.51). In this experiment, the error is characterized by $e^s_{\mathcal{P}} = \max_{\mu\in\mathcal{P}_{\text{test}}} |s(\mu) - \tilde{s}_r(\mu)|/A_s$, where $\tilde{s}_r(\mu) = s_r(\mu)$ or $s^\star_r(\mu)$. The statistical properties of $e^s_{\mathcal{P}}$ for each value of k and $\dim(W_p)$ were obtained with 20 samples of $\mathbf{\Theta}$. Figure 3.4a exhibits the dependence of $e^s_{\mathcal{P}}$ on the size of $\mathbf{\Theta}$ with W_p of dimension $\dim(W_p) = 15$. Furthermore, in Figure 3.4b we provide the maximum value $e^s_{\mathcal{P}}$ from the computed samples for different sizes of $\mathbf{\Theta}$ and W_p. The accuracy of $s^\star_r(\mu)$ can be controlled



(a) **(b)**

(c) **(d)**

(e) **(f)**

Figure 3.2: The errors of the classical minres and Galerkin projections and quantiles of probabilities $p = 1, 0.9, 0.5$ and 0.1 over 20 samples of the errors of the sketched minres and Galerkin projections, versus the number of rows of $\mathbf{\Theta}$. (a) Residual error $\Delta_{\mathcal{P}}$ of standard and sketched minres projection. (b) Residual error $\Delta_{\mathcal{P}}$ of standard and sketched Galerkin projection. (c) Exact error $e_{\mathcal{P}}$ (in $\|\cdot\|_U$) of standard and sketched minres projection. (d) Exact error $e_{\mathcal{P}}$ (in $\|\cdot\|_U$) of standard and sketched Galerkin projection. (e) Quantity of interest error $e^s_{\mathcal{P}}$ of standard and sketched minres projection. (f) Quantity of interest error $e^s_{\mathcal{P}}$ of standard and sketched Galerkin projection.

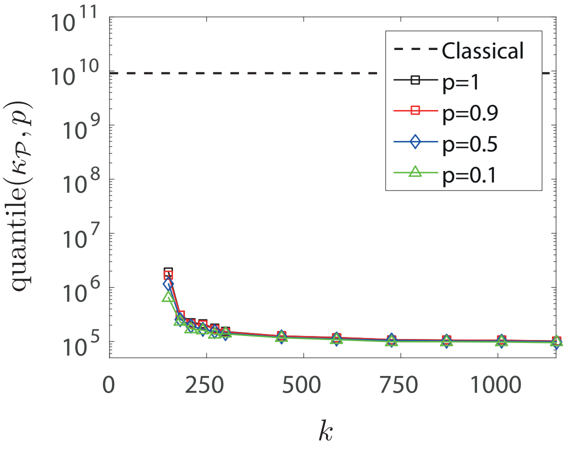


Figure 3.3: The maximal condition number over $\mathcal{P}_{\text{test}}$ of the reduced system associated with the classical minres projection and quantiles of probabilities $p = 1, 0.9, 0.5$ and 0.1 over 20 samples of the maximal condition number of the sketched reduced matrix $\mathbf{V}_r^{\boldsymbol{\Theta}}(\mu)$, versus the size of $\boldsymbol{\Theta}$.

by the quality of W_p for approximation of $\mathbf{u}_r(\mu)$ and the quality of $\|\cdot\|_U^{\boldsymbol{\Theta}}$ for approximation of $\|\cdot\|_U$. When W_p approximates well $\mathbf{u}_r(\mu)$, one can use a random correction with $\boldsymbol{\Theta}$ of rather small size, while in the alternative scenario the usage of a large random sketch is required. In this experiment we see that the quality of the output is nearly preserved with high probability when using W_p of dimension $\dim(W_p) = 20$ and a sketch of size $k = 1000$, or W_p of dimension $\dim(W_p) = 15$ and a sketch of size $k = 10000$. For less accurate W_p, with $\dim(W_p) \leq 10$, the preservation of the quality of the output requires larger sketches of sizes $k \geq 30000$. For most efficiency the dimension for W_p and the size for $\boldsymbol{\Theta}$ should be picked depending on the dimensions r and n of U_r and U, respectively, and the particular computational architecture. The increase of the considered dimension of W_p entails storage and operation with more high-dimensional vectors, while the increase of the sketch entails higher computational cost associated with storage and operation with the sketched matrix $\mathbf{U}_r^{\boldsymbol{\Theta}} = \boldsymbol{\Theta}\mathbf{U}_r$. Let us finally note that for this benchmark the approximate extraction of the quantity of interest with the procedure from Section 3.5 using $\dim(W_p) = 15$ and a sketch of size $k = 10000$, required in about 10 times less amount of storage and complexity than the classical exact extraction.

Certification of the sketch. Next the experimental validation of the procedure for a posteriori certification of the $\boldsymbol{\Theta}$-sketch or the sketched solution (see Section 3.6) is addressed. For this, we generated several $\boldsymbol{\Theta}$ of different sizes k and for each of them computed the sketched minres projections $\mathbf{u}_r(\mu) \in U_r$ for all $\mu \in \mathcal{P}_{\text{test}}$. Thereafter Propositions 3.6.2 and 3.6.3, with $V(\mu) := R_r(U_r;\mu)$ defined by (3.18), were considered for certification of the residual error estimates $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}^{\boldsymbol{\Theta}}$ or the quasi-optimality of $\mathbf{u}_r(\mu)$ in the residual error. Oblivious embeddings of varying sizes

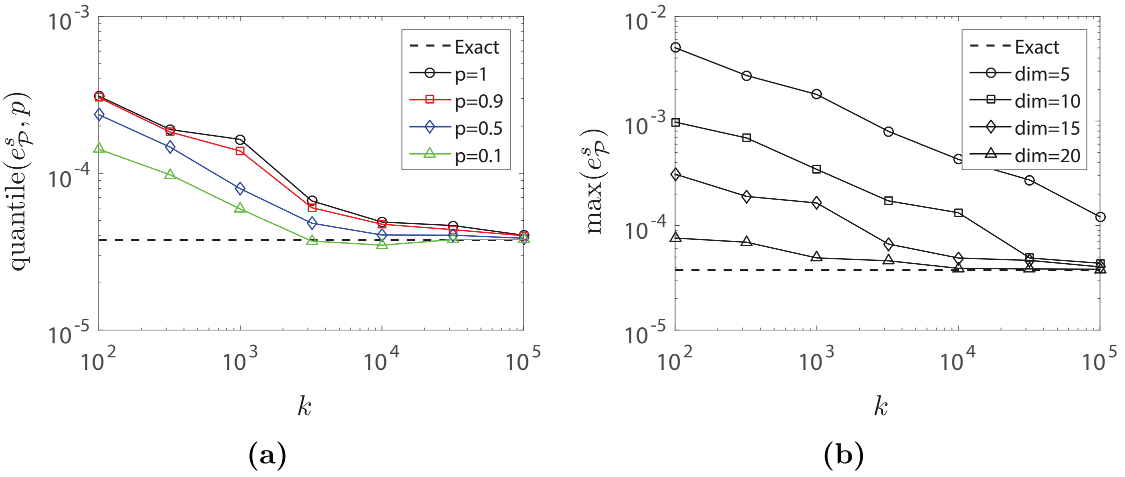


Figure 3.4: The error $e^{\mathrm{s}}_{\mathcal{P}}$ of $s_r(\mu)$ or its efficient approximation $s^\star_r(\mu)$ using W_p and $\mathbf{\Theta}$ of varying sizes. (a) The error of $s_r(\mu)$ and quantiles of probabilities $p = 1, 0.9, 0.5$ and 0.1 over 20 samples of $e^{\mathrm{s}}_{\mathcal{P}}$ associated with $s^\star_r(\mu)$ using W_p with $\dim(W_p) = 10$, versus sketch's size k. (b) The error of $s_r(\mu)$ and maximum over 20 samples of $e^{\mathrm{s}}_{\mathcal{P}}$ associated with $s^\star_r(\mu)$, versus sketch's size k for W_p of varying dimension.

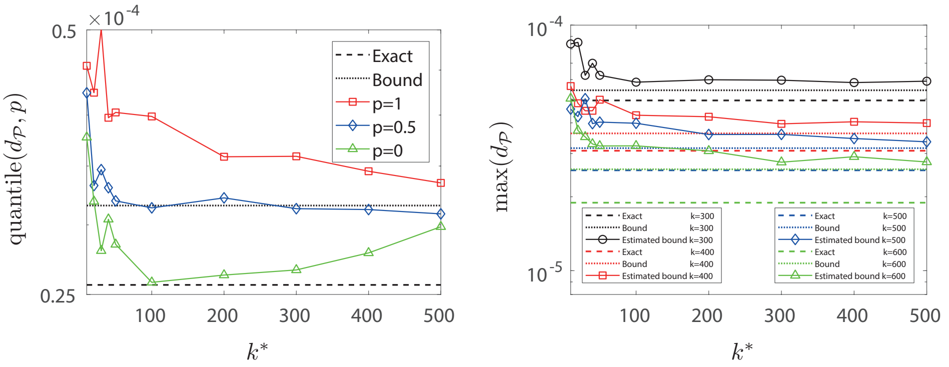
were tested as $\mathbf{\Theta}^*$. For simplicity it was assumed that all considered $\mathbf{\Theta}^*$ satisfy (3.54) with $\varepsilon^* = 0.05$ and small probability of failure δ^*.

By Proposition 3.6.2 the certification of the sketched residual error estimator $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}}_{U'}$ can be performed by comparing it to $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}^*}_{U'}$. More specifically, by (3.56) we have that with probability at least $1-4\delta^*$,

$$\begin{aligned}&|\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^2_{U'} - (\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}}_{U'})^2|^{1/2}\\ &\leq \left(|(\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}^*}_{U'})^2 - (\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}}_{U'})^2| + \frac{\varepsilon^*}{1-\varepsilon^*}(\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}^*}_{U'})^2 \right)^{1/2}.\end{aligned} \tag{3.62}$$

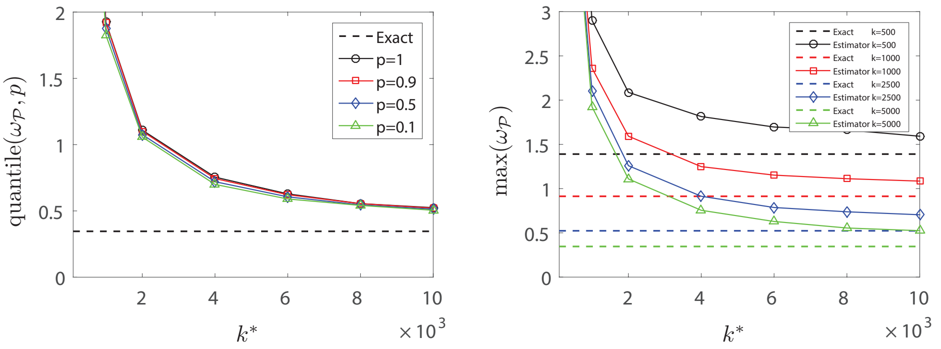
Figure 3.5 depicts $d_{\mathcal{P}} := \max_{\mu \in \mathcal{P}_{\text{test}}} d(\mathbf{u}_r(\mu);\mu)/\|\mathbf{b}(\mu_s)\|_{U'}$, where $d(\mathbf{u}_r(\mu);\mu)$ is the exact discrepancy $|\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^2_{U'} - (\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}}_{U'})^2|^{1/2}$ or its (probabilistic) upper bound in (3.62). For each $\mathbf{\Theta}$ and k^*, 20 realizations of $d_{\mathcal{P}}$ were computed for statistical analysis. We see that (sufficiently) tight upper bounds for $|\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^2_{U'} - (\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}}_{U'})^2|^{1/2}$ were obtained already when $k^* \geq 100$, which is in particular several times smaller than the size of $\mathbf{\Theta}$ required for quasi-optimality of $\mathbf{u}_r(\mu)$. This implies that the certification of the effectivity of the error estimator $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}}_{U'}$ by $\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}^*}_{U'}$ should require negligible computational costs compared to the cost of obtaining the solution (or estimating the error in adaptive algorithms such as greedy algorithms).

By Proposition 3.3.3, the quasi-optimality of $\mathbf{u}_r(\mu)$ can be guaranteed if $\mathbf{\Theta}$ is an ε-embedding for $V(\mu)$. The ε-embedding property of each $\mathbf{\Theta}$ was verified with Proposition 3.6.3. In Figure 3.6 we provide $\omega_{\mathcal{P}} := \max_{\mu \in \mathcal{P}_{\text{test}}} \tilde{\omega}(\mu)$ where $\tilde{\omega}(\mu) = \omega(\mu)$, which is the minimal value for ε such that $\mathbf{\Theta}$ is an ε-embedding for $V(\mu)$, or $\tilde{\omega}(\mu) = \bar{\omega}(\mu)$, which is the upper bound of $\omega(\mu)$ computed with (3.58) using $\mathbf{\Theta}^*$ of varying sizes. For illustration purposes we here allow the value ε in Definition 3.2.2 to be larger than 1. The statistical properties of $\omega_{\mathcal{P}}$ were obtained with 20 realizations for each $\mathbf{\Theta}$ and value of k^*. Figure 3.6a depicts the statistical characterization of $\omega_{\mathcal{P}}$ for $\mathbf{\Theta}$ of size $k = 5000$. The maximal value of $\omega_{\mathcal{P}}$ observed for each k^* and $\mathbf{\Theta}$ is presented in Figure 3.6b. It is observed that with a posteriori estimates from Proposition 3.6.3 using $\mathbf{\Theta}^*$ of size $k^* = 6000$, we here can guarantee with high probability that $\mathbf{\Theta}$ with $k = 5000$ satisfies an ε-embedding property for $\varepsilon \approx 0.6$. The theoretical bounds from Section 2.3.2 for $\mathbf{\Theta}$ to be an ε-embedding for $V(\mu)$ with $\varepsilon = 0.6$ yield much larger sizes, namely, for the probability of failure $\delta \leq 10^{-6}$, they require more than $k = 45700$ rows for Gaussian matrices and $k = 102900$ rows for SRHT. This proves Proposition 3.6.3 to be very useful for the adaptive selection of sizes of random matrices or for the certification of the sketched inner product $\langle \cdot, \cdot \rangle_U^{\mathbf{\Theta}}$ for all vectors in V. Note that the adaptive selection of the size of $\mathbf{\Theta}$ can also be performed without requiring $\mathbf{\Theta}$ to be an ε-embedding for $V(\mu)$ with $\varepsilon < 1$, based on the observation that oblivious embeddings yield preservation of the quality of the minres projection when they are ε-embeddings for $V(\mu)$ with small ε, which is possibly larger than 1 (see Remark 3.7.1).



(a) **(b)**

Figure 3.5: The discrepancy $|\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^2_{U'} - (\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|^{\mathbf{\Theta}}_{U'})^2|^{1/2}$ between the residual error and the sketched error estimator with $\mathbf{\Theta}$, and the upper bound of this value computed with (3.62). (a) The exact discrepancy for $\mathbf{\Theta}$ with $k = 500$ rows, the upper bound (3.62) of this discrepancy taking $\|\cdot\|^{\mathbf{\Theta}^*}_{U'} = \|\cdot\|_{U'}$, and quantiles of probabilities $p = 1, 0.5$ and 0 (i.e., the observed maximum, median and minimum) over 20 realizations of the (probabilistic) upper bound (3.62) versus the size of $\mathbf{\Theta}^*$. (b) The exact discrepancy, the upper bound (3.62) taking $\|\cdot\|^{\mathbf{\Theta}^*}_{U'} = \|\cdot\|_{U'}$, and the maximum of 20 realizations of the (probabilistic) upper bound (3.62) versus the number of rows k^* of $\mathbf{\Theta}^*$ for varying sizes k of $\mathbf{\Theta}$.



(a) **(b)**

Figure 3.6: The minimal value for ε such that $\mathbf{\Theta}$ is an ε-embedding for $V_r(\mu)$ for all $\mu \in \mathcal{P}_{\text{test}}$, and a posteriori random estimator of this value obtained with the procedure from Section 3.6 using $\mathbf{\Theta}^*$ with k^* rows. (a) The minimal value for ε for $\mathbf{\Theta}$ with $k = 5000$ rows and quantiles of probabilities $p = 1, 0.9, 0.5$ and 0.1 over 20 samples of the estimator, versus the size of $\mathbf{\Theta}^*$. (b) The minimal value for ε and the maximum of 20 samples of the estimator, versus the number of rows of $\mathbf{\Theta}^*$ for varying sizes of $\mathbf{\Theta}$.

Remark 3.7.1. *Throughout the paper the quality of $\mathbf{\Theta}$ (e.g., for approximation of minres projection in Section 3.3.2) was characterized by ε-embedding property. However, for this numerical benchmark the sufficient size for $\mathbf{\Theta}$ to be an ε-embedding for $V(\mu)$ is in several times larger than the one yielding an accurate approximation of the minres projection. In particular, $\mathbf{\Theta}$ with $k = 500$ rows with high probability provides an approximation with residual error very close to the minimal one, but it does not satisfy an ε-embedding property (with $\varepsilon < 1$), which is required for guaranteeing the quasi-optimality of $\mathbf{u}_r(\mu)$ with Proposition 3.3.3. A more reliable way for certification of the quality of $\mathbf{\Theta}$ for approximation of the minres projection onto U_r can be derived by taking into account that $\mathbf{\Theta}$ was generated from a distribution of oblivious embeddings. In such a case it is enough to only certify that $\|\cdot\|_U^{\mathbf{\Theta}}$ provides an approximate upper bound of $\|\cdot\|_U$ for all vectors in $V(\mu)$ without the need to guarantee that $\|\cdot\|_U^{\mathbf{\Theta}}$ is an approximate lower bound (that in practice is the main bottleneck). This approach is outlined below.*

We first observe that $\mathbf{\Theta}$ was generated from a distribution of random matrices such that for all $\mathbf{x} \in V(\mu)$, we have

$$\mathbb{P}\left(\left|\|\mathbf{x}\|_U^2 - (\|\mathbf{x}\|_U^{\mathbf{\Theta}})^2\right| \leq \varepsilon_0 \|\mathbf{x}\|_U^2\right) \geq 1 - \delta_0.$$

The values ε_0 and δ_0 can be obtained from the theoretical bounds from Section 2.3.2 or practical experience. Then one can show that for the sketched minres projection

$\mathbf{u}_r(\mu)$ associated with $\mathbf{\Theta}$, the relation

$$\|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'} \leq \sqrt{\frac{1+\varepsilon_0}{1-\omega(\mu)}} \min_{\mathbf{x}\in U_r} \|\mathbf{r}(\mathbf{x};\mu)\|_{U'}, \tag{3.63}$$

holds with probability at least $1-\delta_0$, where $\omega(\mu) < 1$ is the minimal value for ε such that for all $\mathbf{x} \in V(\mu)$

$$(1-\varepsilon)\|\mathbf{x}\|_U^2 \leq (\|\mathbf{x}\|_U^{\mathbf{\Theta}})^2.$$

The quasi-optimality of $\mathbf{u}_r(\mu)$ in the norm $\|\cdot\|_U$ rather than the residual norm can be readily derived from relation (3.63) by using the equivalence between the residual norm and the error in $\|\cdot\|_U$.

In this way a characterization of the quasi-optimality of the sketched minres projection with $\mathbf{\Theta}$ can be obtained from the a posteriori upper bound of $\omega(\mu)$ in (3.63). Note that since $\mathbf{\Theta}$ is an oblivious subspace embedding, the parameters ε_0 and δ_0 do not depend on the dimension of $V(\mu)$, which implies that the considered value for ε_0 should be an order of magnitude less than $\omega(\mu)$. Therefore, it can be a good way to choose ε_0 as $\omega(\mu)$ (or rather its upper bound) multiplied by a small factor, say 0.1.

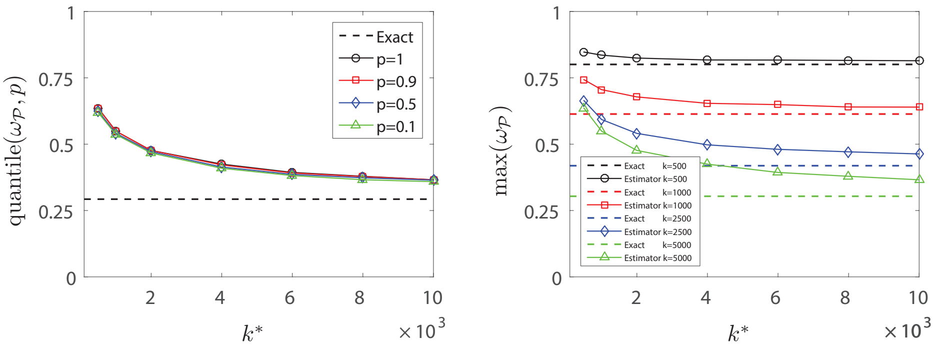
The (probabilistic) upper bound $\bar{\omega}(\mu)$ for $\omega(\mu)$ can be obtained a posteriori by following a similar procedure as the one from Proposition 3.6.3 described for verification of the ε-embedding property. More precisely, we can use similar arguments as in Proposition 3.6.3 to show that

$$\bar{\omega}(\mu) := 1-(1-\varepsilon^*) \min_{\mathbf{x}\in V/\{\mathbf{0}\}} \left(\frac{\|\mathbf{x}\|_U^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U^{\mathbf{\Theta}^*}}\right)^2$$

is an upper bound for $\omega(\mu)$ with probability at least $1-\delta^*$.

Let us now provide experimental validation of the proposed approach. For this we considered same sketching matrices $\mathbf{\Theta}$ as in the previous experiment for validation of the ε-embedding property. For each $\mathbf{\Theta}$ we computed $\omega_{\mathcal{P}} := \max_{\mu\in\mathcal{P}_{\text{test}}} \tilde{\omega}(\mu)$, where $\tilde{\omega}(\mu) = \omega(\mu)$ or its upper bound $\bar{\omega}(\mu)$ using $\mathbf{\Theta}^*$ of different sizes (see Figure 3.7). Again 20 realizations of $\omega_{\mathcal{P}}$ were considered for the statistical characterization of $\omega_{\mathcal{P}}$ for each $\mathbf{\Theta}$ and size of $\mathbf{\Theta}^*$. One can clearly see that the present approach provides better estimation of the quasi-optimality constants than the one with the ε-embedding property. In particular, the quasi-optimality guarantee for $\mathbf{\Theta}$ with $k = 500$ rows is experimentally verified. Furthermore, we see that in all the experiments the a posteriori estimates are lower than 1 even for $\mathbf{\Theta}^*$ of small sizes, yet they are larger than the exact values, which implies efficiency and robustnesses of the method. From Figure 3.7, a good accuracy of a posteriori estimates is with high probability attained for $k^* \geq k/2$.

Computational costs. For this benchmark, random sketching yielded drastic computational savings in the offline stage and considerably improved online efficiency.



(a) **(b)**

Figure 3.7: The minimal value for ε such that $(1-\varepsilon)\|\mathbf{x}\|_U^2 \leq (\|\mathbf{x}\|_U^{\boldsymbol{\Theta}})^2$ holds for all $\mathbf{x} \in V_r(\mu)$ and $\mu \in \mathcal{P}_{\text{test}}$, and a posteriori random estimator of this value using $\boldsymbol{\Theta}^*$ with k^* rows. (a) The minimal value for ε for $\boldsymbol{\Theta}$ with $k = 5000$ rows and quantiles of probabilities $p = 1, 0.9, 0.5$ and 0.1 over 20 realizations of the estimator, versus the size of $\boldsymbol{\Theta}^*$. (b) The minimal value for ε and the maximum of 20 realizations of the estimator versus the number of rows of $\boldsymbol{\Theta}^*$, for varying sizes of $\boldsymbol{\Theta}$.

To verify the gains for the offline stage, we executed two greedy algorithms for the generation of the reduced approximation space of dimension $r = 150$ based on the minres projection and the sketched minres projection, respectively. The standard algorithm resulted in a computational burden after reaching 96-th iteration due to exceeding the limit of RAM (16GB). Note that performing $r = 150$ iterations in this case would require around 25GB of RAM (mainly utilized for storage of the affine factors of $\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_r$). In contrast to the standard method, conducting $r = 150$ iterations of a greedy algorithm with random sketching using $\boldsymbol{\Theta}$ of size $k = 2500$ (and $\boldsymbol{\Gamma}$ of size $k' = 500$) for the sketched minres projection and $\boldsymbol{\Theta}^*$ of size $k^* = 250$ for the error certification, took only 0.65GB of RAM. Moreover, the sketch required only a minor part (0.2GB) of the aforementioned amount of memory, while the major part was consumed by the initialization of the full order model. The sketched greedy algorithm had a total runtime of 1.9 hours, from which 0.8 hours was spent on the computation of 150 snapshots, 0.2 hours on the provisional online solutions and 0.9 hours on random projections. Note that a drastic reduction of the offline runtime would as well be observed even in a computational environment with higher RAM (than 25GB), since random sketching in addition to reducing the memory consumption, also greatly improves the efficiency in terms of complexity and other metrics.

Next the improvement of online computational cost of minres projection is addressed. For this, we computed the reduced solutions on the test set with a standard method, which consists in assembling the reduced system of equations (representing the normal equation) from its affine decomposition (precomputed in the offline stage) and its subsequent solution with the built in Matlab® R2017b linear solver. The online solutions on the test set were additionally computed with the sketched method for comparison of runtimes and storage requirements. For this, for each parameter value, the reduced least-squares problem was assembled from the precomputed affine decompositions of $\mathbf{V}_r^{\mathbf{\Phi}}(\mu)$ and $\mathbf{b}^{\mathbf{\Phi}}(\mu)$ and solved with the normal equation using the built in Matlab® R2017b linear solver. Note that both methods proceeded with the normal equation. The difference was in the way how this equation was obtained. For the standard method it was directly assembled from the affine representation, while for the sketched method it was computed from the sketched matrices $\mathbf{V}_r^{\mathbf{\Phi}}(\mu)$ and $\mathbf{b}^{\mathbf{\Phi}}(\mu)$.

Table 3.2 depicts the runtimes and memory consumption taken by the standard and sketched online stages for varying sizes of the reduced space and $\mathbf{\Phi}$ (for the sketched method). The sketch's sizes were picked such that the associated reduced solutions with high probability had almost (higher by at most a factor of 1.2) optimal residual error. Our approach nearly halved the online runtime for all values of r from Table 3.2. Furthermore, the improvement of memory requirements was even greater. For instance, for $r = 150$ the online memory consumption was divided by 6.8.

Table 3.2: CPU times in seconds and amount of memory in MB taken by the standard and the efficient sketched online solvers for the solutions on the test set.

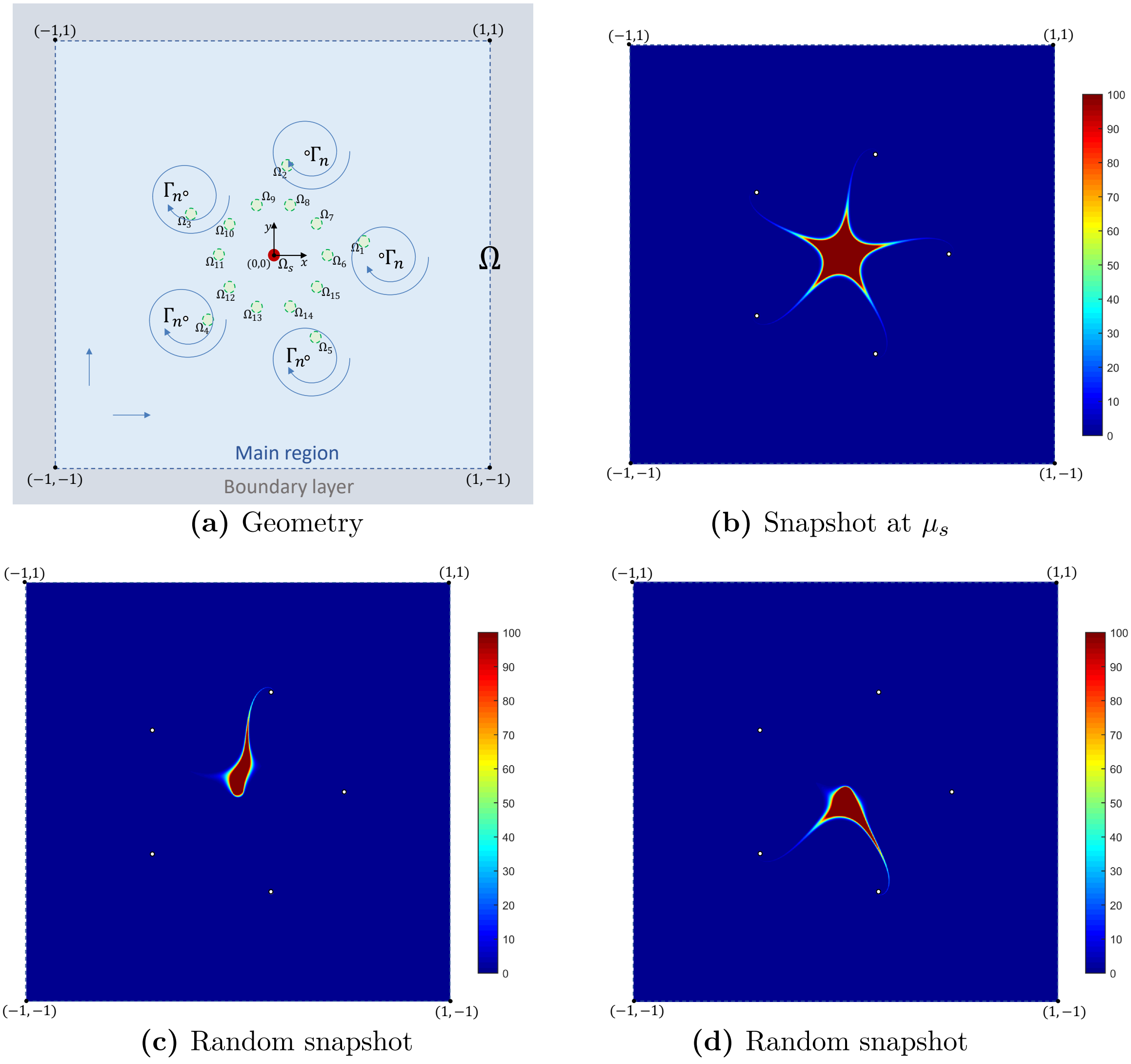
	Standard			Sketched		
	$r = 50$	$r = 100$	$r = 150$	$r = 50$ $k' = 300$	$r = 100$ $k' = 400$	$r = 150$ $k' = 500$
CPU	1.6	5.5	12	0.9	2.8	5.3
Storage	22	87	193	5.8	15	28

3.7.2 Advection-diffusion problem

The dictionary-based approximation method proposed in Section 3.4 is validated on a 2D advection dominated advection-diffusion problem defined on a complex flow. This problem is governed by the following equations

$$\begin{cases} -\epsilon\Delta u + \boldsymbol{\beta}\cdot\nabla u &= f, \quad \text{in } \Omega \\ u &= 0, \quad \text{on } \Gamma_{out} \\ \frac{\partial u}{\partial \boldsymbol{n}} &= 0, \quad \text{on } \Gamma_n. \end{cases} \tag{3.64}$$

where u is the unknown (temperature) field, $\epsilon := 0.0001$ is the diffusion (heat conduction) coefficient and $\boldsymbol{\beta}$ is the advection field. The geometry of the problem is as follows. First we have 5 circular pores of radius 0.01 located at points $\boldsymbol{x}_j = 0.5(\cos(2\pi j/5), \sin(2\pi j/5))$, $1 \leq j \leq 5$. The domain of interest is then defined as the square $[-10,10]^2$ without the pores, i.e, $\Omega := [-10,10]^2/\Omega_n$, with $\Omega_n := \cup_{1\leq j\leq 5}\{\boldsymbol{x} \in [-10,10]^2 : \|\boldsymbol{x}-\boldsymbol{x}_j\|_2 \leq 0.01\}$. The boundaries Γ_n and Γ_{out} are taken as $\partial\Omega_n$ and $\partial\Omega/\partial\Omega_n$, respectively. Furthermore, Ω is (notationally) divided into the main region inside $[-1,1]^2$, and the outer domain playing a role of a boundary layer. Finally, the force term f is nonzero in the disc $\Omega_s := \{\boldsymbol{x} \in \Omega : \|\boldsymbol{x}\|_2 \leq 0.025\}$. The geometric setup of the problem is presented in Figure 3.8a.



(a) Geometry

(b) Snapshot at μ_s

(c) Random snapshot

(d) Random snapshot

Figure 3.8: (a) Geometry of the advection-diffusion problem. (b) The solution field u for parameter value $\mu_s := (0, 0, 0.308, 0.308, 0.308, 0.308, 0.308, 0.616, 0.616, 0.616, 0.616, 0.616)$. (c)-(d) The solution field u for two random samples from $\mathcal{P}$.

The advection field is taken as a potential (divergence-free and curl-free) field consisting of a linear combination of 12 components,

$$\boldsymbol{\beta}(\boldsymbol{x}) = \mu_2 \cos(\mu_1)\hat{\boldsymbol{e}}_x + \mu_2 \sin(\mu_1)\hat{\boldsymbol{e}}_y + \sum_{i=1}^{10} \mu_i \boldsymbol{\beta}_i(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega,$$

where

$$\boldsymbol{\beta}_i(\boldsymbol{x}) = \begin{cases} \frac{-\hat{e}_r(\boldsymbol{x}_i)}{\|\boldsymbol{x}-\boldsymbol{x}_i\|} & \text{for } 1 \leq i \leq 5 \\ \frac{-\hat{\boldsymbol{e}}_\theta(\boldsymbol{x}_{i-5})}{\|\boldsymbol{x}-\boldsymbol{x}_{i-5}\|} & \text{for } 6 \leq i \leq 10. \end{cases} \tag{3.65}$$

The vectors $\hat{\boldsymbol{e}}_x$ and $\hat{\boldsymbol{e}}_y$ are the basis vectors of the Cartesian system of coordinates. The vectors $\hat{\boldsymbol{e}}_r(\mathbf{x}_j)$ and $\hat{\boldsymbol{e}}_\theta(\mathbf{x}_j)$ are the basis vectors of the polar coordinate system with the origin at point $\mathbf{x}_j$, $1 \leq j \leq 5$. Physically speaking, we have here a superposition of two uniform flows and five hurricane flows (each consisting of a sink and a rotational flow) centered at different locations. The source term is

$$f(\boldsymbol{x}) = \begin{cases} \frac{1}{\pi 0.025^2} & \text{for } \boldsymbol{x} \in \Omega_s, \\ 0 & \text{for } \boldsymbol{x} \in \Omega/\Omega_s. \end{cases}$$

We consider a multi-objective scenario, where one aims to approximate the average solution field $s^j(u)$, $1 \leq j \leq 15$, inside sensor Ω_j having a form of a disc of radius 0.025 located as in Figure 3.8a. The objective is to obtain sensor outputs for the parameter values $\mu := (\mu_1, \cdots, \mu_{12}) \in \mathcal{P} := \{[0,\ 2\pi] \times [0,\ 0.028] \times [0.308,\ 0.37]^5 \times [0.616,\ 0.678]^5\}$. Figures 3.8a to 3.8c present solutions $u(\mu)$ for few samples from $\mathcal{P}$.

The discretization of the problem was performed with the classical finite element method. A nonuniform mesh was considered with finer elements near the pores of the hurricanes, and larger ones far from the pores such that each element's Peclet number inside $[-1,1]^2$ was larger than 1 for any parameter value in $\mathcal{P}$. Moreover, it was revealed that for this benchmark the solution field outside the region $[-1,1]^2$ was practically equal to zero for all $\mu \in \mathcal{P}$. Therefore the outer region was discretized with coarse elements. For the discretization we used about 380000 and 20000 degrees of freedom in the main region and the outside boundary layer, respectively, which yielded approximately 400000 degrees of freedom in total.

The solution space is equipped with the inner product

$$\|\mathbf{w}\|_U^2 := \|\boldsymbol{\nabla} w\|_{L_2}^2, \ \mathbf{w} \in U,$$

which is compatible with the H_0^1 inner product.

For this problem, approximation of the solution with a fixed low-dimensional space is ineffective. The problem has to be approached with nonlinear approximation methods with parameter-dependent approximation spaces. For this, the classical *hp*-refinement method is computationally intractable due to high dimensionality of the parameter domain, which makes the dictionary-based approximation to be the most pertinent choice.

The training and test sets $\mathcal{P}_{\text{train}}$ and $\mathcal{P}_{\text{test}}$ were respectively chosen as 20000 and 1000 uniform random samples from $\mathcal{P}$. Then, Algorithm 6 was employed to generate dictionaries of sizes $K = 1500$, $K = 2000$ and $K = 2500$ for the dictionary-based approximation with $r = 100$, $r = 75$ and $r = 50$ vectors, respectively. For comparison, we also performed a greedy reduced basis algorithm (based on sketched minres projection) to generate a fixed reduced approximation space, which in particular coincides with Algorithm 6 with large enough r (here $r = 750$). Moreover, for more efficiency (to reduce the number of online solutions) at i-th iteration of Algorithm 6 and reduced basis algorithm instead of taking μ^{i+1} as a maximizer of $\Delta^{\boldsymbol{\Phi}}(\mathbf{u}_r(\mu);\mu)$ over $\mathcal{P}_{\text{train}}$, we relaxed the problem to finding any parameter-value such that

$$\Delta^{\boldsymbol{\Phi}}(\mathbf{u}_r(\mu^{i+1});\mu^{i+1}) \geq \max_{\mu \in \mathcal{P}_{\text{train}}} \min_{1\leq j \leq i} \Delta^{\boldsymbol{\Phi}}(\mathbf{u}_r^j(\mu);\mu), \tag{3.66}$$

where $\mathbf{u}_r^j(\mu)$ denotes the solution obtained at the j-th iteration. Note that (3.66) improved the efficiency, yet yielding at least as accurate maximizer of the dictionary-based width (defined in (3.25)) as considering $\mu^{i+1} := \arg\max_{\mu \in \mathcal{P}_{\text{train}}} \Delta^{\boldsymbol{\Phi}}(\mathbf{u}_r(\mu);\mu)$. For the error certification purposes, each 250 iterations the solution was computed on the whole training set and μ^{i+1} was taken as $\arg\max_{\mu \in \mathcal{P}_{\text{train}}} \Delta^{\boldsymbol{\Phi}}(\mathbf{u}_r(\mu);\mu)$. Figure 3.9 depicts the observed convergence of the greedy algorithms.

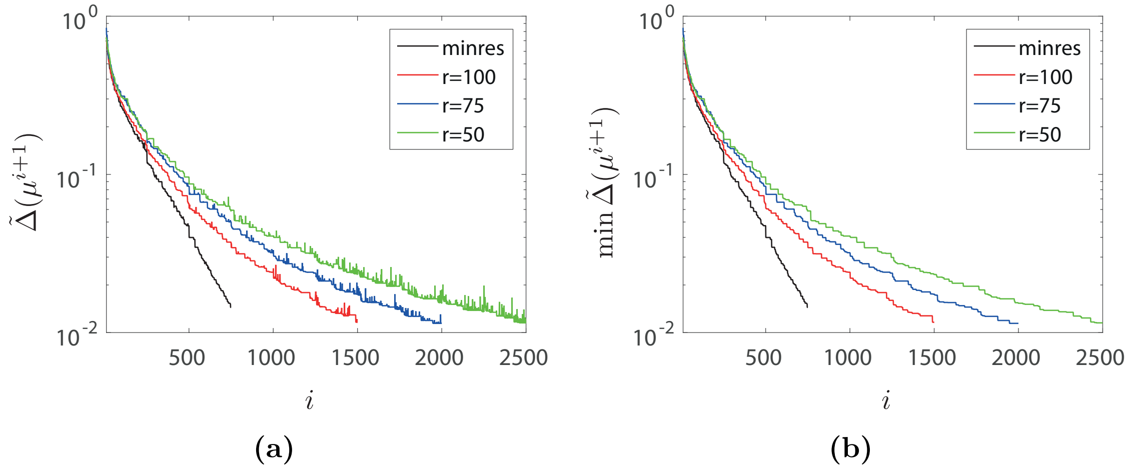


Figure 3.9: Convergences of Algorithm 6 for the dictionary generation for varying values of r, and the reduced basis greedy algorithm based on (sketched) minres projection. (a) The residual-based error indicator $\tilde{\Delta}(\mu^{i+1}) := \|\mathbf{u}_r(\mu^{i+1})\|_{U'}/\|\mathbf{b}\|_{U'}$. (b) The minimal value of the error indicator at parameter value μ^{i+1} at the first i iterations.

We see that at the first r iterations, the error decay for the dictionary generation practically coincides with the error decay of the reduced basis algorithm, which can be explained by the fact that the first r iterations of the two algorithms essentially coincide. The slope of the decay is then preserved for the reduced basis algorithm (even at high iterations), while it slowly subsequently degrades for dictionary-based approximation. The later method should still highly outperform the former one, since its online computational cost scales only linearly with the number of iterations. Furthermore, for the dictionary-based approximation the convergence of the error is moderately noisy. The noise is primarily due to approximating the solutions of online sparse least-squares problems with the orthogonal greedy algorithm, for which the accuracy can be sensitive to the enrichment of the dictionary with new vectors. The quality of online solutions can be improved by the usage of more sophisticated methods for sparse least-squares problems.

As it is clear from Figure 3.9, the obtained dictionaries provide approximations at least as accurate (on the training set) as the minres approximation with a fixed reduced space of dimension $r = 750$. Yet, the dictionary-based approximations are much more online-efficient. Table 3.3 provides the online complexity and storage requirements for obtaining the dictionary-based solutions for all $\mu \in \mathcal{P}_{\text{test}}$ (recall, $\#\mathcal{P}_{\text{test}} = 1000$) with the orthogonal greedy algorithm (Algorithm 5) from a sketch of size $k = 8r$, and the sketched minres solutions with QR factorization with Householder transformations of the sketched reduced matrix in (3.21) from a sketch of size $k = 4r$. In particular, we see that the dictionary-based approximation with $r = 75$ and $K = 2000$ yields a gain in complexity by a factor of 15 and memory consumption by a factor of 1.9. In Table 3.3 we also provide the associated runtimes and required RAM. It is revealed that the dictionary-based approximation with $K = 2000$ and $r = 75$ had an about 4 times speedup. The difference between the gains in terms of complexity and runtime can be explained by superb efficiency of the Matlab® R2017b least-squares solver. It is important to note that even more considerable boost of efficiency could be obtained by better exploitation of the structure of the dictionary-based reduced model, in particular, by representing the sketched matrix $\mathbf{V}_K^{\mathbf{\Theta}}(\mu)$ in a format well suited for the orthogonal greedy algorithm (e.g., a product of a dense matrix by several sparse matrices similarly as in [104, 135]).

Table 3.3: Computational cost of obtaining online solutions for all parameter values from the test set with the reduced basis method (based on sketched minres projection) and the dictionary-based approximations.

	RB, $r = 750$	$K = 1500,\ r = 100$	$K = 2000,\ r = 75$	$K = 2500,\ r = 50$
Complexity in flops	3.1×10^9	0.27×10^9	0.2×10^9	0.12×10^9
Storage in flns	2.9×10^8	1.6×10^8	1.6×10^8	1.3×10^8
CPU in s	400	124	113	100
Storage in MB	234	124	124	104

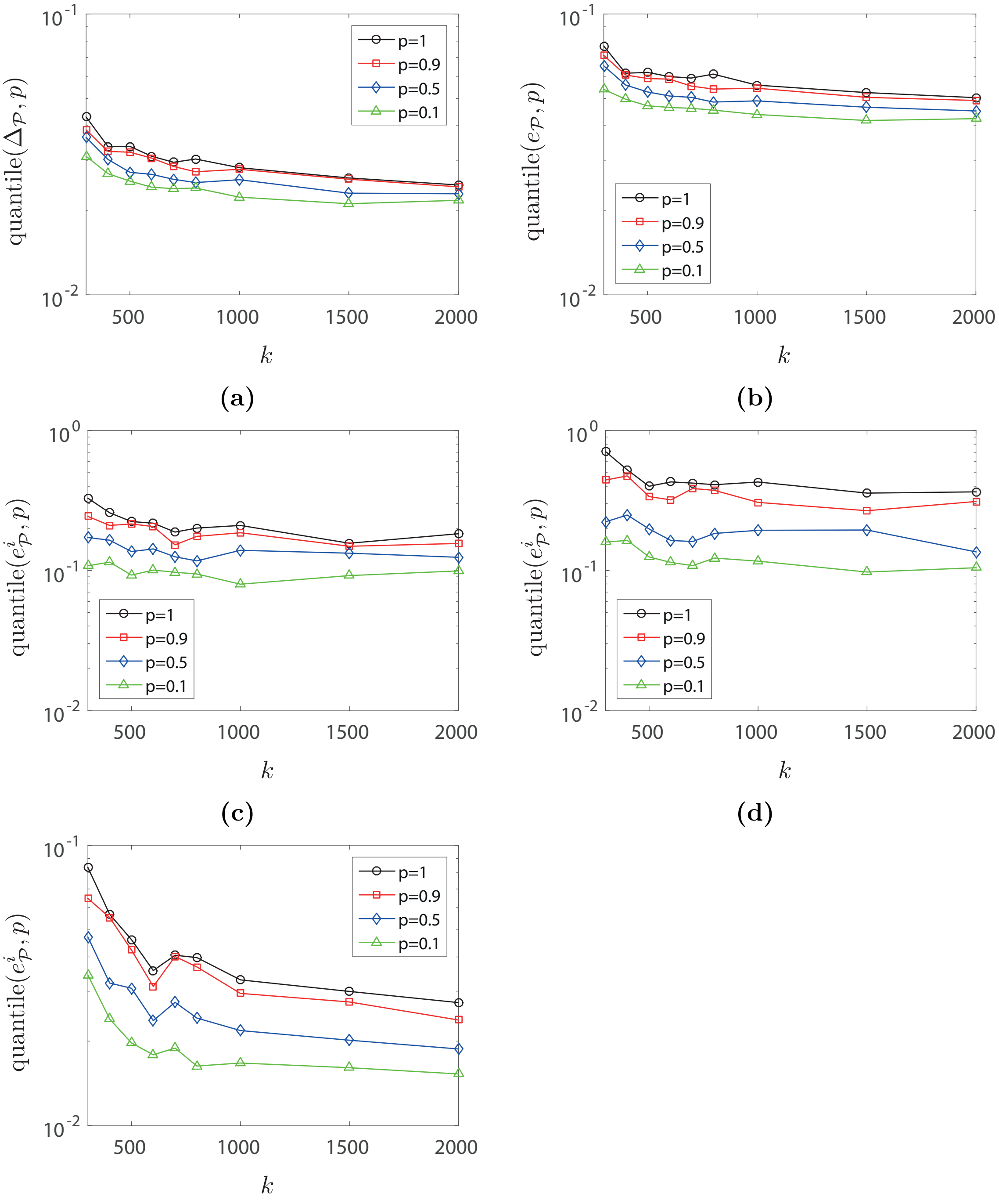
Further we provide statistical analysis of the dictionary-based approximation with $K = 2000$ and $r = 75$. For this we computed the associated dictionary-based solutions $\mathbf{u}_r(\mu)$ for all parameter values in the test set, considering $\mathbf{\Theta}$ of varying sizes. The accuracy of an approximation is characterized by the quantities $\Delta_{\mathcal{P}} := \max_{\mu\in\mathcal{P}_{\text{test}}} \|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'}/\|\mathbf{b}\|_{U'}$, $e_{\mathcal{P}} := \max_{\mu\in\mathcal{P}_{\text{test}}} \|\mathbf{u}(\mu)-\mathbf{u}_r(\mu)\|_U / \max_{\mu\in\mathcal{P}_{\text{test}}} \|\mathbf{u}(\mu)\|_U$ and $e^i_{\mathcal{P}} = \max_{\mu\in\mathcal{P}_{\text{test}}} |s^i(\mathbf{u}(\mu)) - s^i(\mathbf{u}_r(\mu))|$, $1 \leq i \leq 15$. Figure 3.10 depicts the dependence of $\Delta_{\mathcal{P}}$, $e_{\mathcal{P}}$ and $e^i_{\mathcal{P}}$ (for few selected values of i) on the size k of $\mathbf{\Theta}$. For each value of k, the statistical properties of $\Delta_{\mathcal{P}}$, $e_{\mathcal{P}}$ and $e^i_{\mathcal{P}}$ were characterized with 20 samples of $\Delta_{\mathcal{P}}$, $e_{\mathcal{P}}$ and $e^i_{\mathcal{P}}$. It is observed that for $k = 600$, the errors $\Delta_{\mathcal{P}}$ and $e_{\mathcal{P}}$ are concentrated around 0.03 and 0.06, respectively. Moreover, for all tested $k \geq 600$ we obtained nearly the same errors, which suggests preservation of the quality of the dictionary-based approximation by its sketched version at $k = 600$. A (moderate) deviation of the errors in the quantities of interest (even for very large k) can be explained by (moderately) low effectivity of representation of these errors with the error in $\|\cdot\|_U$, which we considered to control.

3.8 Conclusion

In this chapter we have extended the methodology from Chapter 2 to minres methods and proposed a novel nonlinear approximation method to tackle problems with a slow decay of Kolmogorov r-width. Furthermore, we additionally proposed efficient randomized ways for extraction of the quantity of interest and a posteriori certification of the reduced model's sketch. The results from this chapter can be used as a remedy of the drawbacks revealed in Chapter 2.

First, a method to approximate minres projection via random sketching was introduced. For each parameter value, the approximation is obtained by minimization of the ℓ_2-norm of a random projection of the residual. The associated minimization problem can be assembled from a sketch of a low-dimensional space containing the residual vectors and then solved with a standard routine such as QR factorization or (less stable) normal equation. This procedure enables drastic reduction of the computational cost in any modern computational architecture and improvement of numerical stability of the standard method. Precise conditions on the sketch to yield the (essential) preservation of the accuracy of the standard minres projection were provided. The conditions do not depend on the operator's properties, which implies robustness for ill-conditioned and non-coercive problems.

Then we proposed a dictionary-based approximation, where the solution is approximated by a minres projection onto a parameter-dependent reduced space with basis vectors (adaptively) selected from a dictionary. The characterization of the quasi-optimality of the proposed dictionary-based minres projection was provided. For each parameter value the solution can be efficiently approximated by the online solution of a small sparse least-squares problem assembled from random sketches of the dictionary vectors, which entails practical feasibility of the method. It was further shown that the preservation of the quasi-optimality constants of the sparse minres projection by its sketched version can be guaranteed if the random projection satisfies



(a)

(b)

(c)

(d)

(e)

Figure 3.10: Quantiles of probabilities $p = 1, 0.9, 0.5$ and 0.1 over 20 samples of the errors $\Delta_{\mathcal{P}}$, $e_{\mathcal{P}}$, $e^{\mathrm{i}}_{\mathcal{P}}$ of the dictionary-based approximation with $K = 2000$ and $r = 75$, versus the number of rows of $\mathbf{\Theta}$. (a) Residual error $\Delta_{\mathcal{P}}$. (b) Exact error $e_{\mathcal{P}}$. (c) Error $e^i_{\mathcal{P}}$ in the quantity of interest associated with sensor $i = 13$. (d) Error $e^i_{\mathcal{P}}$ in the quantity of interest associated with sensor $i = 8$. (e) Error $e^i_{\mathcal{P}}$ in the quantity of interest associated with sensor $i = 1$.

an ε-embedding property for a collection of low-dimensional spaces containing the set of residuals associated with a dictionary-based approximation. In particular, this condition can be ensured by using random projections constructed with SRHT or Gaussian matrices of sufficiently large sizes depending only logarithmically on the cardinality of the dictionary and the probability of failure.

This chapter has also addressed the efficient post-processing of the approximate solution from its coordinates associated with the given reduced basis (or dictionary). The extraction of the output quantity from an approximate solution can be done by a (sufficiently accurate) projection onto a low-dimensional space with a random sketching correction. This approach can be particularly important for the approximation of the linear output with an expensive extractor of the quantity of interest, as well as quadratic output and primal-dual correction.

Finally, we provided a probabilistic approach for a posteriori certification of the quality of the (random) embedding (and the associated sketch). This procedure is online-efficient and does not require operations on high dimensional vectors but only on their small sketches. It can be particularly useful for an adaptive selection of the size of a random sketching matrix since a priori bounds were revealed to be highly pessimistic.

The great applicability of the proposed methodology was realized on two benchmark problems difficult to tackle with standard methods. The experiments on the invisibility cloak benchmark proved that random sketching indeed provides high computational savings in both offline and online stages compared to the standard minres method while preserving the quality of the output. In particular, we could not even execute the classical greedy algorithm based on minres projection due to exceeding the RAM. Moreover, the improvement of numerical stability of computation (or minimization) of the residual error was also validated. It was verified experimentally that random sketching is much better suited to minres methods than to Galerkin methods. Furthermore, the procedure for a posteriori certification of the sketch was also validated in a series of experiments. It was revealed that this procedure can be used for (adaptive) selection of an (almost) optimal size of random sketching matrices in particular less by an order of magnitude than the theoretical bounds from Chapter 2.

Finally we considered an advection-diffusion benchmark defined on a complex flow. For this problem a slow-decay of the Kolmogorov r-width was revealed, which implied the necessity to use the dictionary-based approximation. It was verified that for this problem the dictionary-based approximation provided a boost of the online stage in more than an order of magnitude in complexity, in about 2 times in terms of memory, and in about 4 times in terms of runtime. Moreover, even higher computational savings could be obtained by representing the sketch of the dictionary in a more favorable format (e.g., as in [104, 135]), which we leave for the future research.

3.9 Appendix

Here we list the proofs of propositions from the chapter.

Proof of Proposition 3.3.1. The statement of the proposition follows directly from the definitions of the constants $\zeta_r(\mu)$ and $\iota_r(\mu)$, that imply

$$\zeta_r(\mu)\|\mathbf{u}(\mu)-\mathbf{u}_r(\mu)\|_U \leq \|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'} \leq \|\mathbf{r}(\mathbf{P}_{U_r}\mathbf{u}(\mu);\mu)\|_{U'} \leq \iota_r(\mu)\|\mathbf{u}(\mu)-\mathbf{P}_{U_r}\mathbf{u}(\mu)\|_U.$$

□

Proof of Proposition 3.3.2. The proof follows the one of Proposition 3.3.1. □

Proof of Proposition 3.3.3. By the assumption on $\mathbf{\Theta}$, we have that

$$\sqrt{1-\varepsilon}\|\mathbf{A}(\mu)\mathbf{x}\|_{U'} \leq \|\mathbf{A}(\mu)\mathbf{x}\|^{\mathbf{\Theta}}_{U'} \leq \sqrt{1+\varepsilon}\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}$$

holds for all $\mathbf{x} \in \mathrm{span}\{\mathbf{u}(\mu)\}+U_r$. Then (3.19) follows immediately. □

Proof of Proposition 3.3.5. Let $\mathbf{a} \in \mathbb{K}^r$ and $\mathbf{x} := \mathbf{U}_r\mathbf{a}$. Then

$$\frac{\|\mathbf{V}^{\mathbf{\Theta}}_r(\mu)\mathbf{a}\|}{\|\mathbf{a}\|} = \frac{\|\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_r\mathbf{a}\|}{\|\mathbf{a}\|} = \frac{\|\mathbf{A}(\mu)\mathbf{x}\|^{\mathbf{\Theta}}_{U'}}{\|\mathbf{x}\|^{\mathbf{\Theta}}_U}.$$

Since $\mathbf{\Theta}$ is an ε-embedding for U_r, we have

$$\sqrt{1-\varepsilon}\|\mathbf{x}\|_U \leq \|\mathbf{x}\|^{\mathbf{\Theta}}_U \leq \sqrt{1+\varepsilon}\|\mathbf{x}\|_U.$$

The statement of the proposition follows immediately. □

Proof of Proposition 3.4.1. Define

$$\mathcal{D}^{(i)}_K = \arg \min_{\#\mathcal{D}_K=K_i} \sup_{\mathbf{u}\in\mathcal{M}^{(i)}} \min_{W_{r_i}\in\mathcal{L}_{r_i}(\mathcal{D}_K)} \|\mathbf{u}-\mathbf{P}_{W_{r_i}}\mathbf{u}\|_U,$$

and

$$\mathcal{D}^*_K = \bigcup_{i=1}^{l} \mathcal{D}^{(i)}_K.$$

The following relations hold:

$$\begin{aligned}
\sum_{i=1}^{l}\sigma_{r_i}(\mathcal{M}^{(i)};K_i) &= \sum_{i=1}^{l} \sup_{\mathbf{u}\in\mathcal{M}^{(i)}} \min_{W_{r_i}\in\mathcal{L}_{r_i}(\mathcal{D}^{(i)}_K)} \|\mathbf{u}-\mathbf{P}_{W_{r_i}}\mathbf{u}\|_U \\
&\geq \sup_{\mu\in\mathcal{P}} \sum_{i=1}^{l} \min_{W_{r_i}\in\mathcal{L}_{r_i}(\mathcal{D}^{(i)}_K)} \|\mathbf{u}^{(i)}(\mu)-\mathbf{P}_{W_{r_i}}\mathbf{u}^{(i)}(\mu)\|_U \\
&\geq \sup_{\mu\in\mathcal{P}} \min_{W_r\in\mathcal{L}_r(\mathcal{D}^*_K)} \sum_{i=1}^{l} \|\mathbf{u}^{(i)}(\mu)-\mathbf{P}_{W_r}\mathbf{u}^{(i)}(\mu)\|_U \\
&\geq \sup_{\mu\in\mathcal{P}} \min_{W_r\in\mathcal{L}_r(\mathcal{D}^*_K)} \|\sum_{i=1}^{l}\mathbf{u}^{(i)}(\mu)-\mathbf{P}_{W_r}\sum_{i=1}^{l}\mathbf{u}^{(i)}(\mu)\|_U \geq \sigma_r(\mathcal{M};K).
\end{aligned}$$

□

Proof of Proposition 3.4.3. Let $U_r^*(\mu) := \arg\min_{W_r \in \mathcal{L}_r(\mathcal{D}_K)} \|\mathbf{u}(\mu) - \mathbf{P}_{W_r}\mathbf{u}(\mu)\|_U$. By definition of $\mathbf{u}_r(\mu)$ and constants $\zeta_{r,K}(\mu)$ and $\iota_{r,K}(\mu)$,

$$\begin{aligned}\zeta_{r,K}(\mu)\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U \leq \|\mathbf{r}(\mathbf{u}_r(\mu);\mu)\|_{U'} &\leq \|\mathbf{r}(\mathbf{P}_{U_r^*(\mu)}\mathbf{u}(\mu);\mu)\|_{U'} \\ &\leq \iota_{r,K}(\mu)\|\mathbf{u}(\mu) - \mathbf{P}_{U_r^*(\mu)}\mathbf{u}(\mu)\|_U,\end{aligned}$$

which ends the proof. □

Proof of Proposition 3.4.4. The proof exactly follows the one of Proposition 3.4.3 by replacing $\|\cdot\|_{U'}$ with $\|\cdot\|_{U'}^{\mathbf{\Theta}}$. □

Proof of Proposition 3.4.5. We have that

$$\sqrt{1-\varepsilon}\|\mathbf{A}(\mu)\mathbf{x}\|_{U'} \leq \|\mathbf{A}(\mu)\mathbf{x}\|_{U'}^{\mathbf{\Theta}} \leq \sqrt{1+\varepsilon}\|\mathbf{A}(\mu)\mathbf{x}\|_{U'}$$

holds for all $\mathbf{x} \in \mathrm{span}\{\mathbf{u}(\mu)\} + W_r$ with $W_r \in \mathcal{L}_r(\mathcal{D}_K)$. The statement of the proposition then follows directly from the definitions of $\zeta_{r,K}(\mu), \iota_{r,K}(\mu), \zeta_{r,K}^{\mathbf{\Theta}}(\mu)$ and $\iota_{r,K}^{\mathbf{\Theta}}(\mu)$.□

Proof of Proposition 3.4.6. Let $\mathbf{U}_r(\mu) \in \mathbb{K}^{n\times r}$ be a matrix whose column vectors are selected from the dictionary $\mathcal{D}_K$. Let $\mathbf{x} \in \mathbb{K}^r$ be an arbitrary vector and $\mathbf{w}(\mu) := \mathbf{U}_r(\mu)\mathbf{x}$. Let $\mathbf{z}(\mu) \in \mathbb{K}^K$ with $\|\mathbf{z}(\mu)\|_0 \leq r$ be a sparse vector such that $\mathbf{U}_K\mathbf{z}(\mu) = \mathbf{w}(\mu)$. Then

$$\begin{aligned}\frac{\|\mathbf{V}_r^{\mathbf{\Theta}}(\mu)\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{A}(\mu)\mathbf{U}_r\mathbf{x}\|}{\|\mathbf{x}\|} &= \frac{\|\mathbf{A}(\mu)\mathbf{w}(\mu)\|_{U'}^{\mathbf{\Theta}}}{\|\mathbf{x}\|} = \frac{\|\mathbf{A}(\mu)\mathbf{w}(\mu)\|_{U'}^{\mathbf{\Theta}}}{\|\mathbf{w}(\mu)\|_U}\frac{\|\mathbf{w}(\mu)\|_U}{\|\mathbf{x}\|} \\ \geq \zeta_{r,K}^{\mathbf{\Theta}}(\mu)\frac{\|\mathbf{U}_r(\mu)\mathbf{x}\|_U}{\|\mathbf{x}\|} &= \zeta_{r,K}^{\mathbf{\Theta}}(\mu)\frac{\|\mathbf{U}_K\mathbf{z}(\mu)\|_U}{\|\mathbf{z}(\mu)\|} \geq \zeta_{r,K}^{\mathbf{\Theta}}(\mu)\Sigma_{r,K}^{\min}.\end{aligned}$$

Similarly,

$$\frac{\|\mathbf{V}_r^{\mathbf{\Theta}}(\mu)\mathbf{x}\|}{\|\mathbf{x}\|} \leq \iota_{r,K}^{\mathbf{\Theta}}(\mu)\frac{\|\mathbf{U}_r(\mu)\mathbf{x}\|_U}{\|\mathbf{x}\|} \leq \iota_{r,K}^{\mathbf{\Theta}}(\mu)\Sigma_{r,K}^{\max}.$$

The statement of the proposition follows immediately. □

Proof of Proposition 3.5.3. Denote $\mathbf{x} := \mathbf{R}_U^{-1}\mathbf{l}(\mu)/\|\mathbf{l}(\mu)\|_{U'}$ and $\mathbf{y} := (\mathbf{u}_r(\mu) - \mathbf{w}_p(\mu))/\|\mathbf{u}_r(\mu) - \mathbf{w}_p(\mu)\|_U$. Let us consider $\mathbb{K} = \mathbb{C}$, which also accounts for the real case, $\mathbb{K} = \mathbb{R}$. Let

$$\omega := \frac{\langle\mathbf{x},\mathbf{y}\rangle_U - \langle\mathbf{x},\mathbf{y}\rangle_U^{\mathbf{\Theta}}}{|\langle\mathbf{x},\mathbf{y}\rangle_U - \langle\mathbf{x},\mathbf{y}\rangle_U^{\mathbf{\Theta}}|}.$$

Observe that $|\omega| = 1$ and $\langle\mathbf{x},\omega\mathbf{y}\rangle_U - \langle\mathbf{x},\omega\mathbf{y}\rangle_U^{\mathbf{\Theta}}$ is a real number.

By a union bound for the probability of success, $\mathbf{\Theta}$ is an ε-embedding for $\mathrm{span}(\mathbf{x}+\omega\mathbf{y})$ and $\mathrm{span}(\mathbf{x}-\omega\mathbf{y})$ with probability at least $1-2\delta$. Then, using the parallelogram identity we obtain

$$\begin{aligned}
4|\langle\mathbf{x},\mathbf{y}\rangle_U - \langle\mathbf{x},\mathbf{y}\rangle_U^{\mathbf{\Theta}}| &= |4\langle\mathbf{x},\omega\mathbf{y}\rangle_U - 4\langle\mathbf{x},\omega\mathbf{y}\rangle_U^{\mathbf{\Theta}}| \\
&= |\|\mathbf{x}+\omega\mathbf{y}\|_U^2 - \|\mathbf{x}-\omega\mathbf{y}\|_U^2 + 4\mathrm{Im}(\langle\mathbf{x},\omega\mathbf{y}\rangle_U) \\
&- \left((\|\mathbf{x}+\omega\mathbf{y}\|_U^{\mathbf{\Theta}})^2 - (\|\mathbf{x}-\omega\mathbf{y}\|_U^{\mathbf{\Theta}})^2 + 4\mathrm{Im}(\langle\mathbf{x},\omega\mathbf{y}\rangle_U^{\mathbf{\Theta}})\right)| \\
&= |\|\mathbf{x}+\omega\mathbf{y}\|_U^2 - (\|\mathbf{x}+\omega\mathbf{y}\|_U^{\mathbf{\Theta}})^2 - \left(\|\mathbf{x}-\omega\mathbf{y}\|_U^2 - (\|\mathbf{x}-\omega\mathbf{y}\|_U^{\mathbf{\Theta}})^2\right) \\
&- 4\mathrm{Im}(\langle\mathbf{x},\omega\mathbf{y}\rangle_U - \langle\mathbf{x},\omega\mathbf{y}\rangle_U^{\mathbf{\Theta}})| \\
&\leq \varepsilon\|\mathbf{x}+\omega\mathbf{y}\|_U^2 + \varepsilon\|\mathbf{x}-\omega\mathbf{y}\|_U^2 = 4\varepsilon.
\end{aligned}$$

We conclude that relation (3.52) holds with probability at least $1-2\delta$. □

Proof of Proposition 3.5.4. We can use a similar proof as in Proposition 3.5.3 using the fact that if $\mathbf{\Theta}$ is an ε-embedding for every subspace in $\mathcal{Y}$, then it satisfies the ε-embedding property for $\mathrm{span}(\mathbf{x}+\omega\mathbf{y})$ and $\mathrm{span}(\mathbf{x}-\omega\mathbf{y})$. □

Proof of Proposition 3.6.2. Using Proposition 3.5.3 with $\mathbf{l}(\mu) := \mathbf{R}_U\mathbf{x}$, $\mathbf{u}_r(\mu) := \mathbf{y}$, $\mathbf{w}_p(\mu) := \mathbf{0}$, $\mathbf{\Theta} := \mathbf{\Theta}^*$, $\varepsilon := \varepsilon^*$ and $\delta := \delta^*$, we have

$$\mathbb{P}(|\langle\mathbf{x},\mathbf{y}\rangle_U - \langle\mathbf{x},\mathbf{y}\rangle_U^{\mathbf{\Theta}^*}| \leq \varepsilon^*\|\mathbf{x}\|_U\|\mathbf{y}\|_U) \geq 1-2\delta^*, \tag{3.67}$$

from which we deduce that

$$\begin{aligned}
|\langle\mathbf{x},\mathbf{y}\rangle_U^{\mathbf{\Theta}^*} - \langle\mathbf{x},\mathbf{y}\rangle_U^{\mathbf{\Theta}}| - \varepsilon^*\|\mathbf{x}\|_U\|\mathbf{y}\|_U &\leq |\langle\mathbf{x},\mathbf{y}\rangle_U - \langle\mathbf{x},\mathbf{y}\rangle_U^{\mathbf{\Theta}}| \\
&\leq |\langle\mathbf{x},\mathbf{y}\rangle_U^{\mathbf{\Theta}^*} - \langle\mathbf{x},\mathbf{y}\rangle_U^{\mathbf{\Theta}}| + \varepsilon^*\|\mathbf{x}\|_U\|\mathbf{y}\|_U
\end{aligned} \tag{3.68}$$

holds with probability at least $1-2\delta^*$. In addition,

$$\mathbb{P}(|\|\mathbf{x}\|_U^2 - (\|\mathbf{x}\|_U^{\mathbf{\Theta}})^2| \leq \varepsilon^*\|\mathbf{x}\|_U^2) \geq 1-\delta^* \tag{3.69}$$

and

$$\mathbb{P}(|\|\mathbf{y}\|_U^2 - (\|\mathbf{y}\|_U^{\mathbf{\Theta}})^2| \leq \varepsilon^*\|\mathbf{y}\|_U^2) \geq 1-\delta^*. \tag{3.70}$$

The statement of the proposition can be now derived by combining (3.68) to (3.70) and using a union bound argument. □

Proof of Proposition 3.6.3. Observe that

$$\omega = \max\left\{1 - \min_{\mathbf{x}\in V/\{\mathbf{0}\}}\left(\frac{\|\mathbf{x}\|_U^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U}\right)^2, \max_{\mathbf{x}\in V/\{\mathbf{0}\}}\left(\frac{\|\mathbf{x}\|_U^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U}\right)^2 - 1\right\}.$$

Let us make the following assumption:

$$1-\min_{\mathbf{x}\in V/\{\mathbf{0}\}}\left(\frac{\|\mathbf{x}\|_U^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U}\right)^2 \geq \max_{\mathbf{x}\in V/\{\mathbf{0}\}}\left(\frac{\|\mathbf{x}\|_U^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U}\right)^2-1.$$

For the alternative case the proof is similar.

Next, we show that $\bar{\omega}$ is an upper bound for ω with probability at least $1-\delta^*$. Define $\mathbf{x}^* := \arg\min_{\mathbf{x}\in V/\{\mathbf{0}\}, \|\mathbf{x}\|_U=1}\|\mathbf{x}\|_U^{\mathbf{\Theta}}$. By definition of $\mathbf{\Theta}^*$,

$$1-\varepsilon^* \leq \left(\|\mathbf{x}^*\|_U^{\mathbf{\Theta}^*}\right)^2 \tag{3.71}$$

holds with probability at least $1-\delta^*$. If (3.71) is satisfied, we have

$$\bar{\omega} \geq 1-(1-\varepsilon^*)\min_{\mathbf{x}\in V/\{\mathbf{0}\}}\left(\frac{\|\mathbf{x}\|_U^{\mathbf{\Theta}}}{\|\mathbf{x}\|_U^{\mathbf{\Theta}^*}}\right)^2 \geq 1-(1-\varepsilon^*)\left(\frac{\|\mathbf{x}^*\|_U^{\mathbf{\Theta}}}{\|\mathbf{x}^*\|_U^{\mathbf{\Theta}^*}}\right)^2 \geq 1-(\|\mathbf{x}^*\|_U^{\mathbf{\Theta}})^2=\omega.$$

□

Proof of Proposition 3.6.4. By definition of ω and the assumption on $\mathbf{\Theta}^*$, for all $\mathbf{x}\in V$, it holds

$$|\|\mathbf{x}\|_U^2-(\|\mathbf{x}\|_U^{\mathbf{\Theta}^*})^2| \leq \omega^*\|\mathbf{x}\|_U^2, \quad \text{and} \quad |\|\mathbf{x}\|_U^2-(\|\mathbf{x}\|_U^{\mathbf{\Theta}})^2| \leq \omega\|\mathbf{x}\|_U^2.$$

The above relations and the definition (3.57) of $\bar{\omega}$ yield

$$\bar{\omega} \leq \max\left\{1-(1-\varepsilon^*)\frac{1-\omega}{1+\omega^*}, (1+\varepsilon^*)\frac{1+\omega}{1-\omega^*}-1\right\} = (1+\varepsilon^*)\frac{1+\omega}{1-\omega^*}-1,$$

which ends the proof. □

Chapter 4

Parameter-dependent preconditioners for model order reduction

The performance of projection-based model order reduction methods for solving parameter-dependent linear systems of equations highly depends on the properties of the operator, which can be improved by preconditioning. This chapter presents an online-efficient procedure to estimate the condition number of a large-scale parameter-dependent (preconditioned) matrix. We also provide an effective way to estimate the quasi-optimality constants of the Petrov-Galerkin projection on a given approximation space of moderately large dimension and the residual-based error estimation. All the estimates are defined by error indicators measuring a discrepancy between the (preconditioned) matrix and the identity (or some positive-definite matrix defining the metric of interest). An effective parameter-dependent preconditioner can be constructed by interpolation of matrix inverse based on minimization of an error indicator. The minimization of the proposed error indicators requires the solution of small least-squares problems which can be efficiently performed online for each parameter value. The obtained preconditioner can be readily used for improving the quality of Petrov-Galerkin projection or for effective error certification without the need to estimate stability constants.

The heavy offline computations are circumvented by using a random sketching technique, which consists in estimating the norms of high-dimensional matrices and vectors by ℓ_2-norms of their random projections in low-dimensional spaces. For this we extend the framework from Chapter 2. Random sketching allows drastic reduction of the computational cost of the offline stage in terms of number of flops, memory consumption, scalability, etc. Moreover, it improves numerical stability. The random projections are obtained using random sketching matrices which are $\ell_2 \to \ell_2$ oblivious subspace embeddings such as rescaled Gaussian matrices and Subsampled Randomized Hadamard Transform. We provide sufficient conditions on the sizes of random sketching matrices to control the accuracy of estimation with a user-specified probability of success.

Contents

4.1 Introduction

We consider a large-scale parameter-dependent system of equations

$$\mathbf{A}(\mu)\mathbf{u}(\mu) = \mathbf{b}(\mu), \ \mu \in \mathcal{P}, \tag{4.1}$$

where $\mathcal{P}$ is the parameter set. Such a system may result from the discretization of a parameter-dependent PDE. We assume that the solution manifold $\{\mathbf{u}(\mu) : \mu \in \mathcal{P}\}$ can be well approximated by a projection onto a space of moderately large dimension. The linear system of equations (4.1) can then be approximately solved using projection-based model order reduction (MOR) methods such as Reduced Basis (RB) method, Proper Orthogonal Decomposition (POD) and (recycling) Krylov methods (see [24, 25, 89, 124, 128] and the references therein). The performance of projection-based methods highly depends on the properties of the matrix $\mathbf{A}(\mu)$, which can be improved by preconditioning.

Let the solution space be characterized by a weighted Euclidean (or Hermitian) inner product $\langle \cdot, \cdot \rangle_U := \langle \mathbf{R}_U \cdot, \cdot \rangle_2$, where $\mathbf{R}_U$ is some self-adjoint positive definite matrix. More details regarding the problem's setting can be found in Section 4.1.2. Let the preconditioner $\mathbf{P}(\mu)$ be an approximate inverse of $\mathbf{A}(\mu)$. Then the (approximate) solution of (4.1) can be obtained from

$$\mathbf{B}(\mu)\mathbf{u}(\mu) = \mathbf{f}(\mu), \ \mu \in \mathcal{P}, \tag{4.2}$$

where $\mathbf{B}(\mu) := \mathbf{R}_U \mathbf{P}(\mu)\mathbf{A}(\mu)$ and $\mathbf{f}(\mu) := \mathbf{R}_U \mathbf{P}(\mu)\mathbf{b}(\mu)$. If $\mathbf{P}(\mu)\mathbf{A}(\mu)$ is close to the identity matrix, then $\mathbf{B}(\mu)$ should have better properties than the original operator $\mathbf{A}(\mu)$, which implies better performances of projection-based methods. In particular, if $\mathbf{P}(\mu) = \mathbf{A}(\mu)^{-1}$ then (4.2) is perfectly conditioned (relatively to the metric induced by $\mathbf{R}_U$). It is important to note that in the context of projection-based MOR, the invertibility of $\mathbf{P}(\mu)$ is not required for obtaining an approximate solution to (4.1). Since we operate only on a subset of vectors it is sufficient to ensure that $\mathbf{P}(\mu)\mathbf{A}(\mu)$ is close to the identity on this subset. Note also that the computation of the explicit form of $\mathbf{B}(\mu)$ can be extremely expensive and has to be avoided. Instead, this matrix should be operated as an implicit map outputting products with vectors.

In the present chapter we provide efficiently computable estimators of the quality of $\mathbf{B}(\mu)$ for the solution of (4.2) with projection-based methods or for residual-based error estimation. Each estimator basically measures a discrepancy between $\mathbf{B}(\mu)$ and $\mathbf{R}_U$ (with respect to a certain semi-norm), and is seen as an error indicator on $\mathbf{P}(\mu)$ as an approximation of the inverse of $\mathbf{A}(\mu)$. The proposed error indicators can be readily employed to efficiently estimate the quasi-optimality constants associated with the given preconditioner or to construct $\mathbf{P}(\mu)$ by interpolation of the inverse of $\mathbf{A}(\mu)$. Unlike the minimization of the condition number of $\mathbf{B}(\mu)$ or the quasi-optimality constants, the minimization of each error indicator over a low-dimensional space of matrices is a small least-squares problem, which can be efficiently solved online. The heavy offline computations are here circumvented with randomized linear algebra. More specifically, a drastic reduction of the computational cost is attained by the usage of the framework from Chapter 2 and its extension to the context of approximation of inner products between matrices. The ℓ_2-embeddings are no longer seen as matrices, but rather as linear maps from a space of matrices to a low-dimensional Euclidean (or Hermitian) space. In Section 4.3 we propose a probabilistic way for the construction of ℓ_2-embeddings for matrices, and in Section 4.4 provide its theoretical characterization.

The construction of an efficient parameter-dependent preconditioner has been addressed in [47, 60, 100, 138, 160]. In particular, in [160] the authors proposed to use randomized linear algebra for the efficient construction of a preconditioner by interpolation of matrix inverse. This principle is taken as the starting point for the present chapter. Randomized linear algebra has also been employed for improving MOR methods in [12, 37, 141].

4.1.1 Contributions

We here consider preconditioners in three different contexts. Besides the multi-purpose context as in [160], where one is interested in estimation (or minimization) of the condition number, we also consider Galerkin projections onto fixed approximation spaces, and residual-based error certification. A detailed presentation of the major contributions is given below.

Preconditioner for multi-purpose context

This work presents a generalization and improvement of the methodology introduced in [160]. First of all, the quality of the preconditioner is characterized with respect to a general norm represented by a self-adjoint positive define matrix instead of the ℓ_2-norm. This is important, for instance, in the context of numerical methods for PDEs to control the quality of an approximation regardless of the used discretization. Secondly, the theoretical bounds from [160] for the size of sketching matrices are considerably improved. For instance our bound for SRHT is linear in the dimension m of the low-dimensional space of operators, and not quadratic as in [160]. Furthermore, thanks to the (extended) framework presented in Chapter 2, we here obtain a great improvement of the efficiency (both offline and online) and numerical stability of the algorithms. More specifically, if $\mathbf{A}(\mu)$ is a $n \times n$ sparse matrix and admits an affine expansion with m_A terms[1], if $\mathbf{P}(\mu)$ is a linear combination of p basis matrices, each requiring $\mathcal{O}(nk_P)$ (for some small k_P) complexity and amount of storage for multiplication by a vector, and if k is the dimension of the sketch, then the

[1] A parameter-dependent quantity $\mathbf{v}(\mu)$ with values in vector space V over a field $\mathbb{K}$ is said to admit an affine representation if $\mathbf{v}(\mu) = \sum_{i=1}^{d} \mathbf{v}_i \lambda_i(\mu)$ with $\lambda_i(\mu) \in \mathbb{K}$ and $\mathbf{v}_i \in V$.

precomputation of our error indicator (using SRHT) takes only $\mathcal{O}(kn(m_A p \log(k) + k_P))$ flops and $\mathcal{O}(n(km_A \log(k) + k_P))$ bytes of memory, while the approach from [160] takes $\mathcal{O}(kn(m_A^2 p + k_P))$ flops and $\mathcal{O}(n(km_A^2 p + k_P))$ bytes of memory. Moreover, we also improve the efficiency and numerical stability of the online stage. The online assembling of the reduced matrix for the computation (or minimization) of the indicator in [160] takes $\mathcal{O}(m_A^2 p^2)$ flops, while our approach essentially consumes only $\mathcal{O}(k' m_A p)$ flops, where $k' = \mathcal{O}(1)$ for the approximation of the indicator or $k' = \mathcal{O}(p)$ for its minimization. Our approach is also less sensitive to round-off errors since we proceed with direct computation (or minimization) of a (sketched) norm and not its square. We also derive a quasi-optimality result for the preconditioned Galerkin projection and error estimation with the proposed error indicator.

Preconditioner for Galerkin projection

The estimation of the operator norm by a Hilbert-Schmidt norm (as used in the multi-purpose context) can be very ineffective. In general a very high overestimation is possible. For numerical methods for PDEs, this may result in a high sensitivity to discretization. We show how to overcome this issue, if the preconditioner is used for a Galerkin projection onto a moderately large approximation space. In such a case the effective error indicators can be obtained by ignoring the component of the residual which is orthogonal to the approximation space (see Section 4.2.2).

Preconditioner for error certification

The error $\|\mathbf{u}(\mu) - \mathbf{u}_r(\mu)\|_U$ of an approximation $\mathbf{u}_r(\mu)$ of the solution $\mathbf{u}(\mu)$ can be estimated by a sketched norm (see Chapter 2 for details) of the preconditioned residual $\mathbf{f}(\mu) - \mathbf{B}(\mu)\mathbf{u}_r(\mu)$. This approach can be linked to the one from [141], which consists in approximating the error by projections of the (unpreconditioned) residual onto approximate solutions of the dual problems $\mathbf{A}(\mu)^{\mathrm{H}}\mathbf{y}_i(\mu) = \mathbf{z}_i$ with random right-hand sides. The difference is that in [141] the authors proposed to tackle the random dual problems separately with RB methods, while we here consider a monolithic approach, approximating solutions by $\mathbf{y}_i(\mu) \approx \mathbf{P}(\mu)^{\mathrm{H}}\mathbf{z}_i$, where $\mathbf{P}(\mu)$ is a preconditioner constructed, for instance, by an interpolation of the operator's inverse based on minimization of an error indicator. Our method has several important advantages over the one in [141]. First, our efficient error certification procedure with the multi-purpose error indicator does not rely on the assumption that the error of the solution(s) of the random dual problem(s) is uniformly small on $\mathcal{P}$ as in [141]. Furthermore, we propose an alternative, more robust approach for error estimation and certification without requiring $\mathbf{y}_i(\mu)$ (or $\mathbf{A}(\mu)^{-1}$) to be well-approximated by low-dimensional spaces. This approach relies on the introduction of an efficiently computable upper bound of a norm of $(\mathbf{B}(\mu) - \mathbf{R}_U)(\mathbf{u}(\mu) - \mathbf{u}_r(\mu))$. A preconditioner for sharp error estimation can be constructed by minimization of this upper bound, without the requirement that $\mathbf{B}(\mu)$ has a small condition number. And finally, in contrast to [141] our methodology yields guarantees of success not only for finite parameter sets $\mathcal{P}$, which can be of particular interest for adaptive algorithms.

4.1.2 Preliminaries

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. The solution space is identified with $U := \mathbb{K}^n$. This space is equipped with the inner product

$$\langle \cdot, \cdot \rangle_U := \langle \mathbf{R}_U \cdot, \cdot \rangle_2,$$

where $\langle \cdot, \cdot \rangle_2$ is the Euclidean (or Hermitian) inner product on $\mathbb{K}^n$ and $\mathbf{R}_U \in \mathbb{K}^{n \times n}$ is some self-adjoint (symmetric if $\mathbb{K} = \mathbb{R}$ and Hermitian if $\mathbb{K} = \mathbb{C}$) positive definite matrix. The dual of U is identified with $U' := \mathbb{K}^n$ and is endowed with the canonical (dual) norm

$$\| \cdot \|_{U'} = \max_{\mathbf{w} \in U} \frac{\langle \cdot, \mathbf{w} \rangle_2}{\|\mathbf{w}\|_U}.$$

This norm is associated with the inner product $\langle \cdot, \cdot \rangle_{U'} := \langle \cdot, \mathbf{R}_U{}^{-1} \cdot \rangle_2$. The solution vector $\mathbf{u}(\mu)$ is seen as an element from U, the matrices $\mathbf{A}(\mu)$ and $\mathbf{R}_U$ are seen as operators from U to U', and $\mathbf{b}(\mu)$ is seen as an element from U'. The parameter set $\mathcal{P}$ can be a subset of $\mathbb{K}^e$ or a subset of an infinite dimensional space such as a function space. See Chapter 1 for more details on the meaning of this semi-discrete setting for numerical methods for PDEs. For problems described simply by algebraic equations the notions of solution spaces and dual spaces can be disregarded.

For finite-dimensional (Hilbert) spaces V and W identified with a Euclidean or a Hermitian space, we denote by $HS(V,W)$ the space of matrices representing operators from V to W. Assuming that V and W are equipped with inner products $\langle \cdot, \cdot \rangle_V$ and $\langle \cdot, \cdot \rangle_W$, respectively, we endow $HS(V,W)$ with the Hilbert-Schmidt inner product

$$\langle \mathbf{X}, \mathbf{Y} \rangle_{HS(V,W)} := \sum_{i=1}^{\dim V} \langle \mathbf{X}\mathbf{v}_i, \mathbf{Y}\mathbf{v}_i \rangle_W,$$

where $\mathbf{X}, \mathbf{Y} : V \to W$ and $\{\mathbf{v}_i : 1 \leq i \leq \dim V\}$ with an orthonormal basis for V. Below we particularize the above setting to specific choices of V and W.

For $V = \ell_2(\mathbb{K}^r)$ and $W = \ell_2(\mathbb{K}^k)$, $HS(V,W)$ is identified with the space of matrices $\mathbb{K}^{k \times r}$ equipped with the Frobenius inner product $\langle \cdot, \cdot \rangle_{HS(\ell_2, \ell_2)} = \langle \cdot, \cdot \rangle_F$.

For $V = \ell_2(\mathbb{K}^r)$ and $W = U$ or $W = U'$, $HS(V,W)$ is identified with the space of matrices $\mathbb{K}^{n \times r}$ equipped with the inner products

$$\langle \cdot, \cdot \rangle_{HS(\ell_2, U)} = \langle \mathbf{R}_U \cdot, \cdot \rangle_F, \text{ or } \langle \cdot, \cdot \rangle_{HS(\ell_2, U')} = \langle \cdot, \mathbf{R}_U^{-1} \cdot \rangle_F,$$

respectively.

Furthermore, $HS(U',U)$ and $HS(U,U')$ are identified with $\mathbb{K}^{n\times n}$. These spaces are seen as spaces of linear operators from U' to U and from U to U', respectively, and are endowed with inner products

$$\langle \cdot,\cdot\rangle_{HS(U',U)} := \langle \mathbf{R}_U \cdot \mathbf{R}_U, \cdot\rangle_F \text{ and } \langle \cdot,\cdot\rangle_{HS(U,U')} := \langle \cdot, \mathbf{R}_U^{-1} \cdot \mathbf{R}_U^{-1}\rangle_F. \tag{4.3}$$

We also let $\|\cdot\|_{HS(\ell_2,U)}, \|\cdot\|_{HS(\ell_2,U')}, \|\cdot\|_{HS(U',U)}$ and $\|\cdot\|_{HS(U,U')}$ be the associated norms.

4.2 Characterization of the quality of a preconditioner

In this section we derive efficiently computable estimates that characterize the quality of a preconditioner. They essentially represent some discrepancy between $\mathbf{P}(\mu)$ and $\mathbf{A}(\mu)^{-1}$. Different error indicators shall be considered depending on the objectives. We also discuss a construction of a preconditioner as a linear combination of some basis matrices.

Further, all considerations are for a fixed parameter value $\mu \in \mathcal{P}$, unless specified otherwise. For clarity of the presentation, the dependencies on μ are dropped out from the equations.

4.2.1 Multi-purpose context

Here we consider the preconditioned system of equations (4.2) and provide an error indicator that characterizes the performance of the preconditioner for projection-based methods such as (possibly adaptive) Galerkin methods, Krylov methods (with or without recycling), RB methods, etc.

The matrix $\mathbf{B} := \mathbf{R}_U\mathbf{P}\mathbf{A}$ in (4.2) can be seen as a linear operator from U to U'. The minimal and the maximal singular values (or inf-sup constant and operator norm) of $\mathbf{B}$ can be defined as follows

$$\alpha(\mathbf{B}) := \min_{\mathbf{v}\in U\backslash\{\mathbf{0}\}} \frac{\|\mathbf{B}\mathbf{v}\|_{U'}}{\|\mathbf{v}\|_U}, \tag{4.4a}$$

$$\beta(\mathbf{B}) := \max_{\mathbf{v}\in U\backslash\{\mathbf{0}\}} \frac{\|\mathbf{B}\mathbf{v}\|_{U'}}{\|\mathbf{v}\|_U}, \tag{4.4b}$$

and the condition number $\kappa(\mathbf{B}) := \frac{\beta(\mathbf{B})}{\alpha(\mathbf{B})}$. The performance of a projection-based method and a residual-based error estimator usually depends on the condition number. A smaller condition number yields better quasi-optimality constants.

The condition number of $\mathbf{B}$ can be characterized by the distance between $\mathbf{B}$ and $\mathbf{R}_U$ measured with the operator norm, i.e., by $\beta(\mathbf{R}_U - \mathbf{B})$. More specifically, it directly follows from the definitions of the minimal and maximal singular values that

$$1-\beta(\mathbf{R}_U-\mathbf{B}) \leq \alpha(\mathbf{B}) \leq \beta(\mathbf{B}) \leq 1+\beta(\mathbf{R}_U-\mathbf{B}). \tag{4.5}$$

The computation (and minimization) of $\beta(\mathbf{B}-\mathbf{R}_U)$ for multiple operators $\mathbf{B}$ can be an unfeasible task. Therefore the condition number of $\mathbf{B}$ shall be approximated with a computable upper bound of $\beta(\mathbf{B}-\mathbf{R}_U)$.

Proposition 4.2.1. *For an operator $\mathbf{C}: U \to U'$ and a vector $\mathbf{v} \in U$, it holds*

$$\|\mathbf{C}\mathbf{v}\|_{U'} \leq \|\mathbf{C}\|_{HS(U,U')}\|\mathbf{v}\|_U. \tag{4.6}$$

Proof. See appendix. □

From Proposition 4.2.1 it follows that $\|\mathbf{R}_U-\mathbf{B}\|_{HS(U,U')}$ is an upper bound of $\beta(\mathbf{B}-\mathbf{R}_U)$, which implies the first main result of this chapter:

Define the following error indicator

$$\boxed{\Delta_{U,U} = \|\mathbf{R}_U-\mathbf{B}\|_{HS(U,U')}.} \tag{4.7}$$

If $\Delta_{U,U} < 1$, then

$$\kappa(\mathbf{B}) \leq \frac{1+\Delta_{U,U}}{1-\Delta_{U,U}}. \tag{4.8}$$

Therefore a good performance of a projection-based method can be guaranteed if $\Delta_{U,U}$ is sufficiently small.

In practice, the condition $\Delta_{U,U} < 1$, which is required for the bound (4.8) to hold, is very hard to reach. Our empirical studies, however, suggest that the operators which come from real applications have a small condition number also when $\Delta_{U,U}$ is small but larger than one.

In general, a good effectivity of $\|\cdot\|_{HS(U,U')}$ as an estimator of the operator norm $\beta(\cdot)$ may not be guaranteed. In some situations, a large overestimation (up to a factor of $n^{1/2}$) happens. This issue can be particularly dramatic for numerical methods for PDEs, where each discrete operator $\mathbf{C}$ (e.g., $\mathbf{C}=\mathbf{R}_U-\mathbf{B}$) represents a finite-dimensional approximation of some differential operator C. The operator norm of C is an upper bound of $\beta(\mathbf{C})$ regardless of the chosen discretization. The norm $\|\mathbf{C}\|_{HS(U,U')}$ is an approximation of the Hilbert-Schmidt norm of C, which can be infinite (if C is not a Hilbert-Schmidt operator). Therefore, even if C has a small operator norm (implying that $\beta(\mathbf{C})$ is also small), $\|\mathbf{C}\|_{HS(U,U')}$ can be highly sensitive to the discretization and go to infinity with the number of degrees of freedom. This implies a possible failure of $\Delta_{U,U}$ for characterizing the quality of the preconditioned operator. This problem can be circumvented for the projection-based MOR context, where the solution is approximated with a moderately large space, or for the residual-based error estimation.

4.2.2 Galerkin projection

Further, we consider the projection-based MOR context where the solution $\mathbf{u}$ in (4.2) is approximated by the Galerkin projection $\mathbf{u}_r$ onto a subspace U_r. The subspace U_r can be constructed with a greedy algorithm for RB method or low-rank approximation of the matrix of solution samples (snapshots) for POD. The basis vectors for U_r can also be chosen a priori by exploiting the structure of the problem. In the context of numerical methods for PDEs, such basis vectors can be obtained by computing the coordinates of the basis functions (associated, for instance, with an approximation on a coarse grid) on the space of functions identified with U.

For given $W \subseteq U$, let $\mathrm{P}_W : U \to W$ denote the orthogonal projection on W with respect to $\|\cdot\|_U$, i.e.,

$$\forall \mathbf{x} \in U, \ \mathrm{P}_W \mathbf{x} = \arg\min_{\mathbf{w}\in W} \|\mathbf{x}-\mathbf{w}\|_U. \tag{4.9}$$

The Galerkin orthogonality condition can be stated as follows

$$\langle \mathbf{B}(\mathbf{u}-\mathbf{u}_r), \mathbf{w}\rangle_2 = 0, \ \forall \mathbf{w} \in U_r, \tag{4.10}$$

or, equivalently [12],

$$\|\mathbf{B}(\mathbf{u}-\mathbf{u}_r)\|_{U_r'} = 0, \tag{4.11}$$

where $\|\cdot\|_{U_r'} := \|\mathrm{P}_{U_r}\mathbf{R}_U^{-1}\cdot\|_U$.

Next we use the following lemma to provide conditions for controlling the accuracy of $\mathbf{u}_r$ summarized in Proposition 4.2.3.

Lemma 4.2.2. *Let $\mathbf{u}_r$ satisfy (4.11). Then*

$$\|\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u}\|_U \leq \|[\mathbf{R}_U - \mathbf{B}](\mathbf{u}-\mathrm{P}_{U_r}\mathbf{u})\|_{U_r'} + \|[\mathbf{R}_U - \mathbf{B}](\mathbf{u}_r-\mathrm{P}_{U_r}\mathbf{u})\|_{U_r'}. \tag{4.12}$$

Proof. See appendix. □

Proposition 4.2.3. *Define*

$$\beta_r(\mathbf{R}_U - \mathbf{B}) := \max_{\mathbf{v}\in U_r\backslash\{\mathbf{0}\}} \frac{\|[\mathbf{R}_U-\mathbf{B}]\mathbf{v}\|_{U_r'}}{\|\mathbf{v}\|_U} \tag{4.13a}$$

$$\bar{\beta}_r(\mathbf{R}_U - \mathbf{B}) := \max_{\mathbf{v}\in U\backslash\{\mathbf{0}\}} \frac{\|[\mathbf{R}_U-\mathbf{B}]\mathbf{v}\|_{U_r'}}{\|\mathbf{v}\|_U}, \tag{4.13b}$$

If $\beta_r(\mathbf{R}_U - \mathbf{B}) < 1$, then the solution $\mathbf{u}_r$ to (4.11) is unique and

$$\|\mathbf{u}-\mathbf{u}_r\|_U \leq \left(1 + \frac{\bar{\beta}_r(\mathbf{R}_U-\mathbf{B})}{1-\beta_r(\mathbf{R}_U-\mathbf{B})}\right)\|\mathbf{u}-\mathrm{P}_{U_r}\mathbf{u}\|_U. \tag{4.14}$$

Proof. See appendix. □

According to Proposition 4.2.3, the quasi-optimality of $\mathbf{u}_r$ can be guaranteed by making sure that the coefficients $\beta_r(\mathbf{R}_U - \mathbf{B})$ and $\bar{\beta}_r(\mathbf{R}_U - \mathbf{B})$ are small enough. We observe that

$$\beta_r(\mathbf{R}_U - \mathbf{B}) \leq \bar{\beta}_r(\mathbf{R}_U - \mathbf{B}) \leq \beta(\mathbf{R}_U - \mathbf{B}) \leq \Delta_{U,U}. \tag{4.15}$$

Moreover, for $U_r = U$ we clearly have $\beta_r(\mathbf{R}_U - \mathbf{B}) = \bar{\beta}_r(\mathbf{R}_U - \mathbf{B}) = \beta(\mathbf{R}_U - \mathbf{B})$. The relation (4.15) and Proposition 4.2.3 imply a characterization of the Galerkin projection with the multi-purpose indicator $\Delta_{U,U}$:

$$\|\mathbf{u} - \mathbf{u}_r\|_U \leq (1 + \frac{\Delta_{U,U}}{1 - \Delta_{U,U}})\|\mathbf{u} - \mathrm{P}_{U_r}\mathbf{u}\|_U. \tag{4.16}$$

Furthermore, the quality of the preconditioner can be better characterized by taking into account that the coefficients $\beta_r(\mathbf{R}_U - \mathbf{B})$ and $\bar{\beta}_r(\mathbf{R}_U - \mathbf{B})$ represent a discrepancy between $\mathbf{B}$ and $\mathbf{R}_U$ measured with the semi-norm $\|\cdot\|_{U_r'}$, which is the restriction of $\|\cdot\|_{U'}$ onto a low-dimensional space. Consequently, the multi-purpose criteria for characterizing the quality of the Galerkin projection can be improved by restriction of $\mathbf{B}$ to U_r. Such considerations lead to error indicators Δ_{U_r,U_r} and $\Delta_{U_r,U}$ defined below.

Proposition 4.2.4. *Define*

$$\boxed{\Delta_{U_r,U_r} := \|\mathbf{U}_r^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{U}_r\|_F} \tag{4.17a}$$

and

$$\boxed{\Delta_{U_r,U} := \|[\mathbf{R}_U - \mathbf{B}]^{\mathrm{H}}\mathbf{U}_r\|_{HS(\ell_2,U')},} \tag{4.17b}$$

where $\mathbf{U}_r : \mathbb{K}^r \to U$ is a matrix whose columns form a basis for U_r. The following relations hold:

$$\frac{1}{\sigma_1^2\sqrt{r}}\Delta_{U_r,U_r} \leq \beta_r(\mathbf{R}_U - \mathbf{B}) \leq \frac{1}{\sigma_r^2}\Delta_{U_r,U_r} \tag{4.18a}$$

$$\frac{1}{\sigma_1\sqrt{r}}\Delta_{U_r,U} \leq \bar{\beta}_r(\mathbf{R}_U - \mathbf{B}) \leq \frac{1}{\sigma_r}\Delta_{U_r,U}, \tag{4.18b}$$

where

$$\sigma_r := \min_{\mathbf{a}\in\mathbb{K}^r/\{\mathbf{0}\}} \frac{\|\mathbf{U}_r\mathbf{a}\|_U}{\|\mathbf{a}\|_2} \text{ and } \sigma_1 := \max_{\mathbf{a}\in\mathbb{K}^r/\{\mathbf{0}\}} \frac{\|\mathbf{U}_r\mathbf{a}\|_U}{\|\mathbf{a}\|_2}$$

are the minimal and the maximal singular values of $\mathbf{U}_r$ with respect to the $\|\cdot\|_U$-norm.

Proof. See appendix. □

Clearly, the bounds in Proposition 4.2.4 are tighter when the columns of $\mathbf{U}_r$ are unit-orthogonal vectors with respect to $\langle \cdot, \cdot \rangle_U$.

Corollary 4.2.5. *Let Δ_{U_r,U_r} and $\Delta_{U_r,U}$ be the error indicators from Proposition 4.2.4. If the columns of $\mathbf{U}_r$ are unit-orthogonal vectors with respect to $\langle \cdot, \cdot \rangle_U$, then*

$$\frac{1}{\sqrt{r}} \Delta_{U_r,U_r} \leq \beta_r(\mathbf{R}_U - \mathbf{B}) \leq \Delta_{U_r,U_r}, \tag{4.19}$$

$$\frac{1}{\sqrt{r}} \Delta_{U_r,U} \leq \bar{\beta}_r(\mathbf{R}_U - \mathbf{B}) \leq \Delta_{U_r,U}. \tag{4.20}$$

Furthermore, it is easy to see that if $\mathbf{U}_r$ has unit-orthogonal columns with respect to $\langle \cdot, \cdot \rangle_U$, then

$$\Delta_{U_r,U_r} \leq \Delta_{U_r,U} \leq \Delta_{U,U}.$$

This fact implies that in this case the quasi-optimality constants obtained with Δ_{U_r,U_r} and $\Delta_{U_r,U}$ shall always be better than the ones obtained with $\Delta_{U,U}$. Note that if $U_r = U$ then $\Delta_{U_r,U_r} = \Delta_{U_r,U} = \Delta_{U,U}$. Unlike the multi-purpose context, here the effectiveness of Δ_{U_r,U_r} and $\Delta_{U_r,U}$ as estimators of $\beta_r(\mathbf{R}_U - \mathbf{B})$ and $\bar{\beta}_r(\mathbf{R}_U - \mathbf{B})$ is guaranteed. For PDEs this implies a robust characterization of the quality of the preconditioned operator regardless of the discretization.

For some problems and computational architectures (e.g., when the basis vectors may not be efficiently maintained) the orthogonalization of the basis with respect to $\langle \cdot, \cdot \rangle_U$ can be very expensive. In these cases one should use $\mathbf{U}_r$ with columns which are only approximately orthogonal. For PDEs such a matrix can be obtained with domain decomposition (i.e., by local orthogonalization). Another possibility (for not too large r) is the approximate orthogonalization of the columns of $\mathbf{U}_r$ with a random sketching technique (see Section 4.3 for details). Note that Δ_{U_r,U_r} is more sensitive to the condition number of $\mathbf{U}_r$ than $\Delta_{U_r,U}$. Consequently, for $\mathbf{U}_r$ with moderate or high condition number it can be more pertinent to characterize the Galerkin projection with the error indicator $\Delta_{U_r,U}$ only, and use the fact that $\beta_r(\mathbf{R}_U - \mathbf{B}) \leq \bar{\beta}_r(\mathbf{R}_U - \mathbf{B})$.

Let us now summarize the results of Propositions 4.2.3 and 4.2.4 in a practical form.

Consider error indicators $\Delta_{U_r,U_r}, \Delta_{U_r,U}$ defined in (4.17) and $\Delta_{U,U}$ defined in (4.7). If $\min\{\Delta_{U,U}, \sigma_r^{-1}\Delta_{U_r,U}, \sigma_r^{-2}\Delta_{U_r,U_r}\} < 1$*, then the solution $\mathbf{u}_r$ to (4.11) is such that*

$$\|\mathbf{u} - \mathbf{u}_r\|_U \leq \left(1 + \frac{\min\{\Delta_{U,U}, \sigma_r^{-1}\Delta_{U_r,U}\}}{1 - \min\{\Delta_{U,U}, \sigma_r^{-1}\Delta_{U_r,U}, \sigma_r^{-2}\Delta_{U_r,U}\}}\right) \|\mathbf{u} - \mathrm{P}_{U_r}\mathbf{u}\|_U, \tag{4.21}$$

where σ_r is the minimal singular value of $\mathbf{U}_r$ with respect to the $\|\cdot\|_U$-norm. Therefore to ensure high quality of the Galerkin projection one can seek a preconditioner that minimizes

$$\left(\gamma_1 \Delta_{U,U}^2 + \gamma_2 \Delta_{U_r,U}^2 + \gamma_3 \Delta_{U_r,U_r}^2\right)^{1/2} \tag{4.22}$$

with (possibly zero) weights γ_1, γ_2 and γ_3 picked depending on the problem.

Again, the condition $\Delta_{U_r,U_r} < \sigma_r^2$ (or $\Delta_{U_r,U} < \sigma_r$) can require too expensive preconditioners and may not be attained for some problems. Without it we do not have any a priori guarantee of high quality of the Galerkin projection. On the other hand, our experimental observations revealed that, in practice, minimizing Δ_{U_r,U_r} and $\Delta_{U_r,U}$ yields reliable preconditioners even when $\Delta_{U_r,U_r} \geq \sigma_r^2$.

4.2.3 Error certification

Let $\mathbf{u}_r$ be an approximation of $\mathbf{u}$. The vector $\mathbf{u}_r$ could be obtained, for example, by projecting $\mathbf{u}$ on an approximation space. Next we address the question of estimating and bounding the error $\|\mathbf{u}-\mathbf{u}_r\|_U$. The standard way is the certification of the error with the residual norm:

$$\|\mathbf{u}-\mathbf{u}_r\|_U \leq \frac{1}{\eta}\|\mathbf{r}(\mathbf{u}_r)\|_{U'}, \tag{4.23}$$

where $\mathbf{r}(\mathbf{u}_r) = \mathbf{b} - \mathbf{A}\mathbf{u}_r$ and η is a computable lower bound of the smallest singular value of $\mathbf{A}$ (the operator norm of $\mathbf{A}^{-1}$). For ill-conditioned operators $\mathbf{A}$, the accuracy of such error estimator is very poor. A straightforward approach to overcome this issue is to replace $\mathbf{A}$ in (4.23) by the preconditioned operator $\mathbf{B}$:

$$\|\mathbf{u}-\mathbf{u}_r\|_U \leq \frac{1}{\eta^*}\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'}, \tag{4.24}$$

with $\mathbf{r}^*(\mathbf{u}_r) = \mathbf{f} - \mathbf{B}\mathbf{u}_r$ and η^* being a lower bound of the smallest singular value of $\mathbf{B}$. The coefficients η and η^* can be obtained theoretically or with the Successive Constraint Method [93]. The above approach can be intractable, since the computation of η^* with classical procedures can be much more expensive than the computation of η. A more efficient certification of the error can be obtained with a multi-purpose error indicator $\Delta_{U,U}$, as proposed in Proposition 4.2.6.

Proposition 4.2.6. *If $\Delta_{U,U} < 1$, then*

$$\frac{1}{1+\Delta_{U,U}}\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'} \leq \|\mathbf{u}-\mathbf{u}_r\|_U \leq \frac{1}{1-\Delta_{U,U}}\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'}. \tag{4.25}$$

Proof. See appendix. □

The advantage of certification with $\Delta_{U,U}$ is that it does not require the usage of expensive methods to estimate the operator's minimal singular value. However, the effectivenesses of the certification with $\Delta_{U,U}$ can be poor (for instance, for PDEs context with non Hilbert-Schmidt operators). Furthermore, both certifications with (4.24) and (4.25) require $\mathbf{B}$ to have a moderate minimal singular value. However, the residual associated with $\mathbf{B}$ can provide a good estimation of the exact error even when the minimal singular value of $\mathbf{B}$ is very small and, possibly, equal to zero (when $\mathbf{B}$ is singular). For instance, imagine a situation when we are able to construct a preconditioner such that $\mathbf{B}$ is close to $\mathbf{R}_U$ when restricted to a specific set of vectors including $\mathbf{u}-\mathbf{u}_r$, but that can highly deviate from $\mathbf{R}_U$ when applied to other vectors. In this case a failure of error certification with (4.24) can be detected. Below, we provide a more robust error certification.

The accuracy of $\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'}$ as an error estimator can be certified by the quantity $\|(\mathbf{R}_U-\mathbf{B})(\mathbf{u}-\mathbf{u}_r)\|_{U'}$, which can be efficiently bounded above by a norm of $(\mathbf{R}_U-\mathbf{B})(\mathbf{u}-\mathbf{u}_r)$ mapped through $\mathbf{A}\mathbf{R}_U^{-1}$. These considerations lead to Proposition 4.2.7.

Proposition 4.2.7. *Define the following error indicator*

$$\boxed{\Delta_{\mathrm{e}} := \|\mathbf{d}(\mathbf{u}_r)\|_{U'},} \tag{4.26}$$

where $\mathbf{d}(\mathbf{u}_r) := [\mathbf{I}-\mathbf{A}\mathbf{P}]\mathbf{r}(\mathbf{u}_r)$. *Then, we have*

$$\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'} - \Delta_{\mathrm{e}}/\eta \leq \|\mathbf{u}-\mathbf{u}_r\|_U \leq \|\mathbf{r}^*(\mathbf{u}_r)\|_{U'} + \Delta_{\mathrm{e}}/\eta. \tag{4.27}$$

Proof. See appendix. □

A great advantage of certification of the error with Proposition 4.2.7 is that such a certification no longer requires $\mathbf{B}$ to have a moderate minimal singular value as in (4.24) and (4.25). The only requirement is that the preconditioner is such that $\mathbf{B}$ is close to $\mathbf{R}_U$ when applied to the vector $\mathbf{u}-\mathbf{u}_r$.

Propositions 4.2.6 and 4.2.7 constitute the main result of this section, which is concluded below.

The error of the approximate solution $\mathbf{u}_r$ *can be estimated by a norm of the preconditioned residual* $\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'}$. *The quality of such an estimation can either be certified using a general error indicator* $\Delta_{U,U}$ *with relation (provided* $\Delta_{U,U}<1$*)*

$$\frac{1}{1+\Delta_{U,U}}\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'} \leq \|\mathbf{u}-\mathbf{u}_r\|_U \leq \frac{1}{1-\Delta_{U,U}}\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'},$$

or using the error indicator Δ_{e} *defined by* (4.26) *with relation*

$$\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'} - \Delta_{\mathrm{e}}/\eta \leq \|\mathbf{u}-\mathbf{u}_r\|_U \leq \|\mathbf{r}^*(\mathbf{u}_r)\|_{U'} + \Delta_{\mathrm{e}}/\eta,$$

where η is a computable lower bound of the minimal singular value of $\mathbf{A}$. It follows that if $\Delta_{U,U}$ or Δ_{e} is small, then $\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'}$ provides sharp estimation of the exact error $\|\mathbf{u}-\mathbf{u}_r\|_U$. The certification with Δ_{e} is more robust than with $\Delta_{U,U}$ but is less efficient since it requires computation of η. Yet, it can be enough to consider for η a rough estimation of the minimal singular value since this coefficient is scaled by Δ_{e}, which may be orders of magnitude less than the error.

4.2.4 Construction of a preconditioner

A parameter-dependent preconditioner $\mathbf{P}(\mu)$ can be obtained by a projection of $\mathbf{A}(\mu)^{-1}$ onto a linear span of some basis matrices $\{\mathbf{Y}_i\}_{i=1}^p$, i.e.,

$$\mathbf{P}(\mu) = \sum_{i=1}^{p} \lambda_i(\mu)\mathbf{Y}_i, \tag{4.28}$$

with coefficients $\lambda_i(\mu)$ computed online by solving a small least-squares problem for a minimization of one of the above error indicators (or rather their efficient approximations given in Section 4.3).

The basis matrices $\mathbf{Y}_i$ can be taken as $\mathbf{A}(\mu^i)^{-1}$ at some interpolation points $\mu^i \in \mathcal{P}$. The set of interpolation points can be obtained simply by random sampling in $\mathcal{P}$. Another way is an iterative greedy selection, at each iteration enriching the set of interpolation points by the parameter value where the error indicator was the largest. For methods where the approximation space U_r is constructed from snapshots $\mathbf{u}(\hat{\mu}^j)$ such as the RB method or POD, the interpolation points can be selected among the parameters $\hat{\mu}^j$, providing recycling of the computations, since each snapshot (typically) requires computation of the implicit inverse (e.g., factorization) of the operator. Finally, for the reduced basis methods where U_r is constructed with a greedy algorithm based on Petrov-Galerkin projection, it can be useful to consider the same interpolation points for the construction of U_r and $\{\mathbf{Y}_i\}_{i=1}^p$. In this case the error indicator for the greedy selection of an interpolation point should be defined as a (weighted) average of the error indicator characterizing the quality of U_r (e.g., an upper bound of the error of the Galerkin projection) and the error indicator characterizing the quality of $\{\mathbf{Y}_i\}_{i=1}^p$ (e.g., one of the error indicators from above). Other strategies for finding the parameter values μ^i can be found in [160].

4.3 Randomized error indicators

In this section we propose a probabilistic approach for drastic reduction of the computational cost and improvement of the numerical stability associated with the computation (or minimization) of the error indicators from Section 4.2. For this we adapt the framework from Chapter 2.

Recall that every error indicator from Section 4.2 is given as a norm of a certain residual matrix (or vector). These norms can be estimated with ℓ_2-norms of the images (so-called sketches) of the residual matrices (or vectors) through a carefully chosen random linear map to a small Euclidean (or Hermitian) space. Such random maps are here constructed using random sketching matrices, so-called $\ell_2 \to \ell_2$ oblivious subspace embeddings (see [12, 157]). They include the rescaled Gaussian matrices, the rescaled Rademacher matrices, Subsampled Randomized Hadamard Transform (SRHT), the Subsampled Randomized Fourier Transform (SRFT) and others.

Let $\mathbf{\Gamma}$, $\mathbf{\Omega}$ and $\mathbf{\Sigma}$ be $\ell_2 \to \ell_2$ oblivious subspace embeddings with sufficiently large numbers of rows, which will be used for the estimation of the error indicators. A detailed analysis of the sizes of random matrices needed to guarantee the given accuracy with probability of failure less than δ is presented in Section 4.4.

Multi-purpose

The error indicator $\Delta_{U,U}(\mu)$ defined by (4.7) can be approximated by

$$\boxed{\Delta^{\mathrm{sk}}_{U,U}(\mu) := \|\mathbf{\Theta}(\mathbf{R}_U^{-1}[\mathbf{R}_U - \mathbf{B}(\mu)]\mathbf{R}_U^{-1})\|_2,} \tag{4.29}$$

where $\mathbf{\Theta}(\cdot)$ is a linear map from the space $HS(U',U)$ (i.e., the space of matrices which are seen as operators from U' to U) to $\mathbb{K}^k$, with $k \ll n$, equipped with ℓ_2-inner product. The map $\mathbf{\Theta}(\cdot)$ can be seen as an oblivious $HS(U',U) \to \ell_2$ subspace embedding.

The map $\mathbf{\Theta}(\cdot)$ is chosen such that $\langle \mathbf{\Theta}(\cdot), \mathbf{\Theta}(\cdot)\rangle_2$ with probability at least $1-\delta$ approximates well $\langle \cdot,\cdot\rangle_{HS(U',U)}$ over an arbitrary subspace of $HS(U',U)$ of small or moderate dimension m. Note that for a fixed parameter value, $\mathbf{B}(\mu)$ belongs to a p-dimensional space, if $\mathbf{P}(\mu)$ has form (4.28). As is indicated in Section 4.4, in this case (if we use $m \geq p+1$) the minimization of $\Delta^{\mathrm{sk}}_{U,U}(\mu)$ over $\lambda_1(\mu),\ldots,\lambda_p(\mu)$ will provide a quasi-optimal minimizer of $\Delta_{U,U}(\mu)$ with high probability.

Let $\mathbf{Q}$ be a matrix such that $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{R}_U$. [2] This matrix can be obtained with a Cholesky factorization or a more efficient approach proposed in Remark 2.2.7. The map $\mathbf{\Theta}(\cdot)$ is obtained by taking

$$\mathbf{\Theta}(\mathbf{X}) := \mathbf{\Gamma}\ \mathtt{vec}(\mathbf{\Omega}\mathbf{Q}\mathbf{X}\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}}),\ \mathbf{X}: U' \to U \tag{4.30}$$

where the operation $\mathtt{vec}(\cdot)$ reshapes (say, column-wise) a matrix to a vector. From Section 4.4 it follows that in (4.30), the random matrices $\mathbf{\Gamma}$, $\mathbf{\Omega}$ and $\mathbf{\Sigma}$ can be chosen as rescaled Gaussian matrices with $\mathcal{O}(m+\log(n)+\log(1/\delta))$ rows. The theoretical bounds for the sizes of SRHT matrices can be higher by logarithmic factors in n,m

[2] The matrix $\mathbf{Q}$ (respectively $\mathbf{Q}^{\mathrm{H}}$) is interpreted as a map from U to ℓ_2 (respectively from ℓ_2 to U').

and $1/\delta$, although in practice SRHT and Gaussian matrices show similar performances. Note that this fact holds not only for the multi-purpose context but also for the Galerkin projection and the error estimation contexts.

Galerkin projection

Let us first comment the approximate orthogonalization of the basis vectors for U_r, which is necessary for Proposition 4.2.4. This can be efficiently done with a random sketching technique (for not too large r). It follows (see Section 4.4) that the orthogonalization of $\mathbf{U}_r$ with respect to the sketched inner product $\langle \mathbf{\Omega}\mathbf{Q}\cdot, \mathbf{\Omega}\mathbf{Q}\cdot\rangle_2$ with probability at least $1-\delta$ yields a matrix with singular values close to 1, if $\mathbf{\Omega}$ is a Gaussian matrix (or SRHT, in practice) with $\mathcal{O}(r+\log(1/\delta))$ rows.

The approximation of the error indicators $\Delta_{U_r,U_r}(\mu)$ and $\Delta_{U_r,U}(\mu)$ defined by (4.17) are given by

$$\boxed{\Delta^{\text{sk}}_{U_r,U_r}(\mu) := \|\mathbf{\Sigma}\mathtt{vec}({\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}(\mu)]\mathbf{U}_r)\|_2} \tag{4.31a}$$

and

$$\boxed{\Delta^{\text{sk}}_{U_r,U}(\mu) := \|\mathbf{\Theta}(\mathbf{R}_U^{-1}[\mathbf{R}_U - \mathbf{B}(\mu)]^{\mathrm{H}}\mathbf{U}_r)\|_2,} \tag{4.31b}$$

where

$$\mathbf{\Theta}(\mathbf{X}) = \mathbf{\Gamma}\mathtt{vec}(\mathbf{\Omega}\mathbf{Q}\mathbf{X}),\ \mathbf{X}: \mathbb{K}^r \to U,$$

is an oblivious $HS(\ell_2, U) \to \ell_2$ embedding of subspaces of matrices. Similarly to the multi-purpose context, the random matrices $\mathbf{\Gamma}$, $\mathbf{\Omega}$ and $\mathbf{\Sigma}$ with sufficiently large numbers of rows are used so that, with probability of failure less than δ, $\langle \mathbf{\Sigma}\text{vec}(\cdot), \mathbf{\Sigma}\text{vec}(\cdot)\rangle_2$ approximates well $\langle \cdot,\cdot\rangle_F$ over an arbitrary m-dimensional subspace of $\mathbb{K}^{r\times r}$ and $\langle \mathbf{\Theta}(\cdot), \mathbf{\Theta}(\cdot)\rangle_2$ approximates well $\langle\cdot,\cdot\rangle_{HS(\ell_2,U)}$ over an arbitrary m-dimensional subspace of matrices with r column vectors interpreted as elements from U. To guarantee this property, it is sufficient to consider rescaled Gaussian matrices (or SRHT, in practice) with $\mathcal{O}(m + \log(r) + \log(1/\delta))$ rows (see Section 4.4 for more details).

Remark 4.3.1. *The offline computations required by the error indicators in* (4.31) *should have only a minor impact on the overall computational costs when the approximation space has a small dimension. In some situations, however, it can be useful to consider larger approximation spaces. Then the computational cost can be further reduced by replacing* $\mathbf{U}_r$ *with its sketch* $\mathbf{U}_r\tilde{\mathbf{\Omega}}^{\mathrm{H}}$, *where* $\tilde{\mathbf{\Omega}}$ *is a small* $\ell_2 \to \ell_2$ *oblivious subspace embedding. Furthermore, the offline precomputation of* $\Delta^{\text{sk}}_{U_r,U_r}(\mu)$ *requires two passes over* $\mathbf{U}_r$ *(or* $\mathbf{U}_r\tilde{\mathbf{\Omega}}^{\mathrm{H}}$*) and can be too costly when* $\mathbf{U}_r$ *(or* $\mathbf{U}_r\tilde{\mathbf{\Omega}}^{\mathrm{H}}$*) is distributed among multiple workstations or when it may not be efficiently stored. In such cases the characterization of the Galerkin projection can be performed with only one error indicator* $\Delta^{\text{sk}}_{U_r,U}(\mu)$ *and using* (4.21)*.*

Error estimation

For the approximation of $\Delta_e(\mu)$ we can use exactly the same procedure as the one in Chapter 2 for the estimation of the residual norm by choosing $\mathbf{d}(\mathbf{u}_r(\mu);\mu)$ as the residual vector. It follows that a good estimation of $\Delta_e(\mu)$ can be efficiently obtained by

$$\boxed{\Delta_{\mathrm{e}}^{\mathrm{sk}}(\mu) := \|\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{d}(\mathbf{u}_r(\mu);\mu)\|_2,} \tag{4.32}$$

where $\mathbf{\Theta} = \mathbf{\Omega}\mathbf{Q}$ is an oblivious $U \to \ell_2$ subspace embedding. Moreover, for efficiency the preconditioned residual norm $\|\mathbf{r}^*(\mathbf{u}_r(\mu);\mu)\|_{U'}$ in Proposition 4.2.7 should also be approximated with random sketching, by

$$\|\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{r}^*(\mathbf{u}_r(\mu);\mu)\|_2. \tag{4.33}$$

We consider the matrix $\mathbf{\Omega}$ with a sufficient number of rows so that $\langle \mathbf{\Theta}\cdot, \mathbf{\Theta}\cdot\rangle_2$ approximates well $\langle\cdot,\cdot\rangle_U$ over an arbitrary m-dimensional subspace of U with probability at least $1-\delta$. This property can be guaranteed to hold with probability at least $1-\delta$, if considering for $\mathbf{\Omega}$ a Gaussian matrix (or SRHT, in practice) with $\mathcal{O}(m+\log(1/\delta))$ rows (see Section 4.4). Note that if $\mathbf{P}(\mu)$ is of the form (4.28), then for a fixed parameter value, the residual $\mathbf{d}(\mathbf{u}_r(\mu);\mu)$ (and $\mathbf{r}^*(\mathbf{u}_r(\mu);\mu)$) belongs to a $p+1$-dimensional space, which particularly implies (if $m \geq p+1$) that the minimizer of $\Delta_{\mathrm{e}}^{\mathrm{sk}}(\mu)$ over $\lambda_1(\mu),\ldots,\lambda_p(\mu)$ is a quasi-optimal minimizer of $\Delta_{\mathrm{e}}(\mu)$.

Below, we summarize the main results of this section from the practical point of view.

The error indicators $\Delta_{U_r,U_r}(\mu)$, $\Delta_{U_r,U}(\mu)$, $\Delta_{U,U}(\mu)$ and the residual error estimator $\|\mathbf{r}^*(\mathbf{u}_r(\mu);\mu)\|_{U'}$ used in Section 4.2 can be respectively estimated by efficiently computable random estimators $\Delta^{\mathrm{sk}}_{U_r,U_r}(\mu)$, $\Delta^{\mathrm{sk}}_{U_r,U}(\mu)$, $\Delta^{\mathrm{sk}}_{U,U}(\mu)$ and $\|\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{r}^*(\mathbf{u}_r(\mu);\mu)\|_2$ which essentially are the ℓ_2-norms of the residual matrices or vectors randomly embedded in a small low-dimensional space. A detailed analysis of the sizes of random sketching matrices $\mathbf{\Gamma}$, $\mathbf{\Omega}$ and $\mathbf{\Sigma}$ needed to guarantee the given accuracy with high probability is presented in Section 4.4. For $\mathbf{P}(\mu)$ defined in (4.28), the sketched error indicators from above (or their quadratic weighted average) can be written in the following form

$$\|\mathbf{W}_p(\mu)\mathbf{a}_p(\mu) - \mathbf{h}(\mu)\|_2, \tag{4.34}$$

where $[\mathbf{a}_p(\mu)]_i = \lambda_i(\mu)$, $1 \leq i \leq p$. The $k \times p$ matrix $\mathbf{W}_p(\mu) = [\mathbf{w}_1(\mu),\ldots,\mathbf{w}_p(\mu)]$ and the vector $\mathbf{h}(\mu)$ represent the sketches of the corresponding large matrices (or vectors). For instance, for the multi-purpose context $\mathbf{w}_i(\mu) = \mathbf{\Theta}(\mathbf{Y}_i\mathbf{A}(\mu)\mathbf{R}_U^{-1})$, $1 \leq i \leq p$, and $\mathbf{h}(\mu) = \mathbf{\Theta}(\mathbf{R}_U^{-1})$. The minimization of (4.34) over $\mathbf{a}_p(\mu)$ can be efficiently and numerically stably performed online for each parameter value with a standard routine such as QR factorization. For this, the affine decompositions of $\mathbf{W}_p(\mu)$ and $\mathbf{h}(\mu)$ have to be precomputed in the offline stage and then used for the efficient assembling of (4.34) for each μ, with a cost independent (or logarithmically dependent) of the *full dimension n. The affine decompositions can be obtained from (given) affine decompositions of $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$ or with empirical interpolation method.*

The computational cost of the offline stage is dominated by two operations: the products of $\mathbf{Y}_i$ (and $\mathbf{R}_U^{-1}$, $\mathbf{Q}$) with multiple vectors and the computation of random projections of explicit matrices and vectors.

With a good choice of random projections $\mathbf{\Gamma}$, $\mathbf{\Omega}$ and $\mathbf{\Sigma}$, the offline computational cost associated with multiplications of these matrices by explicit matrices and vectors should have only a minor impact on the overall cost. Indeed, SRHT matrices have a specific structure allowing products with a low number of flops, while Gaussian matrices are very efficient in terms of scalability of computations for parallel architectures. Moreover, the random matrices can be generated, maintained or transfered (for the computations on multiple computational devices) with a negligible computational cost by using a seeded random number generator. For more details please see Chapters 1 and 2.

As was indicated in [160] the maintenance of the basis matrices for the preconditioner in explicit form can be intractable. In general, one should maintain and operate with $\mathbf{Y}_i$ (and $\mathbf{R}_U^{-1}$) in an efficient implicit form (e.g, obtained with LU or $\mathcal{H}$-factorization), which can be once precomputed and then used for efficient products of $\mathbf{Y}_i$ (and $\mathbf{R}_U^{-1}$) with multiple vectors. Furthermore, for the Galerkin projection and error estimation contexts the (possibly expensive) maintenance and operation with $\mathbf{Y}_i$ can be avoided thanks to the methodology from Chapter 2. As was indicated in Chapter 2, a reduced model can be accurately (with high probability) approximated from small random projections of the approximation space and the associated residuals:

$$\{\mathbf{\Theta}^*\mathbf{x} :\ \mathbf{x} \in U_r\} \text{ and } \{\mathbf{\Theta}^*\mathbf{R}_U^{-1}\mathbf{r}^*(\mathbf{x};\mu) = \mathbf{\Theta}^*\mathbf{R}_U^{-1}(\mathbf{f}(\mu) - \mathbf{B}(\mu)\mathbf{x}) :\ \mathbf{x} \in U_r\},$$

where U_r is a low-dimensional approximation space and $\mathbf{\Theta}^*$ is a small random matrix. Firstly, we see that rather than maintaining and operating with a basis matrix $\mathbf{Y}_i$ in the offline stage, we can precompute its random sketch $\mathbf{\Theta}^*\mathbf{Y}_i$ (along with $\mathbf{w}_i(\mu)$) and operate with the sketch, which can be far more efficient. Furthermore, if $\mathbf{U}_r$ is a matrix whose columns form a basis for U_r, a reduced model and the terms needed for estimation of the preconditioned residual norm can be efficiently evaluated from affine decompositions of small projections $\mathbf{V}_i^{\mathbf{\Theta}^*}(\mu) = \mathbf{\Theta}^*\mathbf{Y}_i\mathbf{A}(\mu)\mathbf{U}_r$, $\mathbf{f}_i^{\mathbf{\Theta}^*}(\mu) = \mathbf{\Theta}^*\mathbf{Y}_i\mathbf{b}(\mu)$ and $\mathbf{\Theta}^*\mathbf{U}_r$. For each $1 \leq i \leq p$ the precomputation of the affine decompositions of $\mathbf{V}_i^{\mathbf{\Theta}^*}(\mu)$, $\mathbf{f}_i^{\mathbf{\Theta}^*}(\mu)$ and $\mathbf{w}_i(\mu)$ requires operations only with $\mathbf{Y}_i$ and no other basis matrices, which implies efficiency in terms of storage and distribution of computations.

Similarly as in Section 3.3.2, the online efficiency of the minimization of (4.34) for a finite test set $\mathcal{P}_{\text{test}}$ of parameter values can be improved by using an extra oblivious $\ell_2 \to \ell_2$ subspace embedding $\mathbf{\Phi}$ (statistically) independent of $\mathcal{P}_{\text{test}}$, $\mathbf{W}_p(\mu)$

and $\mathbf{h}(\mu)$. The minimizer of (4.34) over $\mathbf{a}_p(\mu)$ can be approximated by the minimizer of

$$\|\mathbf{\Phi}(\mathbf{W}_p(\mu)\mathbf{a}_p(\mu) - \mathbf{h}(\mu))\|_2. \tag{4.35}$$

If $\mathbf{\Phi}$ is an $(\varepsilon, \delta(\#\mathcal{P}_{\text{test}})^{-1}, d+1)$ oblivious $\ell_2 \to \ell_2$ subspace embedding (see Section 4.4 for the definition), then the minimizer of (4.35) over $\mathbf{a}_p(\mu)$ is close to optimal with probability at least $1-\delta$. The Gaussian matrices and P-SRHT (in practice) satisfy this property if they have $\mathcal{O}(d + \log(\#\mathcal{P}_{\text{test}}) + \log(1/\delta))$ rows. The size of $\mathbf{\Phi}$ should be several times smaller than the size of $\mathbf{\Theta}$ required to guarantee the accuracy of the sketched error indicators for the whole parameter set $\mathcal{P}$ or for adaptively chosen parameters in the algorithms for the construction of the preconditioner's basis. In the online stage, $\mathbf{\Phi}\mathbf{W}_p(\mu)$ and $\mathbf{\Phi}\mathbf{h}(\mu)$ can be evaluated from their affine decompositions, which can be efficiently precomputed beforehand (in the intermediate online stage) by applying the map $\mathbf{\Phi}$ to the affine terms of $\mathbf{W}_p(\mu)$ and $\mathbf{h}(\mu)$.

4.4 Analysis of random sketching

In this section we provide a theoretical analysis of the sizes of the random sketching matrices to guarantee the quasi-optimality of the randomized error indicators from Section 4.3 with probability at least $1-\delta$. For this we first introduce a general framework and then particularize it to each error indicator from Section 4.3 individually.

4.4.1 ℓ_2-embeddings for vectors and matrices

Let X be a space of vectors or matrices equipped with an inner product $\langle \cdot, \cdot \rangle_X$. We will consider different cases: the space of vectors $X = U$ equipped with $\langle \cdot, \cdot \rangle_X = \langle \cdot, \cdot \rangle_U$, the space of matrices $X = HS(\ell_2, U)$ equipped with $\langle \cdot, \cdot \rangle_X = \langle \cdot, \cdot \rangle_{HS(\ell_2,U)} = \langle \mathbf{R}_U \cdot, \cdot \rangle_F$, or the space of matrices $X = HS(U', U)$ equipped with $\langle \cdot, \cdot \rangle_X = \langle \cdot, \cdot \rangle_{HS(U',U)} = \langle \mathbf{R}_U \cdot \mathbf{R}_U, \cdot \rangle_F$. See Section 4.1.2 for details.

Let V be a subspace of X. Let $\mathbf{\Theta}$ be a linear map from X to $\ell_2(\mathbb{K}^k)$ with $k \leq \dim(X)$, which is seen as an $X \to \ell_2$ subspace embedding.

Definition 4.4.1. *If $\mathbf{\Theta}$ satisfies*

$$\forall \mathbf{X}, \mathbf{Y} \in V, \ |\langle \mathbf{X}, \mathbf{Y} \rangle_X - \langle \mathbf{\Theta}(\mathbf{X}), \mathbf{\Theta}(\mathbf{Y}) \rangle_2| \leq \varepsilon \|\mathbf{X}\|_X \|\mathbf{Y}\|_X, \tag{4.36}$$

for some $\varepsilon \in [0,1)$, then it is called a $X \to \ell_2$ ε-subspace embedding for V.

An ε-embedding can be efficiently constructed with a probabilistic approach. Consider $\mathbf{\Theta}$ to be drawn from a certain random distribution of linear maps.

Definition 4.4.2. *The map $\mathbf{\Theta}$ is called a (ε,δ,m) oblivious $X\to\ell_2$ subspace embedding if for any m-dimensional subspace V_m of X it holds*

$$\mathbb{P}(\mathbf{\Theta} \textit{ is a } X\to\ell_2 \ \varepsilon\textit{-subspace embedding for } V_m)\geq 1-\delta. \tag{4.37}$$

A random matrix which is a (ε,δ,m) oblivious $X\to\ell_2$ subspace embedding, with $X=\mathbb{K}^n$ and $\langle\cdot,\cdot\rangle_X=\langle\cdot,\cdot\rangle_2$, is refereed to as a (ε,δ,m) oblivious $\ell_2\to\ell_2$ subspace embedding. Some distributions of matrices are known to be (ε,δ,m) oblivious $\ell_2\to\ell_2$ subspace embeddings. In this work, from the oblivious $\ell_2\to\ell_2$ subspace embeddings, we shall only explore the rescaled Gaussian and the SRHT distributions. A $k\times n$ rescaled Gaussian matrix has i.i.d. entries with mean 0 and variance k^{-1}. Assuming that n is the power of 2, a $k\times n$ SRHT matrix is defined as $k^{-1/2}(\mathbf{R}\mathbf{H}_n\mathbf{D})\in\mathbb{R}^{k\times n}$, where $\mathbf{R}\in\mathbb{R}^{k\times n}$ are the first k rows of an uniform random permutation of rows of the identity matrix, $\mathbf{H}_n\in\mathbb{R}^{n\times n}$ is a Walsh-Hadamard matrix and $\mathbf{D}\in\mathbb{R}^{n\times n}$ is a random diagonal matrix with random entries such that $\mathbb{P}([\mathbf{D}]_{i,i}=\pm1)=1/2$. The partial-SRHT (P-SRHT) is used when n is not necessarily a power of 2, and is defined as the first n columns of a SRHT matrix of size s, were s is the power of 2 and $n\leq s<2n$.

From Section 2.3.1 it follows that the rescaled Gaussian distribution with

$$k\geq 7.87\varepsilon^{-2}(C6.9m+\log(1/\delta)), \tag{4.38a}$$

where $C=1$ for $\mathbb{K}=\mathbb{R}$ or $C=2$ for $\mathbb{K}=\mathbb{C}$, and the P-SRHT distribution with

$$k\geq 2(\varepsilon^2-\varepsilon^3/3)^{-1}\left[\sqrt{m}+\sqrt{8\log(6n/\delta)}\right]^2\log(3m/\delta), \tag{4.38b}$$

respectively, are (ε,δ,m) oblivious $\ell_2\to\ell_2$ subspace embeddings. These random matrices can be used for the construction of (ε,δ,m) oblivious $X\to\ell_2$ subspace embeddings proposed in the next propositions and corollary. Let $\mathbf{Q}$ be a matrix such that $\mathbf{Q}^{\mathrm{H}}\mathbf{Q}=\mathbf{R}_U$. Note that $\mathbf{Q}$ and $\mathbf{Q}^{\mathrm{H}}$ are seen as operators from U to ℓ_2 and from ℓ_2 to U', respectively.

Corollary 4.4.3 (Proposition 2.3.11). *Let $\mathbf{\Omega}$ be a (ε,δ,m) oblivious $\ell_2\to\ell_2$ subspace embedding. The random matrix*

$$\mathbf{\Theta}:=\mathbf{\Omega}\mathbf{Q}$$

is a (ε,δ,m) oblivious $U\to\ell_2$ subspace embedding of subspaces of U.

Proposition 4.4.4. *The random map*

$$\mathbf{\Theta}(\mathbf{X}):=\mathbf{\Gamma}\ \mathtt{vec}(\mathbf{\Omega}\mathbf{Q}\mathbf{X}),\ \mathbf{X}:\mathbb{K}^r\to U,$$

where $\mathbf{\Gamma}$ and $\mathbf{\Omega}$ are $(\varepsilon_{\mathbf{\Gamma}},\delta_{\mathbf{\Gamma}},m)$ and $(\varepsilon_{\mathbf{\Omega}},\delta_{\mathbf{\Omega}},m)$ oblivious $\ell_2\to\ell_2$ subspace embeddings, is a (ε,δ,m) oblivious $HS(\ell_2,U)\to\ell_2$ subspace embedding of subspaces of matrices with r columns representing the vectors in U with $\varepsilon=(1+\varepsilon_{\mathbf{\Omega}})(1+\varepsilon_{\mathbf{\Gamma}})-1$ and $\delta=r\delta_{\mathbf{\Omega}}+\delta_{\mathbf{\Gamma}}$.

Proof. See appendix. □

Proposition 4.4.5. *The random map*

$$\mathbf{\Theta}(\mathbf{X}) := \mathbf{\Gamma}\ \texttt{vec}(\mathbf{\Omega}\mathbf{Q}\mathbf{X}\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}}),\ \mathbf{X}: U' \to U,$$

where $\mathbf{\Gamma}$*,* $\mathbf{\Omega}$ *and* $\mathbf{\Sigma}$ *are* $(\varepsilon_{\mathbf{\Gamma}}, \delta_{\mathbf{\Gamma}}, m)$*,* $(\varepsilon_{\mathbf{\Omega}}, \delta_{\mathbf{\Omega}}, m)$ *and* $(\varepsilon_{\mathbf{\Sigma}}, \delta_{\mathbf{\Sigma}}, m)$ *oblivious* $\ell_2 \to \ell_2$ *subspace embeddings, is a* (ε, δ, m) *oblivious* $HS(U', U) \to \ell_2$ *subspace embedding of matrices representing operators from* U' *to* U *with* $\varepsilon = (1+\varepsilon_{\mathbf{\Gamma}})(1+\varepsilon_{\mathbf{\Sigma}})(1+\varepsilon_{\mathbf{\Omega}}) - 1$ *and* $\delta = \min(k_{\mathbf{\Sigma}}\delta_{\mathbf{\Omega}} + n\delta_{\mathbf{\Sigma}},\ k_{\mathbf{\Omega}}\delta_{\mathbf{\Sigma}} + n\delta_{\mathbf{\Omega}}) + \delta_{\mathbf{\Gamma}}$*, where* $k_{\mathbf{\Omega}}$ *and* $k_{\mathbf{\Sigma}}$ *are the numbers of rows of* $\mathbf{\Omega}$ *and* $\mathbf{\Sigma}$*, respectively.*

Proof. See appendix. □

4.4.2 Analysis of randomized error indicators

Let us now apply the above setting to the randomized error indicators from Section 4.3.

Multi-purpose

We consider a situation where the preconditioned operators $\mathbf{B}(\mu)$ lie in a space V of operators from U to U' and V has a low dimension m. This is the case when the preconditioner has the form (4.28) and is obtained by minimization of an error indicator. Then for a fixed parameter value one may choose $V = \operatorname{span}\{\mathbf{R}_U \mathbf{Y}_i \mathbf{A}(\mu) : 1 \leq i \leq p\}$ with $m = p$. This is also the case when the preconditioner is provided (and therefore belongs to a one-dimensional space) with the objective to estimate the condition number (and quasi-optimality constants).

Corollary 4.4.6. *If* $\mathbf{\Theta}$ *is an* $(\varepsilon, \delta, m+1)$ *oblivious* $HS(U', U) \to \ell_2$ *subspace embedding, then we have*

$$\mathbb{P}\left(\forall \mathbf{B} \in V,\ |\Delta_{U,U}{}^2 - \Delta^{\mathrm{sk}}_{U,U}{}^2| \leq \varepsilon \Delta_{U,U}{}^2\right) \geq 1 - \delta. \tag{4.39}$$

Proof. This is an immediate corollary of Definitions 4.4.1 and 4.4.2 and the fact that $\|\mathbf{R}_U - \mathbf{B}\|_{HS(U,U')} = \|\mathbf{R}_U^{-1}[\mathbf{R}_U - \mathbf{B}]\mathbf{R}_U^{-1}\|_{HS(U',U)}$. □

Observe that (4.39) implies with high probability the quasi-optimality of the minimizer of $\Delta^{\mathrm{sk}}_{U,U}$ over V (or a subspace of V) as a minimizer of $\Delta_{U,U}$. A random map $\mathbf{\Theta}$, which is a $(\varepsilon, \delta, m+1)$ oblivious $HS(U', U) \to \ell_2$ subspace embedding, can be constructed using Proposition 4.4.5 with Gaussian matrices or P-SRHT as $\ell_2 \to \ell_2$ subspace embeddings. The conditions (4.38) can be used for a priori selection of the sizes of random sketching matrices.

Galerkin projection

As was discussed in Section 4.3, when the orthogonalization of the basis for U_r with respect to $\langle \cdot, \cdot \rangle_U$ is expensive, the basis should be orthogonalized approximately, which can be done with the random sketching technique.

Corollary 4.4.7. *If the columns of* $\mathbf{U}_r$ *are unit-orthogonal vectors with respect to* $\langle \mathbf{\Omega} \mathbf{Q} \cdot, \mathbf{\Omega} \mathbf{Q} \cdot \rangle_2$*, where* $\mathbf{\Omega}$ *is* (ε, δ, r) *oblivious* $\ell_2 \to \ell_2$ *embedding, then with probability at least* $1-\delta$*, all singular values of* $\mathbf{U}_r$ *are bounded below by* $\sqrt{1-\varepsilon}$ *and are bounded above by* $\sqrt{1+\varepsilon}$*.*

Proof. This is an immediate corollary of Definitions 4.4.1 and 4.4.2 and Corollary 4.4.3. □

It follows that orthogonalization of $\mathbf{U}_r$ with respect to $\langle \mathbf{\Omega} \mathbf{Q} \cdot, \mathbf{\Omega} \mathbf{Q} \cdot \rangle_2$ will with high probability provide constants $\Delta_{U_r,U_r}(\mu)$ and $\Delta_{U_r,U}(\mu)$ that yield accurate estimation of the quasi-optimality constants of the Galerkin projection onto U_r.

Next we address the estimation of $\Delta_{U_r,U_r}(\mu)$ and $\Delta_{U_r,U}(\mu)$. As for the multi-purpose context, we consider a scenario where the operators $\mathbf{B}(\mu)$ of interest lie in a certain fixed m-dimensional space V of operators from U to U'.

Corollary 4.4.8. *If* $\mathbf{\Sigma}$ *is* $(\varepsilon, \delta, m+1)$ *oblivious* $\ell_2 \to \ell_2$ *subspace embedding and* $\mathbf{\Theta}$ *is* $(\varepsilon, \delta, m+1)$ *oblivious* $HS(\ell_2, U) \to \ell_2$ *subspace embedding, then*

$$\mathbb{P}\left(\forall \mathbf{B} \in V,\ |{\Delta_{U_r,U_r}}^2 - {\Delta^{\mathrm{sk}}_{U_r,U_r}}^2| \leq \varepsilon {\Delta_{U_r,U_r}}^2 \right) \geq 1-\delta \tag{4.40a}$$

and

$$\mathbb{P}\left(\forall \mathbf{B} \in V,\ |{\Delta_{U_r,U}}^2 - {\Delta^{\mathrm{sk}}_{U_r,U}}^2| \leq \varepsilon {\Delta_{U_r,U}}^2 \right) \geq 1-\delta. \tag{4.40b}$$

Proof. This is an immediate corollary of Definitions 4.4.1 and 4.4.2 and the fact that $\|[\mathbf{R}_U - \mathbf{B}]^{\mathrm{H}} \mathbf{U}_r\|_{HS(\ell_2, U')} = \|\mathbf{R}_U^{-1} [\mathbf{R}_U - \mathbf{B}]^{\mathrm{H}} \mathbf{U}_r\|_{HS(\ell_2, U)}$. □

Relations (4.39) and (4.40) imply the quasi-optimality of the minimizer of (for some given weights $\gamma_1, \gamma_2, \gamma_3$)

$$\left(\gamma_1 {\Delta^{\mathrm{sk}}_{U,U}}^2 + \gamma_2 {\Delta^{\mathrm{sk}}_{U_r,U}}^2 + \gamma_3 {\Delta^{\mathrm{sk}}_{U_r,U_r}}^2 \right)^{1/2}$$

as a minimizer of

$$\left(\gamma_1 {\Delta_{U,U}}^2 + \gamma_2 {\Delta_{U_r,U}}^2 + \gamma_3 {\Delta_{U_r,U_r}}^2 \right)^{1/2}$$

over V (or a subspace of V) with high probability. The random map $\mathbf{\Theta}$ can be constructed using Proposition 4.4.4. The oblivious $\ell_2 \to \ell_2$ subspace embeddings ($\mathbf{\Gamma}$

and $\mathbf{\Omega}$) used for the construction of $\mathbf{\Theta}$, and the random matrix $\mathbf{\Sigma}$ can be readily taken as Gaussian or P-SRHT matrices with sufficiently large numbers of rows chosen according to (4.38).

There are two ways to guarantee a success of the sketched estimation of an error indicator (or a weighted average of error indicators) with probability at least $1-\delta^*$ for all parameter values in $\mathcal{P}$, simultaneously. If $\mathcal{P}$ is finite then one can simply consider success for each parameter value, separately, and then use a union bound argument, therefore using $\delta = \delta^*/\#\mathcal{P}$ and $m = p$ for the selection of sizes of random matrices. A second way is to exploit the fact that the set $\bigcup_{\mu \in \mathcal{P}} \{\mathbf{B}(\mu)\}$ is a subset of some low-dimensional space. For instance, if $\mathbf{A}(\mu)$ has affine expansion with m_A terms and $\mathbf{P}(\mu)$ is of the form (4.28) then there exists such a space with dimension $m \leq m_A p$. This space can be readily chosen as the space V in the above considerations.

Let us underline that the two characterizations of the probability of success only hold if $\{\mathbf{Y}_i\}_{i=1}^p$ is (statistically) independent of $\mathbf{\Gamma}$, $\mathbf{\Omega}$ and $\mathbf{\Sigma}$, which is not the case in the greedy algorithm for the selection of the parameter values for the interpolation of matrix inverse in Section 4.2.4. For the adaptive algorithms, one has to consider all possible outcomes with another union bound for the probability of success. In particular, if the training set has cardinality M, then there can exist up to $\binom{M}{p}$ possible outcomes of the greedy selection of p basis matrices and a success has to be guaranteed for each of them. In practice, this implies an increase of the probability of failure by a factor of $\binom{M}{p}$. Luckily the required sizes of random matrices depend only logarithmically on the probability of failure, therefore the replacement of δ by $\delta\binom{M}{p}$ shall not catastrophically affect the computational costs.

Error estimation

Consider error estimation with Proposition 4.2.7. Assume that the residual vectors $\mathbf{r}^*(\mathbf{u}_r(\mu);\mu)$ and $\mathbf{d}(\mathbf{u}_r(\mu);\mu)$ are contained in some fixed subspaces $R^* \subseteq U'$ and $D \subseteq U'$ of low dimensions $m_\mathbf{r}$ and $m_\mathbf{d}$, respectively. This situation appears for a fixed parameter value with $m_\mathbf{r} = m_\mathbf{d} = p+1$, if $\mathbf{P}(\mu)$ is of the form (4.28).

Corollary 4.4.9. *If $\mathbf{\Theta}$ is a $(\varepsilon,\delta,m_\mathbf{r})$ oblivious $U \to \ell_2$ subspace embedding, then we have*

$$\mathbb{P}\left(\forall \mathbf{r}^* \in R^*,\ \left|\|\mathbf{r}^*\|_{U'}^2 - \|\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{r}^*\|_2^2\right| \leq \varepsilon \|\mathbf{r}^*\|_{U'}^2\right) \geq 1-\delta.$$

Furthermore, if $\mathbf{\Theta}$ is a $(\varepsilon,\delta,m_\mathbf{d})$ oblivious $U \to \ell_2$ subspace embedding,

$$\mathbb{P}\left(\forall \mathbf{d} \in D,\ |{\Delta_\mathrm{e}}^2 - {\Delta_\mathrm{e}^\mathrm{sk}}^2| \leq \varepsilon {\Delta_\mathrm{e}}^2\right) \geq 1-\delta.$$

Proof. These two statements follow immediately from Definitions 4.4.1 and 4.4.2 and the fact that

$$\|\mathbf{r}^*\|_{U'} = \|\mathbf{R}_U^{-1}\mathbf{r}^*\|_U \text{ and } \|\mathbf{d}\|_{U'} = \|\mathbf{R}_U^{-1}\mathbf{d}\|_U.$$

□

The above relations imply with high probability the (essential) preservation of the quality of the error certification with Proposition 4.2.7, when $\|\mathbf{r}^*(\mathbf{u}_r(\mu);\mu)\|_{U'}$ and $\Delta_{\mathrm{e}}(\mu)$ are substituted by their efficient sketched estimations $\|\mathbf{\Theta}\mathbf{R}_U^{-1}\mathbf{r}^*(\mathbf{u}_r(\mu);\mu)\|_2$ and $\Delta_{\mathrm{e}}^{\mathrm{sk}}(\mu)$. Furthermore, for the projection-based construction of the preconditioner proposed in Section 4.2.4, a quasi-optimal minimizer of $\Delta_{\mathrm{e}}(\mu)$ over $\lambda_1(\mu),\ldots,\lambda_p(\mu)$ can with high probability be obtained by minimization of $\Delta_{\mathrm{e}}^{\mathrm{sk}}(\mu)$. The random matrix $\mathbf{\Theta}$ can here be constructed with Corollary 4.4.3.

If the approximation $\mathbf{u}_r(\mu)$ of $\mathbf{u}(\mu)$ is provided and $\mathcal{P}$ is finite, then the success of error estimation (or certification) for all $\{\mathbf{u}_r(\mu) : \mu \in \mathcal{P}\}$, simultaneously, can be guaranteed by considering error estimation for each $\mathbf{u}_r(\mu)$ separately and using a union bound argument. It follows that for having a probability of success of at least $1-\delta^*$, we can choose $\delta = \delta^*/\#\mathcal{P}$ and $m_{\mathbf{r}} = m_{\mathbf{d}} = p+1$ for the selection of sizes of random matrices, if $\mathbf{P}(\mu)$ is of the form (4.28).

In some cases one may want to obtain an effective upper bound of the error $\|\mathbf{u}_r - \mathbf{u}(\mu)\|_U$ for all vectors $\mathbf{u}_r$ from a r-dimensional approximation space U_r. For each parameter value, the success of estimation for all $\mathbf{u}_r \in U_r$ can be guaranteed by exploiting the fact that $\mathbf{r}^*(\mathbf{u}_r;\mu)$ and $\mathbf{d}(\mathbf{u}_r;\mu)$ lie in low-dimensional spaces R^* and D of dimensions at most $m_{\mathbf{r}} = m_{\mathbf{d}} = rp+1$ (if $\mathbf{P}(\mu)$ has the form (4.28)). To guarantee the success for all parameter values in $\mathcal{P}$, we can again use a union bound argument by choosing $\delta = \delta^*/\#\mathcal{P}$.

Alternatively, for infinite $\mathcal{P}$ we can choose R^* and D as low-dimensional spaces which contain $\bigcup_{\mu\in\mathcal{P}}\{\mathbf{r}^*(\mathbf{u}_r(\mu);\mu)\}$ and $\bigcup_{\mu\in\mathcal{P}}\{\mathbf{d}(\mathbf{u}_r(\mu);\mu)\}$ for the estimation for the given $\mathbf{u}_r(\mu)$, or $\bigcup_{\mu\in\mathcal{P}}\bigcup_{\mathbf{u}_r\in U_r}\{\mathbf{r}^*(\mathbf{u}_r;\mu)\}$ and $\bigcup_{\mu\in\mathcal{P}}\bigcup_{\mathbf{u}_r\in U_r}\{\mathbf{d}(\mathbf{u}_r;\mu)\}$ for the estimation for all $\mathbf{u}_r \in U_r$. Note that such low-dimensional spaces exist if $\mathbf{A}(\mu)$ and $\mathbf{b}(\mu)$ have affine representations with a small number of terms and $\mathbf{P}(\mu)$ is of the form (4.28).

4.5 Appendix

This section provides the proofs of propositions from the chapter.

Proof of Proposition 4.2.1. We have,

$$\|\mathbf{C}\mathbf{v}\|_{U'} = \|(\mathbf{R}_U^{-1/2}\mathbf{C}\mathbf{R}_U^{-1/2})(\mathbf{R}_U^{1/2}\mathbf{v})\|_2 \leq \|\mathbf{R}_U^{-1/2}\mathbf{C}\mathbf{R}_U^{-1/2}\|_F\|\mathbf{R}_U^{1/2}\mathbf{v}\|_2 = \|\mathbf{C}\|_{HS(U,U')}\|\mathbf{v}\|_U,$$

which ends the proof. □

Proof of Lemma 4.2.2. We have,

$$\begin{aligned}\|\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u}\|_U &= \|\mathrm{P}_{U_r}(\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u})\|_U \\ &\leq \|\mathrm{P}_{U_r}\mathbf{R}_U^{-1}\mathbf{B}(\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u})\|_U + \|\mathrm{P}_{U_r}\mathbf{R}_U^{-1}[\mathbf{R}_U - \mathbf{B}](\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u})\|_U \\ &= \|\mathbf{B}(\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u})\|_{U_r'} + \|[\mathbf{R}_U - \mathbf{B}](\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u})\|_{U_r'} \\ &\leq \|\mathbf{B}(\mathbf{u} - \mathrm{P}_{U_r}\mathbf{u})\|_{U_r'} + \|[\mathbf{R}_U - \mathbf{B}](\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u})\|_{U_r'} \\ &\leq \|[\mathbf{R}_U - \mathbf{B}](\mathbf{u} - \mathrm{P}_{U_r}\mathbf{u})\|_{U_r'} + \|[\mathbf{R}_U - \mathbf{B}](\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u})\|_{U_r'},\end{aligned}$$

which completes the proof. □

Proof of Proposition 4.2.3. The relation (4.14) and the uniqueness of $\mathbf{u}_r$ directly follow from Lemma 4.2.2 and the definitions of constants $\beta_r(\mathbf{R}_U - \mathbf{B})$ and $\bar{\beta}_r(\mathbf{R}_U - \mathbf{B})$. More specifically, Lemma 4.2.2 implies that

$$\begin{aligned}\|\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u}\|_U &\leq \|[\mathbf{R}_U - \mathbf{B}](\mathbf{u} - \mathrm{P}_{U_r}\mathbf{u})\|_{U_r'} + \|[\mathbf{R}_U - \mathbf{B}](\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u})\|_{U_r'} \\ &\leq \bar{\beta}_r(\mathbf{R}_U - \mathbf{B})\|\mathbf{u} - \mathrm{P}_{U_r}\mathbf{u}\|_U + \beta_r(\mathbf{R}_U - \mathbf{B})\|\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u}\|_U,\end{aligned}$$

which combined with the inequality

$$\|\mathbf{u}_r - \mathrm{P}_{U_r}\mathbf{u}\|_U \geq \|\mathbf{u} - \mathbf{u}_r\|_U - \|\mathbf{u} - \mathrm{P}_{U_r}\mathbf{u}\|_U$$

yields (4.14). The uniqueness of $\mathbf{u}_r$ classically follows from the argument that if $\mathbf{u}_r \in U_r$ and $\mathbf{v}_r \in U_r$ satisfy the Galerkin orthogonality condition, then

$$\begin{aligned}0 = \|\mathbf{B}(\mathbf{u} - \mathbf{u}_r)||_{U_r'} + \|\mathbf{B}(\mathbf{u} - \mathbf{v}_r)||_{U_r'} &\geq \|\mathbf{B}(\mathbf{v}_r - \mathbf{u}_r)||_{U_r'} \\ &\geq \|\mathbf{R}_U(\mathbf{v}_r - \mathbf{u}_r)||_{U_r'} - \|(\mathbf{R}_U - \mathbf{B})(\mathbf{v}_r - \mathbf{u}_r)||_{U_r'} \\ &\geq (1 - \beta_r(\mathbf{R}_U - \mathbf{B}))\|\mathbf{v}_r - \mathbf{u}_r\|_U,\end{aligned}$$

which implies that $\mathbf{v}_r = \mathbf{u}_r$. □

Proof of Proposition 4.2.4. Let $\mathbf{T}$ be such that $\mathbf{U}_r^* := \mathbf{U}_r\mathbf{T}$ has unit-orthogonal columns with respect to $\langle\cdot,\cdot\rangle_U$, and $\mathrm{range}(\mathbf{U}_r^*) = U_r$. The identity $\mathrm{P}_{U_r} = \mathbf{U}_r^*\mathbf{U}_r^{*\mathrm{H}}\mathbf{R}_U$ implies that

$$\|\cdot\|_{U_r'} = \|\mathrm{P}_{U_r}\mathbf{R}_U^{-1}\cdot\|_U = \|\mathbf{U}_r^*\mathbf{U}_r^{*\mathrm{H}}\cdot\|_U = \|\mathbf{U}_r^{*\mathrm{H}}\cdot\|_2 = \|\mathbf{T}^{\mathrm{H}}\mathbf{U}_r{}^{\mathrm{H}}\cdot\|_2.$$

From this fact we obtain the following expressions for $\beta_r(\mathbf{R}_U - \mathbf{B})$ and $\bar{\beta}_r(\mathbf{R}_U - \mathbf{B})$:

$$\beta_r(\mathbf{R}_U - \mathbf{B}) = \max_{\mathbf{v}\in U_r\backslash\{\mathbf{0}\}} \frac{\|\mathbf{T}^{\mathrm{H}}{\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{v}\|_2}{\|\mathbf{v}\|_U} = \|\mathbf{T}^{\mathrm{H}}{\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{U}_r\mathbf{T}\|_2 \quad (4.41a)$$

and

$$\bar{\beta}_r(\mathbf{R}_U - \mathbf{B}) := \max_{\mathbf{v}\in U\backslash\{\mathbf{0}\}} \frac{\|\mathbf{T}^{\mathrm{H}}{\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{v}\|_2}{\|\mathbf{v}\|_U} = \|\mathbf{T}^{\mathrm{H}}{\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{R}_U^{-1/2}\|_2. \quad (4.41b)$$

It can be shown that the minimal and the maximal singular values of $\mathbf{T}$ are equal to σ_1^{-1} and σ_r^{-1}, respectively. Then for any matrices $\mathbf{X}$ and $\mathbf{X}^*$ with $\mathbf{X}^* = \mathbf{X}\mathbf{T}$ or $\mathbf{T}^{\mathrm{H}}\mathbf{X}$, it holds

$$\sigma_1^{-1}\|\mathbf{X}\|_2 \leq \|\mathbf{X}^*\|_2 \leq \sigma_r^{-1}\|\mathbf{X}\|_2. \quad (4.42)$$

By choosing in (4.42), first $\mathbf{X} = \mathbf{T}^{\mathrm{H}}{\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{U}_r$ with $\mathbf{X}^* = \mathbf{X}\mathbf{T}$, and then $\mathbf{X} = {\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{U}_r$ with $\mathbf{X}^* = \mathbf{T}^{\mathrm{H}}\mathbf{X}$, and using (4.41a) we obtain

$$\sigma_1^{-2}\|{\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{U}_r\|_2 \leq \beta_r(\mathbf{R}_U - \mathbf{B}) \leq \sigma_r^{-2}\|{\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{U}_r\|_2.$$

At the same time, by choosing $\mathbf{X} = {\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{R}_U^{-1/2}$ with $\mathbf{X}^* = \mathbf{T}^{\mathrm{H}}\mathbf{X}$ in (4.42), and using (4.41b) we get

$$\sigma_1^{-1}\|{\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{R}_U^{-1/2}\|_2 \leq \bar{\beta}_r(\mathbf{R}_U - \mathbf{B}) \leq \sigma_r^{-1}\|{\mathbf{U}_r}^{\mathrm{H}}[\mathbf{R}_U - \mathbf{B}]\mathbf{R}_U^{-1/2}\|_2.$$

These two relations combined with the fact that for a matrix $\mathbf{X}$ with r rows,

$$r^{-1/2}\|\mathbf{X}\|_F \leq \|\mathbf{X}\|_2 \leq \|\mathbf{X}\|_F,$$

result in (4.18). □

Proof of Proposition 4.2.6. By definition of $\alpha(\mathbf{B})$ and $\beta(\mathbf{B})$, we have

$$\beta(\mathbf{B})^{-1}\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'} \leq \|\mathbf{u} - \mathbf{u}_r\| \leq \alpha(\mathbf{B})^{-1}\|\mathbf{r}^*(\mathbf{u}_r)\|_{U'}. \quad (4.43)$$

Moreover, since $\Delta_{U,U}$ is an upper bound of $\beta(\mathbf{R}_U - \mathbf{B})$, the following inequalities hold

$$1 - \Delta_{U,U} \leq 1 - \beta(\mathbf{B} - \mathbf{R}_U) \leq \alpha(\mathbf{B}) \leq \beta(\mathbf{B}) \leq 1 + \beta(\mathbf{B} - \mathbf{R}_U) \leq 1 + \Delta_{U,U}. \quad (4.44)$$

The result of the proposition follows immediately from (4.43) and (4.44). □

Proof of Proposition 4.2.7. We have,

$$\begin{aligned}\|\mathbf{u}-\mathbf{u}_r\|_U &= \|\mathbf{R}_U(\mathbf{u}-\mathbf{u}_r)\|_{U'}\\ &\leq \|\mathbf{B}(\mathbf{u}-\mathbf{u}_r)\|_{U'} + \|[\mathbf{R}_U-\mathbf{B}](\mathbf{u}-\mathbf{u}_r)\|_{U'}\\ &= \|\mathbf{r}^*(\mathbf{u}_r)\|_{U'} + \|[\mathbf{I}-\mathbf{PA}](\mathbf{u}-\mathbf{u}_r)\|_U\\ &\leq \|\mathbf{r}^*(\mathbf{u}_r)\|_{U'} + \frac{1}{\eta}\|\mathbf{A}[\mathbf{I}-\mathbf{PA}](\mathbf{u}-\mathbf{u}_r)\|_{U'}\\ &= \|\mathbf{r}^*(\mathbf{u}_r)\|_{U'} + \frac{1}{\eta}\|[\mathbf{I}-\mathbf{AP}]\mathbf{r}(\mathbf{u}_r)\|_{U'}.\end{aligned}$$

The lower bound in (4.27) is derived similarly. □

Proof of Proposition 4.4.4. Let $V_m \subset HS(\ell_2, U)$ be any m-dimensional space of matrices with r columns representing vectors in U. Let $V_m^i \subset U$ denote the subspace spanned by the i-th column vectors of matrices from V_m, i.e.,

$$V_m^i := \{\mathbf{V}_m\mathbf{e}_i : \mathbf{V}_m \in V_m\},$$

where $\mathbf{e}_i$ denotes the i-th column of the identity matrix.

By Corollary 4.4.3, $\mathbf{\Omega Q}$ is a $\varepsilon_{\mathbf{\Omega}}$-embedding for each V_m^i with probability at least $1-\delta_{\mathbf{\Omega}}$. This fact combined with a union bound argument imply that

$$|\|\mathbf{V}_m\mathbf{e}_i\|_U^2 - \|\mathbf{\Omega Q V}_m\mathbf{e}_i\|_2^2| \leq \varepsilon_{\mathbf{\Omega}}\|\mathbf{V}_m\mathbf{e}_i\|_U^2$$

holds with probability at least $1-r\delta_{\mathbf{\Omega}}$ for all $\mathbf{V}_m \in V_m$ and $1 \leq i \leq r$. From the above relation and identities

$$\|\mathbf{V}_m\|_{HS(\ell_2,U)}^2 = \sum_{i=1}^r \|\mathbf{V}_m\mathbf{e}_i\|_U^2 \text{ and } \|\mathtt{vec}(\mathbf{\Omega Q V}_m)\|_2^2 = \|\mathbf{\Omega Q V}_m\|_F^2 = \sum_{i=1}^r \|\mathbf{\Omega Q V}_m\mathbf{e}_i\|_2^2,$$

we obtain that

$$\mathbb{P}(\forall \mathbf{V}_m \in V_m,\ |\|\mathbf{V}_m\|_{HS(\ell_2,U)}^2 - \|\mathtt{vec}(\mathbf{\Omega Q V}_m)\|_2^2| \leq \varepsilon_{\mathbf{\Omega}}\|\mathbf{V}_m\|_{HS(\ell_2,U)}^2) \geq 1-r\delta_{\mathbf{\Omega}}. \tag{4.45}$$

Moreover, by definition of $\mathbf{\Gamma}$,

$$\mathbb{P}(\forall \mathbf{V}_m \in V_m,\ |\|\mathtt{vec}(\mathbf{\Omega Q V}_m)\|_2^2 - \|\mathbf{\Gamma}\mathtt{vec}(\mathbf{\Omega Q V}_m)\|_2^2| \leq \varepsilon_{\mathbf{\Gamma}}\|\mathtt{vec}(\mathbf{\Omega Q V}_m)\|_2^2) \geq 1-\delta_{\mathbf{\Gamma}}. \tag{4.46}$$

By (4.45) and (4.46) and a union bound for the probability of success,

$$\begin{aligned}&|\|\mathbf{V}_m\|_{HS(\ell_2,U)}^2 - \|\mathbf{\Gamma}\mathtt{vec}(\mathbf{\Omega Q V}_m)\|_2^2|\\ &\leq |\|\mathbf{V}_m\|_{HS(\ell_2,U)}^2 - \|\mathtt{vec}(\mathbf{\Omega Q V}_m)\|_2^2| + |\|\mathtt{vec}(\mathbf{\Omega Q V}_m)\|_2^2 - \|\mathbf{\Gamma}\mathtt{vec}(\mathbf{\Omega Q V}_m)\|_2^2|\\ &\leq \varepsilon_{\mathbf{\Omega}}\|\mathbf{V}_m\|_{HS(\ell_2,U)}^2 + \varepsilon_{\mathbf{\Gamma}}\|\mathtt{vec}(\mathbf{\Omega Q V}_m)\|_2^2 \leq \varepsilon_{\mathbf{\Omega}}\|\mathbf{V}_m\|_{HS(\ell_2,U)}^2 + \varepsilon_{\mathbf{\Gamma}}(1+\varepsilon_{\mathbf{\Omega}})\|\mathbf{V}_m\|_{HS(\ell_2,U)}^2\\ &= \varepsilon\|\mathbf{V}_m\|_U^2\end{aligned}$$

holds with probability at least $1-\delta$ for all $\mathbf{V}_m \in V_m$. It can be easily shown by using the parallelogram identity that the above statement is equivalent to

$$\mathbb{P}(\forall \mathbf{X}, \mathbf{Y} \in V_m, \left|\langle \mathbf{X}, \mathbf{Y}\rangle_{HS(\ell_2,U)} - \langle \mathbf{\Theta}(\mathbf{X}), \mathbf{\Theta}(\mathbf{Y})\rangle_2\right| \leq \varepsilon \|\mathbf{X}\|_{HS(\ell_2,U)}\|\mathbf{Y}\|_{HS(\ell_2,U)}) \geq 1-\delta,$$

which implies the statement of the proposition. □

Proof of Proposition 4.4.5. Let us first assume that

$$k_{\mathbf{\Sigma}}\delta_{\mathbf{\Omega}} + n\delta_{\mathbf{\Sigma}} \leq k_{\mathbf{\Omega}}\delta_{\mathbf{\Sigma}} + n\delta_{\mathbf{\Omega}}.$$

Let V_m be a m-dimensional space of operators from U' to U. Define

$$W_m^i = \{\mathbf{V}_m^{\mathrm{H}}\mathbf{R}_U^{1/2}\mathbf{e}_i : \mathbf{V}_m \in V_m\} \subset U,$$

where $\mathbf{e}_i$ denotes the i-th column of the identity matrix. By Corollary 4.4.3 and a union bound argument, we have that $\mathbf{\Sigma}\mathbf{Q}$ is an ε-embedding for all W_m^i, $1 \leq i \leq n$, with probability at least $1-n\delta_{\mathbf{\Sigma}}$. This implies that

$$|\|\mathbf{V}_m^{\mathrm{H}}\mathbf{R}_U^{1/2}\mathbf{e}_i\|_U^2 - \|\mathbf{\Sigma}\mathbf{Q}[\mathbf{V}_m^{\mathrm{H}}\mathbf{R}_U^{1/2}\mathbf{e}_i]\|_2^2| \leq \varepsilon_{\mathbf{\Sigma}}\|\mathbf{V}_m^{\mathrm{H}}\mathbf{R}_U^{1/2}\mathbf{e}_i\|_U^2 \tag{4.47}$$

holds with probability at least $1-n\delta_{\mathbf{\Sigma}}$ for all $\mathbf{V}_m \in V_m$ and $1 \leq i \leq n$. By (4.47) and the following identities

$$\|\mathbf{V}_m\|^2_{HS(U',U)} = \|\mathbf{V}_m^{\mathrm{H}}\|^2_{HS(U',U)} = \sum_{i=1}^n \|\mathbf{V}_m^{\mathrm{H}}\mathbf{R}_U^{1/2}\mathbf{e}_i\|_U^2$$

and

$$\|\mathbf{V}_m\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}}\|^2_{HS(\ell_2,U)} = \|\mathbf{R}_U^{1/2}\mathbf{V}_m\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}}\|_F^2 = \sum_{i=1}^n \|\mathbf{\Sigma}\mathbf{Q}[\mathbf{V}_m^{\mathrm{H}}\mathbf{R}_U^{1/2}\mathbf{e}_i]\|_2^2,$$

we deduce that

$$\mathbb{P}(\forall \mathbf{V}_m \in V_m, |\|\mathbf{V}_m\|^2_{HS(U',U)} - \|\mathbf{V}_m\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}}\|^2_{HS(\ell_2,U)}| \leq \varepsilon_{\mathbf{\Sigma}}\|\mathbf{V}_m\|^2_{HS(U',U)}) \geq 1-n\delta_{\mathbf{\Sigma}}. \tag{4.48}$$

Furthermore, by Proposition 4.2.7, the linear map

$$\mathbf{\Gamma}\mathtt{vec}(\mathbf{\Omega}\mathbf{Q}\mathbf{X}),\ \mathbf{X} : \mathbb{K}^{k_{\mathbf{\Sigma}}} \to U$$

is a $(\varepsilon^*, \delta^*, m)$ oblivious $HS(\ell_2, U) \to \ell_2$ subspace embedding of subspaces of matrices with $k_{\mathbf{\Sigma}}$ columns with $\varepsilon^* = (1+\varepsilon_{\mathbf{\Omega}})(1+\varepsilon_{\mathbf{\Gamma}}) - 1$ and $\delta^* = k_{\mathbf{\Sigma}}\delta_{\mathbf{\Omega}} + \delta_{\mathbf{\Gamma}}$. Consequently, with probability at least $1-\delta^*$, for all $\mathbf{V}_m \in V_m$, it holds

$$|\|\mathbf{V}_m\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}}\|^2_{HS(\ell_2,U)} - \|\mathbf{\Gamma}\mathtt{vec}(\mathbf{\Omega}\mathbf{Q}\mathbf{V}_m\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}})\|_2^2| \leq \varepsilon^*\|\mathbf{V}_m\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}}\|^2_{HS(\ell_2,U)}. \tag{4.49}$$

From (4.48) and (4.49) and a union bound for the probability of success, we obtain that

$$\begin{aligned}
&|\|\mathbf{V}_m\|^2_{HS(U',U)} - \|\mathbf{\Gamma}\mathtt{vec}(\mathbf{\Omega}\mathbf{Q}\mathbf{V}_m\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}})\|^2_2| \\
&\le |\|\mathbf{V}_m\|^2_{HS(U',U)} - \|\mathbf{V}_m\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}}\|^2_{HS(\ell_2,U)}| + |\|\mathbf{V}_m\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}}\|^2_{HS(\ell_2,U)} - \|\mathbf{\Gamma}\mathtt{vec}(\mathbf{\Omega}\mathbf{Q}\mathbf{V}_m\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}})\|^2_2| \\
&\le \varepsilon_{\mathbf{\Sigma}}\|\mathbf{V}_m\|^2_{HS(U',U)} + \varepsilon^*\|\mathbf{V}_m\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}}\|^2_{HS(\ell_2,U)} \\
&\le \varepsilon_{\mathbf{\Sigma}}\|\mathbf{V}_m\|^2_{HS(U',U)} + \varepsilon^*(1+\varepsilon_{\mathbf{\Sigma}})\|\mathbf{V}_m\|^2_{HS(U',U)} \\
&= \varepsilon\|\mathbf{V}_m\|^2_{HS(U',U)}
\end{aligned}$$

holds with probability at least $1-(\delta^*+n\delta_{\mathbf{\Sigma}})$ for all $\mathbf{V}_m \in V_m$. This statement with the parallelogram identity imply that, with probability at least $1-\delta$, for all $\mathbf{X},\mathbf{Y} \in V_m$

$$\left|\langle \mathbf{X},\mathbf{Y}\rangle_{HS(U',U)} - \langle \mathbf{\Theta}(\mathbf{X}),\mathbf{\Theta}(\mathbf{Y})\rangle_{HS(U',U)}\right| \le \varepsilon\|\mathbf{X}\|_{HS(U',U)}\|\mathbf{Y}\|_{HS(U',U)},$$

which completes the proof for the case

$$k_{\mathbf{\Sigma}}\delta_{\mathbf{\Omega}} + n\delta_{\mathbf{\Sigma}} \le k_{\mathbf{\Omega}}\delta_{\mathbf{\Sigma}} + n\delta_{\mathbf{\Omega}}.$$

For the alternative case, we can apply the proof of the first case by interchanging $\mathbf{\Omega}$ with $\mathbf{\Sigma}$ and considering a reshaping operator $\mathtt{vec}^*(\cdot) := \mathtt{vec}(\cdot^{\mathrm{H}})$ instead of the operator $\mathtt{vec}(\cdot)$ to show that the linear map

$$\mathbf{\Gamma}\mathtt{vec}^*(\mathbf{\Sigma}\mathbf{Q}\mathbf{X}\mathbf{Q}^{\mathrm{H}}\mathbf{\Omega}^{\mathrm{H}}),\ \mathbf{X}: U' \to U,$$

is a (ε,δ,m) oblivious $HS(U',U) \to \ell_2$ subspace embedding. Since the Frobenius inner product (and $\langle\cdot,\cdot\rangle_{HS(U',U)}$) of two matrices is equal to the Frobenius inner product (and $\langle\cdot,\cdot\rangle_{HS(U',U)}$) of the (Hermitian-)transposed matrices, the linear map

$$\mathbf{\Gamma}\mathtt{vec}^*(\mathbf{\Sigma}\mathbf{Q}\mathbf{X}^{\mathrm{H}}\mathbf{Q}^{\mathrm{H}}\mathbf{\Omega}^{\mathrm{H}}),\ \mathbf{X}: U' \to U,$$

is also a (ε,δ,m) oblivious $HS(U',U) \to \ell_2$ subspace embedding. The proof is completed by noticing that

$$\mathbf{\Gamma}\mathtt{vec}^*(\mathbf{\Sigma}\mathbf{Q}\mathbf{X}^{\mathrm{H}}\mathbf{Q}^{\mathrm{H}}\mathbf{\Omega}^{\mathrm{H}}) = \mathbf{\Gamma}\mathtt{vec}((\mathbf{\Sigma}\mathbf{Q}\mathbf{X}^{\mathrm{H}}\mathbf{Q}^{\mathrm{H}}\mathbf{\Omega}^{\mathrm{H}})^{\mathrm{H}}) = \mathbf{\Gamma}\mathtt{vec}(\mathbf{\Omega}\mathbf{Q}\mathbf{X}\mathbf{Q}^{\mathrm{H}}\mathbf{\Sigma}^{\mathrm{H}}).$$

□

Bibliography

[1] R. Abgrall, D. Amsallem, and R. Crisovan. “Robust model reduction by L_1-norm minimization and approximation via dictionaries: application to nonlinear hyperbolic problems”. *Advanced Modeling and Simulation in Engineering Sciences* 3.1 (2016), pp. 1–16 (cit. on p. 14).

[2] D. Achlioptas. “Database-friendly random projections: Johnson-Lindenstrauss with binary coins”. *Journal of computer and System Sciences* 66.4 (2003), pp. 671–687 (cit. on pp. 16, 19, 29, 42, 51, 76).

[3] N. Ailon and E. Liberty. “Fast dimension reduction using Rademacher series on dual BCH codes”. *Discrete & Computational Geometry* 42.4 (2009), pp. 615–630 (cit. on pp. 16, 42, 49).

[4] H. Al Daas, L. Grigori, P. Hénon, and P. Ricoux. “Recycling Krylov subspaces and reducing deflation subspaces for solving sequence of linear systems” (2018) (cit. on p. 6).

[5] A. Alla and J. N. Kutz. “Randomized model order reduction”. *Advances in Computational Mathematics* 45.3 (2016), pp. 1251–1271. ISSN: 1572-9044. DOI: 10.1007/s10444-018-09655-9. URL: https://doi.org/10.1007/s10444-018-09655-9 (cit. on pp. 15, 21, 23, 30, 31).

[6] B. Almroth, P. Stern, and F. A. Brogan. “Automatic choice of global shape functions in structural analysis”. *AIAA Journal* 16.5 (1978), pp. 525–528 (cit. on p. 6).

[7] D. Amsallem, C. Farhat, and M. Zahr. “On the robustness of residual minimization for constructing POD-based reduced-order CFD models”. *21st AIAA Computational Fluid Dynamics Conference* (2013) (cit. on p. 91).

[8] D. Amsallem and B. Haasdonk. “PEBL-ROM: Projection-error based local reduced-order models”. *Advanced Modeling and Simulation in Engineering Sciences* 3.1 (2016), p. 6 (cit. on p. 103).

[9] I. Babuška, F. Nobile, and R. Tempone. “A stochastic collocation method for elliptic partial differential equations with random input data”. *SIAM Journal on Numerical Analysis* 45.3 (2007), pp. 1005–1034 (cit. on p. 4).

[10] I. Babuška, R. Tempone, and G. E. Zouraris. “Galerkin finite element approximations of stochastic elliptic partial differential equations”. *SIAM Journal on Numerical Analysis* 42.2 (2004), pp. 800–825 (cit. on p. 4).

[11] C. G. Baker, K. A. Gallivan, and P. Van Dooren. “Low-rank incremental methods for computing dominant singular subspaces”. *Linear Algebra and its Applications* 436.8 (2012), pp. 2866–2888 (cit. on p. 31).

[12] O. Balabanov and A. Nouy. “Randomized linear algebra for model reduction. Part I: Galerkin methods and error estimation”. *arXiv preprint arXiv:1803.02602* (2019) (cit. on pp. 15, 18, 19, 20, 21, 27, 96, 152, 157, 163).

[13] O. Balabanov and A. Nouy. “Randomized linear algebra for model reduction. Part II: minimal residual methods and dictionary-based approximation”. *arXiv preprint arXiv:1910.14378* (2019) (cit. on pp. 13, 15, 18, 21, 89).

[14] E. Balmes. “Parametric families of reduced finite element models. Theory and applications”. *Mechanical Systems and Signal Processing* 10.4 (1996), pp. 381–394 (cit. on p. 6).

[15] R. Baraniuk, M. Davenport, R. A. DeVore, and M. Wakin. “A simple proof of the restricted isometry property for random matrices”. *Constructive Approximation* 28.3 (2008), pp. 253–263 (cit. on p. 16).

[16] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera. “An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations”. *Comptes Rendus Mathématique* 339.9 (2004), pp. 667–672 (cit. on p. 10).

[17] A. Barrett and G. Reddien. “On the reduced basis method”. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* 75.7 (1995), pp. 543–549 (cit. on p. 6).

[18] V. Barthelmann, E. Novak, and K. Ritter. “High dimensional polynomial interpolation on sparse grids”. *Advances in Computational Mathematics* 12.4 (2000), pp. 273–288 (cit. on p. 4).

[19] M. Bebendorf. *Hierarchical matrices*. Springer, 2008 (cit. on p. 48).

[20] M. Bebendorf. “Why finite element discretizations can be factored by triangular hierarchical matrices”. *SIAM Journal on Numerical Analysis* 45.4 (2007), pp. 1472–1494 (cit. on p. 48).

[21] S. Becker, J. Bobin, and E. J. Candès. “NESTA: A fast and accurate first-order method for sparse recovery”. *SIAM Journal on Imaging Sciences* 4.1 (2011), pp. 1–39 (cit. on p. 16).

[22] R. E. Bellman. *Adaptive control processes: a guided tour*. Vol. 42. 7-8. Princeton university press, 2015, pp. 364–365. DOI: 10.1002/zamm.19620420718 (cit. on p. 4).

[23] R. E. Bellman. “Dynamic programming”. *Science* 153.3731 (1966), pp. 34–37 (cit. on p. 4).

[24] P. Benner, A. Cohen, M. Ohlberger, and K. Willcox, eds. *Model Reduction and Approximation: Theory and Algorithms*. SIAM, Philadelphia, PA, 2017 (cit. on pp. 3, 4, 29, 91, 151).

[25] P. Benner, S. Gugercin, and K. Willcox. "A survey of projection-based model reduction methods for parametric dynamical systems". *SIAM review* 57.4 (2015), pp. 483–531 (cit. on pp. 3, 4, 6, 29, 151).

[26] G. Berkooz, P. Holmes, and J. L. Lumley. "The proper orthogonal decomposition in the analysis of turbulent flows". *Annual Review of Fluid Mechanics* 25.1 (1993), pp. 539–575 (cit. on p. 6).

[27] K. Bertoldi, V. Vitelli, J. Christensen, and M. van Hecke. "Flexible mechanical metamaterials". *Nature Reviews Materials* 2.11 (2017), p. 17066 (cit. on p. 125).

[28] P. Binev, A. Cohen, W. Dahmen, R. A. DeVore, G. Petrova, and P. Wojtaszczyk. "Convergence rates for greedy algorithms in reduced basis methods". *SIAM Journal on Mathematical Analysis* 43.3 (2011), pp. 1457–1472 (cit. on pp. 8, 14).

[29] E. G. Boman, B. Hendrickson, and S. Vavasis. "Solving elliptic finite element systems in near-linear time with support preconditioners". *SIAM Journal on Numerical Analysis* 46.6 (2008), pp. 3264–3284 (cit. on p. 48).

[30] J. Bourgain, J. Lindenstrauss, and V. Milman. "Approximation of zonoids by zonotopes". *Acta mathematica* 162.1 (1989), pp. 73–141 (cit. on p. 76).

[31] C. Boutsidis and A. Gittens. "Improved matrix algorithms via the subsampled randomized Hadamard transform". *SIAM Journal on Matrix Analysis and Applications* 34.3 (2013), pp. 1301–1340 (cit. on pp. 16, 49, 78, 85).

[32] C. Boutsidis, D. P. Woodruff, and P. Zhong. "Optimal principal component analysis in distributed and streaming models". *Proceedings of the 48th annual ACM symposium on Theory of Computing*. ACM. 2016, pp. 236–249 (cit. on p. 23).

[33] T. Braconnier, M. Ferrier, J.-C. Jouhaud, M. Montagnac, and P. Sagaut. "Towards an adaptive POD/SVD surrogate model for aeronautic design". *Computers & Fluids* 40.1 (2011), pp. 195–209 (cit. on pp. 15, 23, 30).

[34] S. Brugiapaglia, F. Nobile, S. Micheletti, and S. Perotto. "A theoretical study of COmpRessed SolvING for advection-diffusion-reaction problems". *Mathematics of Computation* 87.309 (2018), pp. 1–38 (cit. on p. 21).

[35] A. Buffa, Y. Maday, A. T. Patera, C. Prud'homme, and G. Turinici. "A priori convergence of the greedy algorithm for the parametrized reduced basis method". *ESAIM: Mathematical Modelling and Numerical Analysis* 46.3 (2012), pp. 595–603 (cit. on pp. 8, 14).

[36] A. Buhr, C. Engwer, M. Ohlberger, and S. Rave. “A numerically stable a posteriori error estimator for reduced basis approximations of elliptic equations”. *arXiv preprint arXiv:1407.8005* (2014) (cit. on p. 50).

[37] A. Buhr and K. Smetana. “Randomized Local Model Order Reduction”. *SIAM Journal on Scientific Computing* 40 (2018), pp. 2120–2151. DOI: 10.1137/17M1138480 (cit. on pp. 21, 30, 92, 152).

[38] T. Bui-Thanh, K. Willcox, and O. Ghattas. “Model reduction for large-scale systems with high-dimensional parametric input space”. *SIAM Journal on Scientific Computing* 30.6 (2008), pp. 3270–3288 (cit. on pp. 13, 91).

[39] N. Cagniart, Y. Maday, and B. Stamm. “Model order reduction for problems with large convection effects”. *Contributions to Partial Differential Equations and Applications*. Springer, 2019, pp. 131–150 (cit. on pp. 14, 15).

[40] E. J. Candes and T. Tao. “Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?” *IEEE Transactions on Information Theory* 52.12 (Dec. 2006), pp. 5406–5425. ISSN: 1557-9654. DOI: 10.1109/TIT.2006.885507 (cit. on p. 16).

[41] C. Canuto, T. Tonn, and K. Urban. “A posteriori error analysis of the reduced basis method for nonaffine parametrized nonlinear PDEs”. *SIAM Journal on Numerical Analysis* 47.3 (2009), pp. 2001–2022 (cit. on p. 11).

[42] K. Carlberg, C. Bou-Mosleh, and C. Farhat. “Efficient non-linear model reduction via a least-squares Petrov–Galerkin projection and compressive tensor approximations”. *International Journal for Numerical Methods in Engineering* 86.2 (2011), pp. 155–181 (cit. on p. 15).

[43] F. Casenave, A. Ern, and T. Lelièvre. “Accurate and online-efficient evaluation of the a posteriori error bound in the reduced basis method”. *ESAIM: Mathematical Modelling and Numerical Analysis* 48.1 (2014), pp. 207–229 (cit. on pp. 15, 50).

[44] T. F. Chan and J. J. Shen. *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*. Vol. 94. SIAM, 2005 (cit. on p. 4).

[45] S. Chaturantabut and D. C. Sorensen. “Nonlinear model reduction via discrete empirical interpolation”. *SIAM Journal on Scientific Computing* 32.5 (2010), pp. 2737–2764 (cit. on pp. 11, 15).

[46] H. Chen and C. T. Chan. “Acoustic cloaking and transformation acoustics”. *Journal of Physics D: Applied Physics* 43.11 (2010), p. 113001 (cit. on p. 124).

[47] Y. Chen, S. Gottlieb, and Y. Maday. “Parametric analytical preconditioning and its applications to the reduced collocation methods”. *Comptes Rendus Mathématique* 352.7-8 (2014), pp. 661–666 (cit. on p. 152).

[48] Y. Cheng, F. Yang, J. Y. Xu, and X. J. Liu. “A multilayer structured acoustic cloak with homogeneous isotropic materials”. *Applied Physics Letters* 92.15 (2008), p. 151913 (cit. on p. 124).

[49] A. Chkifa, A. Cohen, and C. Schwab. “Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs”. *Journal de Mathématiques Pures et Appliquées* 103.2 (2015), pp. 400–428 (cit. on p. 4).

[50] K. L. Clarkson and D. P. Woodruff. “Low-rank approximation and regression in input sparsity time”. *Journal of the ACM (JACM)* 63.6 (2017), p. 54 (cit. on p. 16).

[51] K. L. Clarkson and D. P. Woodruff. “Numerical linear algebra in the streaming model”. *Proceedings of the forty-first annual ACM symposium on Theory of computing*. ACM. 2009, pp. 205–214 (cit. on p. 23).

[52] B. Cockburn, G. E. Karniadakis, and C.-W. Shu. *Discontinuous Galerkin methods: theory, computation and applications*. Vol. 11. Springer Science & Business Media, 2012 (cit. on p. 3).

[53] A. Cohen, W. Dahmen, and R. A. DeVore. “Compressed sensing and best k-term approximation”. *Journal of the American mathematical society* 22.1 (2009), pp. 211–231 (cit. on p. 16).

[54] A. Cohen, W. Dahmen, and R. A. DeVore. “Orthogonal matching pursuit under the restricted isometry property”. *Constructive Approximation* 45.1 (2017), pp. 113–127 (cit. on p. 16).

[55] A. Cohen and R. A. DeVore. “Approximation of high-dimensional parametric PDEs”. *Acta Numerica* 24 (2015), pp. 1–159 (cit. on p. 4).

[56] A. Cohen and R. A. DeVore. “Kolmogorov widths under holomorphic mappings”. *IMA Journal of Numerical Analysis* 36.1 (2015), pp. 1–12 (cit. on p. 14).

[57] A. Cohen, R. A. DeVore, and C. Schwab. “Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs”. *Analysis and Applications* 9.01 (2011), pp. 11–47 (cit. on p. 14).

[58] M. B. Cohen, J. Nelson, and D. P. Woodruff. “Optimal approximate matrix product in terms of stable rank”. *arXiv preprint arXiv:1507.02268* (2015) (cit. on p. 16).

[59] W. Dahmen, C. Plesken, and G. Welper. “Double greedy algorithms: Reduced basis methods for transport dominated problems”. *ESAIM: Mathematical Modelling and Numerical Analysis* 48.3 (2014), pp. 623–663 (cit. on p. 13).

[60] C. Desceliers, R. Ghanem, and C. Soize. “Polynomial chaos representation of a stochastic preconditioner”. *International journal for numerical methods in engineering* 64.5 (2005), pp. 618–634 (cit. on p. 152).

[61] R. A. DeVore. “Nonlinear approximation and its applications”. *Multiscale, Nonlinear and Adaptive Approximation*. Springer, 2009, pp. 169–201 (cit. on p. 106).

[62] R. A. DeVore. “The theoretical foundation of reduced basis methods”. *Model Reduction and approximation: Theory and Algorithms* (2014), pp. 137–168 (cit. on p. 14).

[63] R. A. DeVore and A. Kunoth. *Multiscale, nonlinear and adaptive approximation*. Springer, 2009 (cit. on p. 4).

[64] M. Dihlmann, S. Kaulmann, and B. Haasdonk. “Online reduced basis construction procedure for model reduction of parametrized evolution systems”. *IFAC Proceedings Volumes* 45.2 (2012), pp. 112–117 (cit. on pp. 13, 14, 15, 94, 110).

[65] P. Drineas, M. Magdon-Ismail, M. W. Mahoney, and D. P. Woodruff. “Fast approximation of matrix coherence and statistical leverage”. *Journal of Machine Learning Research* 13 (2012), pp. 3475–3506 (cit. on p. 16).

[66] P. Drineas and M. W. Mahoney. “Randomized numerical linear algebra”. *Communications of the ACM* 59.6 (2016), pp. 80–90 (cit. on p. 16).

[67] M. Drohmann, B. Haasdonk, and M. Ohlberger. “Reduced basis approximation for nonlinear parametrized evolution equations based on empirical operator interpolation”. *SIAM Journal on Scientific Computing* 34.2 (2012), pp. 937–969 (cit. on pp. 11, 15).

[68] J. L. Eftang, D. J. Knezevic, and A. T. Patera. “An hp certified reduced basis method for parametrized parabolic partial differential equations”. *Mathematical and Computer Modelling of Dynamical Systems* 17.4 (2011), pp. 395–422 (cit. on pp. 12, 93, 102).

[69] J. L. Eftang, A. T. Patera, and E. M. Rønquist. “An ”hp” certified reduced basis method for parametrized elliptic partial differential equations”. *SIAM Journal on Scientific Computing* 32.6 (2010), pp. 3170–3200 (cit. on pp. 12, 93, 102).

[70] H. C. Elman, D. J. Silvester, and A. J. Wathen. *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics*. Numerical Mathematics and Scientific Computation, 2014 (cit. on p. 48).

[71] B. Engquist and L. Ying. “Sweeping preconditioner for the Helmholtz equation: hierarchical matrix representation”. *Communications on pure and applied mathematics* 64.5 (2011), pp. 697–735 (cit. on p. 48).

[72] R. Everson and L. Sirovich. “Karhunen–Loeve procedure for gappy data”. *JOSA A* 12.8 (1995), pp. 1657–1664 (cit. on p. 11).

[73] L. Fick, Y. Maday, A. T. Patera, and T. Taddei. “A reduced basis technique for long-time unsteady turbulent flows”. *arXiv preprint arXiv:1710.03569* (2017) (cit. on p. 14).

[74] A. Frieze, R. Kannan, and S. Vempala. “Fast Monte-Carlo algorithms for finding low-rank approximations”. *Journal of the ACM* 51.6 (2004), pp. 1025–1041 (cit. on p. 16).

[75] F. Fritzen, B. Haasdonk, D. Ryckelynck, and S. Schöps. “An algorithmic comparison of the Hyper-Reduction and the Discrete Empirical Interpolation Method for a nonlinear thermal problem”. *Mathematical and Computational Applications* 23.1 (2018), pp. 1–25 (cit. on p. 12).

[76] R. G. Ghanem and P. D. Spanos. *Stochastic finite elements: a spectral approach.* Courier Corporation, 2003 (cit. on p. 4).

[77] M. A. Grepl, Y. Maday, N. C. Nguyen, and A. T. Patera. “Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations”. *ESAIM: Mathematical Modelling and Numerical Analysis* 41.3 (2007), pp. 575–605 (cit. on p. 11).

[78] D. Gross and V. Nesme. “Note on sampling without replacing from a finite collection of matrices”. *arXiv preprint arXiv:1001.2738* (2010) (cit. on p. 86).

[79] S. Gugercin and A. C. Antoulas. “A survey of model reduction by balanced truncation and some new results”. *International Journal of Control* 77.8 (2004), pp. 748–766 (cit. on p. 6).

[80] L. Gui-rong. *Smoothed particle hydrodynamics: a meshfree particle method.* World Scientific, 2003 (cit. on p. 3).

[81] B. Haasdonk. “Convergence rates of the POD–greedy method”. *ESAIM: Mathematical Modelling and Numerical Analysis* 47.3 (2013), pp. 859–873 (cit. on p. 14).

[82] B. Haasdonk. “Reduced basis methods for parametrized PDEs – A tutorial introduction for stationary and instationary problems”. *Model reduction and approximation: theory and algorithms* 15 (2017), p. 65 (cit. on pp. 6, 9, 34, 50, 57, 99).

[83] B. Haasdonk, M. Dihlmann, and M. Ohlberger. “A training set and multiple bases generation approach for parameterized model reduction based on adaptive grids in parameter space”. *Mathematical and Computer Modelling of Dynamical Systems* 17.4 (2011), pp. 423–442 (cit. on p. 13).

[84] B. Haasdonk and M. Ohlberger. “Reduced basis method for explicit finite volume approximations of nonlinear conservation laws” (2008) (cit. on p. 11).

[85] B. Haasdonk, M. Ohlberger, and G. Rozza. “A reduced basis method for evolution schemes with parameter-dependent explicit operators”. *Electronic Transactions on Numerical Analysis* 32 (2008), pp. 145–161 (cit. on pp. 6, 11).

[86] W. Hackbusch. *Hierarchical matrices: algorithms and analysis*. Vol. 49. Springer, 2015 (cit. on p. 48).

[87] N. Halko, P.-G. Martinsson, Y. Shkolnisky, and M. Tygert. “An algorithm for the principal component analysis of large data sets”. *SIAM Journal on Scientific computing* 33.5 (2011), pp. 2580–2594 (cit. on p. 16).

[88] N. Halko, P.-G. Martinsson, and J. A. Tropp. “Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions”. *SIAM review* 53.2 (2011), pp. 217–288 (cit. on pp. 15, 16, 18, 23, 29, 31, 40, 50, 56, 96, 97).

[89] J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*. 1st ed. Springer Briefs in Mathematics. Switzerland: Springer, 2015, p. 135. ISBN: 978-3-319-22469-5. DOI: `10.1007/978-3-319-22470-1` (cit. on pp. 3, 4, 29, 151).

[90] C. Himpe, T. Leibner, and S. Rave. “Hierarchical approximate proper orthogonal decomposition”. *SIAM Journal on Scientific Computing* 40.5 (2018), pp. 3267–3292 (cit. on pp. 15, 23, 30).

[91] A. Hochman, J. F. Villena, A. G. Polimeridis, L. M. Silveira, J. K. White, and L. Daniel. “Reduced-order models for electromagnetic scattering problems”. *IEEE Transactions on Antennas and Propagation* 62.6 (2014), pp. 3150–3162 (cit. on pp. 15, 21, 23, 30).

[92] T. J. Hughes. *The finite element method: linear static and dynamic finite element analysis*. Courier Corporation, 2012 (cit. on p. 3).

[93] D. B. P. Huynh, G. Rozza, S. Sen, and A. T. Patera. “A successive constraint linear optimization method for lower bounds of parametric coercivity and inf–sup stability constants”. *Comptes Rendus Mathématique* 345.8 (2007), pp. 473–478 (cit. on pp. 9, 99, 160).

[94] P. Indyk and R. Motwani. “Approximate nearest neighbors: towards removing the curse of dimensionality”. *Proceedings of the thirtieth annual ACM symposium on Theory of computing*. ACM. 1998, pp. 604–613 (cit. on p. 16).

[95] W. B. Johnson and J. Lindenstrauss. “Extensions of Lipschitz mappings into a Hilbert space”. *Contemporary mathematics* 26.189-206 (1984), p. 1 (cit. on p. 16).

[96] M. Kadic, T. Bückmann, R. Schittny, P. Gumbsch, and M. Wegener. “Pentamode metamaterials with independently tailored bulk modulus and mass density”. *Physical Review Applied* 2.5 (2014), p. 054007 (cit. on p. 125).

[97] S. Kaulmann and B. Haasdonk. “Online greedy reduced basis construction using dictionaries”. *VI International Conference on Adaptive Modeling and Simulation (ADMOS 2013)*. 2013, pp. 365–376 (cit. on pp. 13, 14, 94, 110).

[98] M. E. Kilmer and E. De Sturler. “Recycling subspace information for diffuse optical tomography”. *SIAM Journal on Scientific Computing* 27.6 (2006), pp. 2140–2166 (cit. on p. 6).

[99] D. J. Knezevic and J. W. Peterson. “A high-performance parallel implementation of the certified reduced basis method”. *Computer Methods in Applied Mechanics and Engineering* 200.13 (2011), pp. 1455–1466 (cit. on pp. 15, 23, 32).

[100] D. Kressner, M. Plešinger, and C. Tobler. “A preconditioned low-rank CG method for parameter-dependent Lyapunov matrix equations”. *Numerical Linear Algebra with Applications* 21.5 (2014), pp. 666–684 (cit. on p. 152).

[101] K. Kunisch and S. Volkwein. “Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics”. *SIAM Journal on Numerical analysis* 40.2 (2002), pp. 492–515 (cit. on p. 6).

[102] K. Kunisch and S. Volkwein. “Galerkin proper orthogonal decomposition methods for parabolic problems”. *Numerische Mathematik* 90.1 (2001), pp. 117–148 (cit. on p. 6).

[103] E. Kushilevitz, R. Ostrovsky, and Y. Rabani. “Efficient search for approximate nearest neighbor in high dimensional spaces”. *SIAM Journal on Computing* 30.2 (2000), pp. 457–474 (cit. on p. 16).

[104] L. Le Magoarou and R. Gribonval. “Flexible multilayer sparse approximations of matrices and applications”. *IEEE Journal of Selected Topics in Signal Processing* 10.4 (2016), pp. 688–700 (cit. on pp. 140, 143).

[105] Y. T. Lee and A. Sidford. “Efficient Accelerated Coordinate Descent Methods and Faster Algorithms for Solving Linear Systems”. *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*. Oct. 2013, pp. 147–156. DOI: 10.1109/FOCS.2013.24 (cit. on p. 48).

[106] Y. Maday, N. C. Nguyen, A. T. Patera, and G. S. Pau. “A general, multipurpose interpolation procedure: the magic points”. *Communications on Pure & Applied Analysis* 8 (2009), p. 383. URL: http://aimsciences.org//article/id/30a3894d-b0c8-4e29-8b8d-2611be32876f (cit. on pp. 10, 47).

[107] Y. Maday, A. T. Patera, and D. V. Rovas. “A blackbox reduced-basis output bound method for noncoercive linear problems”. *Studies in Mathematics and Applications* 31 (2002) (cit. on p. 13).

[108] Y. Maday, A. T. Patera, and G. Turinici. “A priori convergence theory for reduced-basis approximations of single-parameter elliptic partial differential equations”. *Journal of Scientific Computing* 17.1-4 (2002), pp. 437–446 (cit. on p. 14).

[109] Y. Maday and B. Stamm. “Locally adaptive greedy approximations for anisotropic parameter reduced basis spaces”. *SIAM Journal on Scientific Computing* 35.6 (2013), pp. 2417–2441 (cit. on pp. 13, 103).

[110] M. W. Mahoney et al. “Randomized algorithms for matrices and data”. *Foundations and Trends® in Machine Learning* 3.2 (2011), pp. 123–224 (cit. on pp. 16, 91).

[111] P.-G. Martinsson. “A fast direct solver for a class of elliptic partial differential equations”. *Journal of Scientific Computing* 38.3 (2009), pp. 316–330 (cit. on p. 48).

[112] B. McWilliams, G. Krummenacher, M. Lucic, and J. M. Buhmann. “Fast and robust least squares estimation in corrupted linear models”. *Advances in Neural Information Processing Systems*. 2014, pp. 415–423 (cit. on p. 15).

[113] B. Moore. “Principal component analysis in linear systems: Controllability, observability, and model reduction”. *IEEE transactions on automatic control* 26.1 (1981), pp. 17–32 (cit. on p. 6).

[114] F. Negri, A. Manzoni, and D. Amsallem. “Efficient model reduction of parametrized systems by matrix discrete empirical interpolation”. *Journal of Computational Physics* 303 (2015), pp. 431–454 (cit. on p. 12).

[115] N. C. Nguyen, A. T. Patera, and J. Peraire. “A ’best points’ interpolation method for efficient approximation of parametrized functions”. *International journal for numerical methods in engineering* 73.4 (2008), pp. 521–543 (cit. on p. 11).

[116] F. Nobile, R. Tempone, and C. G. Webster. “An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data”. *SIAM Journal on Numerical Analysis* 46.5 (2008), pp. 2411–2442 (cit. on p. 4).

[117] A. K. Noor. “On making large nonlinear problems small”. *Computer methods in applied mechanics and engineering* 34.1-3 (1982), pp. 955–985 (cit. on p. 6).

[118] A. K. Noor. “Recent advances in reduction methods for nonlinear problems”. *Computational Methods in Nonlinear Structural and Solid Mechanics*. Elsevier, 1981, pp. 31–44 (cit. on p. 6).

[119] A. Nouy. “Low-rank tensor methods for model order reduction”. *Handbook of Uncertainty Quantification* (2017), pp. 857–882 (cit. on p. 4).

[120] M. Ohlberger and S. Rave. “Nonlinear reduced basis approximation of parameterized evolution equations via the method of freezing”. *Comptes Rendus Mathématique* 351.23-24 (2013), pp. 901–906 (cit. on pp. 14, 15).

[121] M. Ohlberger and S. Rave. “Reduced Basis Methods: Success, Limitations and Future Challenges”. *Proceedings of the Conference Algoritmy* (2016), pp. 1–12 (cit. on p. 14).

[122] G. M. Oxberry, T. Kostova-Vassilevska, W. Arrighi, and K. Chand. “Limited-memory adaptive snapshot selection for proper orthogonal decomposition”. *International Journal for Numerical Methods in Engineering* 109.2 (2017), pp. 198–217 (cit. on pp. 15, 23, 30).

[123] C. H. Papadimitriou, P. Raghavan, H. Tamaki, and S. Vempala. “Latent semantic indexing: A probabilistic analysis”. *Journal of Computer and System Sciences* 61.2 (2000), pp. 217–235 (cit. on p. 16).

[124] M. L. Parks, E. De Sturler, G. Mackey, D. D. Johnson, and S. Maiti. “Recycling Krylov subspaces for sequences of linear systems”. *SIAM Journal on Scientific Computing* 28.5 (2006), pp. 1651–1674 (cit. on pp. 6, 151).

[125] A. T. Patera, G. Rozza, et al. “Reduced basis approximation and a posteriori error estimation for parametrized partial differential equations” (2007) (cit. on p. 6).

[126] B. Peherstorfer, K. Willcox, and M. Gunzburger. “Survey of multifidelity methods in uncertainty propagation, inference, and optimization”. *SIAM Review* 60.3 (2018), pp. 550–591 (cit. on p. 4).

[127] C. Prud’Homme, D. V. Rovas, K. Veroy, L. Machiels, Y. Maday, A. T. Patera, and G. Turinici. “Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods”. *Journal of Fluids Engineering* 124.1 (2002), pp. 70–80 (cit. on p. 6).

[128] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced basis methods for partial differential equations: an introduction.* Vol. 92. Springer, 2015 (cit. on pp. 3, 4, 29, 34, 151).

[129] A. Quarteroni, G. Rozza, et al. *Reduced order methods for modeling and computational reduction.* Vol. 9. Springer, 2014 (cit. on p. 3).

[130] J. Reiss, P. Schulze, J. Sesterhenn, and V. Mehrmann. “The shifted proper orthogonal decomposition: A mode decomposition for multiple transport phenomena”. *SIAM Journal on Scientific Computing* 40.3 (2018), pp. 1322–1344 (cit. on p. 14).

[131] V. Rokhlin and M. Tygert. “A fast randomized algorithm for overdetermined linear least-squares regression”. *Proceedings of the National Academy of Sciences* 105.36 (2008), pp. 13212–13217 (cit. on pp. 16, 18, 95, 96).

[132] G. Rozza, D. B. P. Huynh, and A. T. Patera. “Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations”. *Archives of Computational Methods in Engineering* 15.3 (2008), pp. 229–275 (cit. on pp. 6, 36).

[133] G. Rozza, H. Malik, N. Demo, M. Tezzele, M. Girfoglio, G. Stabile, and A. Mola. “Advances in Reduced Order Methods for Parametric Industrial Problems in Computational Fluid Dynamics”. *Proceedings of 6th European Conference on Computational Mechanics (ECCM 6) and 7th European Conference on Computational Fluid Dynamics (ECFD 7)* (2018), pp. 59–76 (cit. on p. 3).

[134] G. Rozza and K. Veroy. “On the stability of the reduced basis method for Stokes equations in parametrized domains”. *Computer methods in applied mechanics and engineering* 196.7 (2007), pp. 1244–1260 (cit. on p. 13).

[135] R. Rubinstein, M. Zibulevsky, and M. Elad. “Double sparsity: Learning sparse dictionaries for sparse signal approximation”. *IEEE Transactions on signal processing* 58.3 (2009), pp. 1553–1564 (cit. on pp. 140, 143).

[136] M. Rudelson and R. Vershynin. “Sampling from large matrices: An approach through geometric functional analysis”. *Journal of the ACM (JACM)* 54.4 (2007), p. 21 (cit. on p. 16).

[137] D. Ryckelynck. “A priori hyperreduction method: an adaptive approach”. *Journal of computational physics* 202.1 (2005), pp. 346–366 (cit. on p. 12).

[138] N. D. Santo, S. Deparis, A. Manzoni, and A. Quarteroni. “Multi space reduced basis preconditioners for large-scale parametrized PDEs”. *SIAM Journal on Scientific Computing* 40.2 (2018), A954–A983 (cit. on p. 152).

[139] T. Sarlós. “Improved approximation algorithms for large matrices via random projections”. *47th Annual IEEE Symposium on Foundations of Computer Science*. 2006, pp. 143–152 (cit. on pp. 16, 29, 41).

[140] L. Sirovich. “Turbulence and the dynamics of coherent structures. I. Coherent structures”. *Quarterly of applied mathematics* 45.3 (1987), pp. 561–571 (cit. on p. 37).

[141] K. Smetana, O. Zahm, and A. T. Patera. “Randomized residual-based error estimators for parametrized equations”. *arXiv preprint arXiv:1807.10489* (2018) (cit. on pp. 15, 20, 21, 92, 152, 153, 154).

[142] T. Taddei, S. Perotto, and A. Quarteroni. “Reduced basis techniques for nonlinear conservation laws”. *ESAIM: Mathematical Modelling and Numerical Analysis* 49.3 (2015), pp. 787–814 (cit. on p. 14).

[143] T. Taddei. “An offline/online procedure for dual norm calculations of parameterized functionals: empirical quadrature and empirical test spaces”. *arXiv preprint arXiv:1805.08100* (2018) (cit. on p. 15).

[144] V. N. Temlyakov. "Nonlinear Kolmogorov widths". *Mathematical Notes* 63.6 (1998), pp. 785–795 (cit. on pp. 12, 103).

[145] D. Torlo, F. Ballarin, and G. Rozza. "Stabilized weighted reduced basis methods for parametrized advection dominated problems with random inputs". *SIAM/ASA Journal on Uncertainty Quantification* 6.4 (2018), pp. 1475–1502 (cit. on p. 14).

[146] J. A. Tropp. "Improved analysis of the subsampled randomized Hadamard transform". *Advances in Adaptive Data Analysis* 3.01n02 (2011), pp. 115–126 (cit. on pp. 16, 78, 85, 86).

[147] J. A. Tropp. "User-friendly tail bounds for sums of random matrices". *Foundations of computational mathematics* 12.4 (2012), pp. 389–434 (cit. on pp. 16, 86).

[148] J. A. Tropp et al. "An introduction to matrix concentration inequalities". *Foundations and Trends® in Machine Learning* 8.1-2 (2015), pp. 1–230 (cit. on p. 16).

[149] J. A. Tropp and A. C. Gilbert. "Signal recovery from random measurements via orthogonal matching pursuit". *IEEE Transactions on information theory* 53.12 (2007), pp. 4655–4666 (cit. on pp. 16, 91, 106).

[150] J. A. Tropp, A. Yurtsever, M. Udell, and V. Cevher. "Practical sketching algorithms for low-rank matrix approximation". *SIAM Journal on Matrix Analysis and Applications* 38.4 (2017), pp. 1454–1485 (cit. on p. 16).

[151] R. Vershynin. "Introduction to the non-asymptotic analysis of random matrices". *Compressed Sensing: Theory and Applications* (2010) (cit. on p. 16).

[152] H. K. Versteeg and W. Malalasekera. *An introduction to computational fluid dynamics: the finite volume method*. Pearson education, 2007 (cit. on p. 3).

[153] S. Voronin and P.-G. Martinsson. "RSVDPACK: An implementation of randomized algorithms for computing the singular value, interpolative, and CUR decompositions of matrices on multi-core and GPU architectures". *arXiv preprint arXiv:1502.05366* (2015) (cit. on pp. 15, 23).

[154] S. Wang, E. d. Sturler, and G. H. Paulino. "Large-scale topology optimization using preconditioned Krylov subspace methods with recycling". *International journal for numerical methods in engineering* 69.12 (2007), pp. 2441–2468 (cit. on p. 6).

[155] G. Welper. "Transformed Snapshot Interpolation with High Resolution Transforms". *arXiv preprint arXiv:1901.01322* (2019) (cit. on p. 15).

[156] K. Willcox. "Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition". *Computers & fluids* 35.2 (2006), pp. 208–226 (cit. on p. 15).

[157] D. P. Woodruff et al. “Sketching as a tool for numerical linear algebra”. *Foundations and Trends® in Theoretical Computer Science* 10.1–2 (2014), pp. 1–157 (cit. on pp. 15, 16, 18, 20, 23, 29, 38, 40, 41, 54, 76, 91, 96, 163).

[158] J. Xia, S. Chandrasekaran, M. Gu, and X. S. Li. “Superfast multifrontal method for large structured linear systems of equations”. *SIAM Journal on Matrix Analysis and Applications* 31.3 (2009), pp. 1382–1411 (cit. on p. 48).

[159] J. Yang, X. Meng, and M. W. Mahoney. “Implementing randomized matrix algorithms in parallel and distributed environments”. *Proceedings of the IEEE* 104.1 (2015), pp. 58–92 (cit. on p. 23).

[160] O. Zahm and A. Nouy. “Interpolation of inverse operators for preconditioning parameter-dependent equations”. *SIAM Journal on Scientific Computing* 38.2 (2016), pp. 1044–1074 (cit. on pp. 13, 14, 20, 25, 30, 92, 152, 153, 162, 166).

Titre : Algèbre linéaire randomisée pour la réduction de l'ordre des modèles

Mot clés : réduction de modèle, projections aléatoires, préconditionneur

Résumé : Cette thèse introduit des nouvelles approches basées sur l'algèbre linéaire aléatoire pour améliorer l'efficacité et la stabilité des méthodes de réduction de modèles basées sur des projections pour la résolution d'équations dépendant de paramètres.

Notre méthodologie repose sur des techniques de projections aléatoires ("random sketching") qui consistent à projeter des vecteurs de grande dimension dans un espace de faible dimension. Un modèle réduit est ainsi construit de manière efficace et numériquement stable à partir de projections aléatoires de l'espace d'approximation réduit et des espaces des résidus associés.

Notre approche permet de réaliser des économies de calcul considérables dans pratiquement toutes les architectures de calcul modernes. Par exemple, elle peut réduire le nombre de flops et la consommation de mémoire et améliorer l'efficacité du flux de données (caractérisé par l'extensibilité ou le coût de communication). Elle peut être utilisée pour améliorer l'efficacité et la stabilité des méthodes de projection de Galerkin ou par minimisation de résidu. Elle peut également être utilisée pour estimer efficacement l'erreur et post-traiter la solution du modèle réduit.

De plus, l'approche par projection aléatoire rend viable numériquement une méthode d'approximation basée sur un dictionnaire, où pour chaque valeur de paramètre, la solution est approchée dans un sous-espace avec une base sélectionnée dans le dictionnaire.

Nous abordons également la construction efficace (par projections aléatoires) de préconditionneurs dépendant de paramètres, qui peuvent être utilisés pour améliorer la qualité des projections de Galerkin ou des estimateurs d'erreur pour des problèmes à opérateurs mal conditionnés.

Pour toutes les méthodes proposées, nous fournissons des conditions précises sur les projections aléatoires pour garantir des estimations précises et stables avec une probabilité de succès spécifiée par l'utilisateur. Pour déterminer la taille des matrices aléatoires, nous fournissons des bornes a priori ainsi qu'une procédure adaptative plus efficace basée sur des estimations a posteriori.

Title: Randomized linear algebra for model order reduction

Keywords: model reduction, random sketching, subspace embedding, sparse approximation, preconditioner

Abstract: Solutions to high-dimensional parameter-dependent problems are in great demand in the contemporary applied science and engineering. The standard approximation methods for parametric equations can require computational resources that are exponential in the dimension of the parameter space, which is typically referred to as the curse of dimensionality. To break the curse of dimensionality one has to appeal to nonlinear methods that exploit the structure of the solution map, such as projection-based model order reduction methods.

This thesis proposes novel methods based on randomized linear algebra to enhance the efficiency and stability of projection-based model order reduction methods for solving parameter-dependent equations. Our methodology relies on random projections (or random sketching). Instead of operating with high-dimensional vectors we first efficiently project them into a low-dimensional space. The reduced model is then efficiently and numerically stably constructed from the projections of the reduced approximation space and the spaces of associated residuals.

Our approach allows drastic computational savings in basically any modern computational architecture. For instance, it can reduce the number of flops and memory consumption and improve the efficiency of the data flow (characterized by scalability or communication costs). It can be employed for improving the efficiency and numerical stability of classical Galerkin and minimal residual methods. It can also be used for the efficient estimation of the error, and post-processing of the solution of the reduced order model. Furthermore, random sketching makes computationally feasible a dictionary-based approximation method, where for each parameter value the solution is approximated in a subspace with a basis selected from a dictionary of vectors. We also address the efficient construction (using random sketching) of parameter-dependent preconditioners that can be used to improve the quality of Galerkin projections or for effective error certification for problems with ill-conditioned operators. For all proposed methods we provide precise conditions on the random sketch to guarantee accurate and stable estimations with a user-specified probability of success. A priori estimates to determine the sizes of the random matrices are provided as well as a more effective adaptive procedure based on a posteriori estimates.