Hierarchical methods for deterministic and stochastic partial differential equations

Paul Mycek

▶ **To cite this version:**

Paul Mycek. Hierarchical methods for deterministic and stochastic partial differential equations. Numerical Analysis [math.NA]. Université de Bordeaux, 2024. tel-04521927

HAL Id: tel-04521927
https://hal.science/tel-04521927v1

Submitted on 26 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

université de BORDEAUX Inria CERFACS
CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

Habilitation à diriger des recherches

Université de Bordeaux
École Doctorale de Mathématiques et d'Informatique
Spécialité : Mathématiques Appliquées et Calcul Scientifique

Méthodes hiérarchiques pour des équations aux dérivées partielles déterministes et stochastiques

Hierarchical methods for deterministic and stochastic partial differential equations

par

Paul Mycek

Soutenue à Toulouse le 8 février 2024

Composition du jury

Laurent Debreu	Directeur de recherche	Inria	Rapporteur
Christophe Prud'homme	Professeur	Univ. Strasbourg	Rapporteur
Robert Scheichl	Professeur	Univ. Heidelberg	Rapporteur
Nathalie Bartoli	Directrice de recherche	Onera	Examinatrice
Grégory Pinon	Professeur	Univ. Le Havre Normandie	Examinateur
Carmen Rodrigo	Professeure associée	Univ. Zaragoza	Examinatrice
Ulrich Rüde	Professeur	FAU Erlangen-Nürnberg	Examinateur
Elisabeth Ullmann	Professeure	TU Munich	Examinatrice
Luc Giraud	Directeur de recherche	Inria	Directeur de recherche

Méthodes hiérarchiques pour des équations aux dérivées partielles déterministes et stochastiques

Paul Mycek

Résumé

Les méthodes hiérarchiques sont devenues un outil essentiel pour la simulation numérique efficace de phénomènes physiques avec une fidélité croissante sur des calculateurs à haute performance modernes. Des méthodes bien établies, telles que les méthodes multigrilles et de décomposition de domaine, se sont avérées performantes, de par leur capacité de passage à l'échelle, pour résoudre certains problèmes de grande taille de manière parallèle sur des supercalculateurs. D'autre part, la propagation des incertitudes dans les simulations numériques a reçu une attention croissante ces dernières années. Pour des simulations onéreuses en temps de calcul de problèmes fortement non-linéaires (par rapport à des paramètres d'entrée incertains), des méthodes avancées de quantification des incertitudes doivent être dévelopées. Dans ce manuscrit, nous exposons des idées méthodologiques et algorithmiques pour répondre à des questions spécifiques liées aux méthodes hiérarchiques. Premièrement, nous présentons des approches de décomposition de domaine résilientes pour la résolution numérique d'équations aux dérivées partielles déterministes et stochastiques, ainsi que des stratégies de Monte Carlo accélérées par des modèles de substitution pour la propagation des incertitudes dans des méthodes de décomposition de domaine. Deuxièmement, nous présentons de nouveaux solveurs multigrilles pour des systèmes linéaires issus de discrétisations hybrides d'ordre élevé. Troisièmement, nous proposons des approches hiérarchiques, plus précisément des méthodes multiniveaux et multifidélité, pour la propagation d'incertitudes dans les simulateurs numériques coûteux. Enfin, nous proposons également des idées pour la résolution numérique efficace de suites de systèmes d'équations linéaires, rencontrées typiquement lors de la résolution de problèmes instationnaires et/ou non linéaires, ou lors de l'échantillonnage d'équations aux dérivées partielles stochastiques.

Hierarchical methods for deterministic and stochastic partial differential equations

Paul Mycek

Abstract

Hierarchical methods have become essential tools for the efficient numerical simulation of physical phenomena with ever-growing fidelity on modern high-performance supercomputers. Well-established methods, such as multigrid and domain decomposition methods, owing to their scaling capabilities, have proven powerful to solve certain classes of large-scale problems in a parallel fashion on supercomputers. On the other hand, the propagation of uncertainties in numerical simulations has received increasing attention in recent years. For computationally demanding simulations of highly non-linear problems (w.r.t. uncertain input parameters), advanced uncertainty quantification methods need to be designed. In this manuscript, we present methodological and algorithmic ideas to address specific questions related to hierarchical methods. First, we present resilient domain decomposition approaches for the numerical solution of deterministic and stochastic partial differential equations, as well as surrogate-assisted Monte Carlo strategies for the propagation of uncertainties in domain decomposition solvers. Second, we introduce novel multigrid solvers for the linear systems arising from hybridizable high-order discretizations. Third, we propose hierarchical approaches, namely multilevel and multifidelity methods, for the propagation of uncertainties in expensive numerical simulators. Finally, we also propose ideas for the efficient numerical solving of sequences of systems of linear equations that typically arise when solving time-dependent and/or non-linear problems, or when sampling stochastic partial differential equations.

Contents

Chapter 1

Introduction

Since the beginning of my scientific career as a researcher, most of my research activities have revolved around high-performance computing (HPC), with a focus on hierarchical methods and algorithms for the numerical solution of deterministic and stochastic partial differential equations (PDEs), for large-scale applications primarily in computational fluid dynamics. With a somewhat loose definition of the term hierarchical[1], we include here domain decomposition methods under this designation.

My early years of research were largely influenced by the challenges that the forthcoming generation of exascale supercomputers was expected to face. In particular, the question of resilience of such extreme scale computers to faults was a major concern for the scientific computing community. Indeed, the fault-coping mechanisms implemented on petascale computer systems at the time were challenged by the anticipated increase in the occurrence of faults with the increased number of transistors on exascale systems [M1]. This context led me to work on the design of resilient algorithms for the numerical solution of (deterministic and stochastic) PDEs, based on a reformulation of standard overlapping domain decomposition approaches and the use of robust regression techniques [M2–M7]. This also gave me the opportunity to work more generally on domain decomposition methods for the propagation of uncertainties in discretized elliptic PDEs, using surrogate models to accelerate the Monte Carlo sampling of such problems [M8–M10].

My research interests have since evolved towards multifidelity techniques for the estimation of statistics of the output of costly numerical simulators, which represents a major part of my current research work [M11–M14]. Multifidelity estimation has gained increasing attention in recent years as a means to tackle high-dimensional uncertainty propagation in complex systems for which high-fidelity sampling or high-fidelity surrogate modeling alone do not yield accurate enough estimates at afford-

[1] for the sake of a more concise manuscript title!

able computational cost. A typical example is the problem of data assimilation for numerical weather prediction applications, where uncertainties need to be propagated through a large-scale, high-dimensional, chaotic numerical model.

The multilevel aspect of certain multifidelity techniques, along with the high-performance computing and numerical linear algebra ecosystem at Cerfacs, naturally prompted me to get involved in multigrid methods. This materialized through the co-supervision of a PhD thesis on the design of efficient multigrid solvers for the hybrid high-order (HHO) discretization [M15–M18].

Tractable uncertainty propagation through discretized PDEs often requires the ability to efficiently solve (possibly massively long) sequences of systems of linear equations with multiple left- and right-hand sides. This led me to work on block iterative solvers and deflation techniques, bringing together the worlds of uncertainty quantification and high-performance numerical linear algebra [M19]. Incidentally, this question of solving sequences of systems had already come up during my PhD [M20] in the context of a time-dependent problem for the modeling of marine current turbine farms, for which a block-Jacobi preconditioner was proposed [M21].

This manuscript is intended as an extended summary of selected parts of my past research. For that reason, technical details are kept to a minimum, with an effort to employ consistent and standard notation, so that readers acquainted with the subject can (hopefully) appreciate the core ideas of each research aspect. The curious reader may, however, find further details in the supporting articles, preprints or reports, which are systematically cited. The manuscript concludes with medium- and long-term research perspectives for the coming years.

Chapter 2

Domain decomposition methods

Domain decomposition (DD) methods [1–4] refer to a class of divide-and-conquer methodologies for the numerical solution of partial differential equations (PDEs). DD methods are based on the partitioning of the domain of the PDE into smaller subdomains, and are, by design, well-suited for parallel computing. This chapter presents two aspects of my research activities related to DD methods.

The first aspect concerns the design of a resilient DD solver for extreme scale computing. Resilience to faults was a major concern for the HPC community in the 2010s, as the upcoming exascale computer systems were expected to have significantly higher fault rates than petascale systems [M1]. It was then anticipated that the mean time between faults would become too short for the fault-coping mechanisms implemented on petascale systems to remain efficient. For example, in the case of the popular checkpoint-restart mechanism, which consists of periodically saving the state of the system and restoring it upon detection of a fault, the time required for checkpointing could be close to the mean time between faults, resulting in the system spending more time saving and restoring states than performing actual computation [5]. Owing to their natural parallel aspect and their ability to scale to many subdomains, DD methods represent a particularly promising candidate framework for solving large-scale discretized elliptic PDEs on exascale computer systems. They are also well-suited to address the question of resilience, in that faults are, to some extent, localized to subdomains. In section 2.1, we present a resilient DD solver based on a reformulation of the original PDE as an optimization problem, exploiting an extreme-scale computing paradigm that states that computation is nearly free, while communication is severely expensive. The proposed algorithm relies on two resilience-enabling programming models formally introduced by [6], namely selective reliability programming and skeptical programming. Although the issue of resilience has now become a secondary concern for the community, it is still closely tied to another critical aspect of extreme-scale computer systems, which is

their energy consumption. Indeed, the energy consumption of a supercomputer may be reduced by lowering its supply voltage at the expense of higher fault rates. In the long run, potentially for the post-exascale era, should it exist, the issue of resilience may resurface in that context, provided that the HPC community embraces a new paradigm of performing scientific computations on unreliable computer systems. Such considerations are briefly discussed in the context of our algorithm in section 2.1.5.

The second aspect concerns the design of DD strategies for the numerical solution of stochastic elliptic PDEs. While a stochastic counterpart to multigrid methods, namely multilevel Monte Carlo (MLMC) methods [7–9], has been introduced and successfully employed for a variety of stochastic problems, few stochastic versions of deterministic DD methods have been proposed to solve stochastic elliptic PDEs [10–14]. The advantage of such DD methods could lie in leveraging some form of locality in the random input parameters to mitigate the curse of dimensionality within individual subdomains, so that local surrogate models can be constructed at affordable cost. In section 2.2, we present stochastic DD strategies based on local Karhunen-Loève (KL) approximations of the random coefficient field of the elliptic PDE and accelerated Monte Carlo sampling based on local polynomial expansions.

2.1 A resilient domain decomposition method for extreme scale computing

For extreme scale computing, and more generally for reliable scientific computations on unreliable computer systems, a major challenge lies in the ability to overcome the occurrence of faults. In this section, we present the ingredients of a resilient DD solver for PDEs of the form

$$\begin{cases} \mathcal{L}u = f \text{ in } \mathcal{D}, \\ u|_{\partial\mathcal{D}} = g. \end{cases} \tag{2.1}$$

We will restrict our exposition to the case where $\mathcal{L}$ is a linear partial differential operator. The more general case is addressed in [M2].

2.1.1 General formulation

The global domain $\mathcal{D}$ is partitioned into N overlapping, open and bounded subdomains $\{\mathcal{D}_i\}_{i=1}^N$ such that $\cup_{i=1}^N \mathcal{D}_i = \mathcal{D}$. We denote by $\partial\mathcal{D}$ (resp. $\partial\mathcal{D}_i$) the boundary of $\mathcal{D}$ (resp. $\mathcal{D}_i$), and by $\Gamma_i^{\text{in}} := \partial\mathcal{D}_i \cap \mathcal{D} = \partial\mathcal{D}_i \backslash \partial\mathcal{D}$ the inner boundary of $\mathcal{D}_i$ (i.e., excluding the boundary $\partial\mathcal{D}$ of the global domain $\mathcal{D}$). We consider a Dirichlet prob-

lem of the form (2.1) in the global domain $\mathcal{D}$, and we focus on the solution $u|_\Gamma$ to this problem at the boundaries of the inner boundaries of the subdomains, defined by $\Gamma^{\text{in}} := \cup_{i=1}^N \Gamma_i^{\text{in}} = (\cup_{i=1}^N \partial\mathcal{D}_i) \cap \mathcal{D} = (\cup_{i=1}^N \partial\mathcal{D}_i) \backslash \partial\mathcal{D}$. Because of the overlapping decomposition, each subdomain $\mathcal{D}_i$ includes a set $\Gamma_i^{\text{neigh}} := (\cup_{j=1}^N \partial\mathcal{D}_j) \cap \mathcal{D}_i = \Gamma^{\text{in}} \cap \mathcal{D}_i$ of boundary parts belonging to neighboring subdomains.

Our approach [M2], which may be seen as an accelerated overlapping Schwarz algorithm, relies on the functions $\mathcal{F}_i$ that map, on each subdomain $\mathcal{D}_i$, local Dirichlet boundary data v_i^Γ to the solution $v|_{\Gamma_i^{\text{neigh}}}$, at the neighboring interfaces Γ_i^{neigh}, of a local Dirichlet problem with boundary condition $v|_{\Gamma_i^{\text{in}}} = v_i^\Gamma$. Specifically,

$$\mathcal{F}_i \colon v_i^\Gamma \mapsto v|_{\Gamma_i^{\text{neigh}}}, \text{ where } v \text{ is such that } \begin{cases} \mathcal{L}v = f \text{ in } \mathcal{D}_i, \\ v|_{\Gamma_i^{\text{in}}} = v_i^\Gamma, \\ v|_{\Gamma_i^{\text{ext}}} = g|_{\Gamma_i^{\text{ext}}}. \end{cases} \tag{2.2}$$

We then seek u^Γ defined on Γ^{in} satisfying compatibility conditions at the interfaces of the subdomains, namely, $\forall i = 1, \ldots, N,\ u^\Gamma|_{\Gamma_i^{\text{in}}} = \mathcal{F}_i u^\Gamma|_{\Gamma_i^{\text{neigh}}}$, which may eventually be recast as a fixed-point problem of the form $u^\Gamma = \mathcal{F}u^\Gamma$. Provided that the original problem (2.1) is well-posed, the solution u^Γ of this reduced problem uniquely matches the solution $u|_\Gamma$ of (2.1) at the inner boundaries. Note that the additive Schwarz method [15] amounts to solving the fixed-point problem by the iterations $u_{(k+1)}^\Gamma = \mathcal{F}u_{(k)}^\Gamma$.

For linear PDEs (i.e., if $\mathcal{L}$ is a linear operator), the boundary-to-boundary mappings $\mathcal{F}_i$ are affine mappings, so that the fixed-point problem can be recast as a trace problem of the form $Su^\Gamma = b^\Gamma$, where b^Γ is defined on Γ^{in} and S is a linear operator. Upon discretization, we obtain a reduced system of linear equations $\mathbf{S}\boldsymbol{u}^\Gamma = \boldsymbol{b}^\Gamma$, where $\boldsymbol{u}^\Gamma$ denotes the vector of solution values at the discretization points along Γ^{in}.

Our approach consists in *learning* the boundary-to-boundary mappings $\mathcal{F}_i$ in a resilient manner to build the fixed-point operator $\mathcal{F}$. The process is illustrated in Figure 2.1. Specifically, we rely on a sampling strategy that involves solving the target PDE equation locally within each subdomain for sampled values of the boundary conditions on that subdomain (stage 4 of Fig. 2.1). Using these input-output sample pairs, we resort to robust regression techniques to find resilient approximations of the mappings from potentially corrupted samples (stage 4 of Fig. 2.1). Finally, the reduced problem $\mathbf{S}\boldsymbol{u}^\Gamma = \boldsymbol{b}^\Gamma$ is assembled and solved (stage 5 of Fig. 2.1). One of the important features of the algorithm is that the construction of the boundary-to-boundary mappings can be done for each subdomain *independently* from all the others. This allows us to satisfy data locality, which is one of the main factors contributing to scalability on extreme scale machines. In that sense, our approach

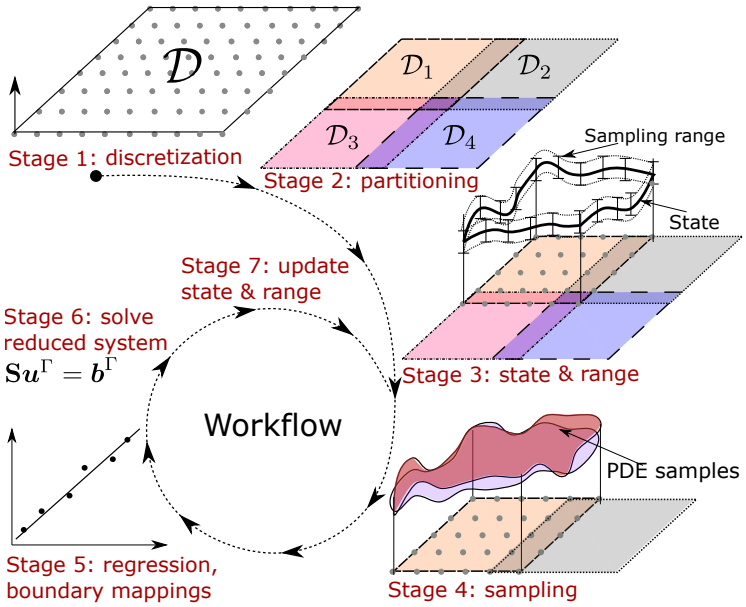


Figure 2.1: Schematic of the workflow of the resilient domain decomposition solver described in [M2]. Adapted from [M3, M4].

differs considerably from the classical iterative overlapping Schwarz algorithms.

In this framework, robust regression techniques are key to achieving resilience. Because least squares regression is known to be very sensitive to outliers (see [M2] for examples in the context of bit-flips), we instead resort to least absolute deviations regression, which is known to be much more robust [16]. In practice, we solve this regression problem using the *iteratively reweighted least squares* algorithm (see, e.g.,[17]).

For linear PDEs, the boundary-to-boundary mappings could theoretically be determined by solving exactly as many independent local PDEs as the number of unknown affine coefficients (i.e., the number of unknown boundary points plus one) on each subdomain $\mathcal{D}_i$. However, because of the potential occurrence of faults, we allow ourselves to slightly oversample the boundary conditions by a factor $\rho > 1$. In other words, on a subdomain $\mathcal{D}_i$ with N_i^Γ unknown boundary points, we define a target number of samples $s_i^* = \rho(N_i^\Gamma + 1)$ to be used for the regression problems on subdomain $\mathcal{D}_i$. The nominal sampling rate ρ thus determines the amount of extra work required to ensure the resilience of the approach.

2.1.2 Server-client implementation and selective reliability model

The algorithm described previously was implemented using a task-based paradigm within a server-client (SC) structure [M3, M5] where MPI processes are grouped into servers and clients.

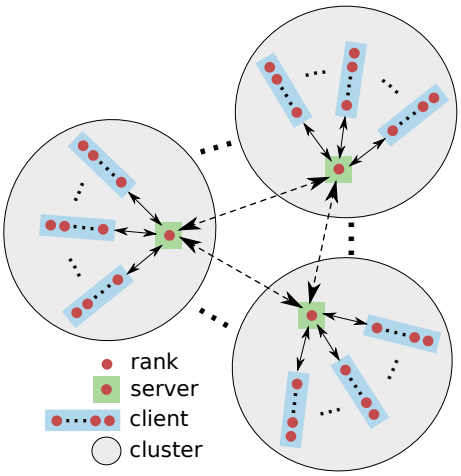


Figure 2.2: Schematic of the server-client implementation. Adapted from [M5].

Our SC model follows the *selective reliability* programming model introduced by [6, 18], which consists of declaring specific data and compute regions to be highly reliable (more than the default reliability level of the system). Specifically, in our model, the servers are assumed to be safe (or "sandboxed") units holding the data, whereas the clients are designed solely to accept and perform work without any assumption on their reliability. Clearly, the SC framework involves substantial data exchange between servers and clients, but a key advantage of this structure relies in its inherent resiliency to hard faults, provided that the MPI framework is made fault-tolerant, e.g., using the *User Level Failure Mitigation* (ULFM) prototype for MPI [19]. Since the actual data is safely held by the servers, the SC is inherently resilient to clients crashing (partial or complete node failures), since this only translates into missing tasks. The asynchronous nature of the SC model is beneficial to reduce the communication wait times.

Figure 2.2 shows a schematic of our SC structure. We adopt a cluster-based model, namely the MPI ranks are grouped into separate clusters, with each cluster containing a server and, for resource balancing purposes, the same number of clients. These clusters are designed such that all servers can communicate between each other, while the clients within a cluster are only visible to the server within that cluster. Moreover, within any given cluster, clients are independent, i.e., they cannot communicate with each other. The data is distributed among the servers, and these are assumed to be highly resilient (safe or under a sandbox model implementation). The sandbox model assumed for the servers can be supported by either hardware or software. The former assumption is supported by hardware specifications on the variable levels of resilience that can be allowed within large computer systems. In the case of software support, a sandbox effect can be accomplished by a programming model relying on data redundancy and strategic synchronization [18, 20, 21].

Since the servers hold the data, they are responsible for generating work in the form of tasks, dispatching them to their pool of available clients, as well as receiving and processing tasks. To optimize communication, clients are designed such that it is their root process that receives a new task to perform. After receiving the task, the root process then broadcasts it to all the other ranks in the client, so that the client as a whole works in parallel to solve the task. This paradigm can be exploited in certain hardware configurations because leveraging local communication within a client is more efficient than having the server communicate a task to all the MPI ranks in a client. One example is the case where all ranks of a client live in the same node, so that they can exploit in-node parallelism and faster memory access. All communications between a server and its clients are implemented with non-blocking operations, allowing us to overlap them on the server side with the computational operations involved in the creation and processing of the tasks.

Due to separation of state from computation, and because of the nature of the algorithm, sampling and regression are executed by the clients through tasks. On the other hand, because the servers hold the state, they are responsible for performing the solve of the reduced system and the subdomains updating. The safety of the servers precludes any potential data corruption during these operations.

This resilient solver was validated on 1D [M2] and 2D [M3, M5] elliptic problems, where tests up to about 110,000 cores showed excellent parallel scalability (greater than 90%) and perfect resilience with limited overhead.

2.1.3 Skeptical programming approach using a priori bounds

Another key model introduced in [6, 22] is the so-called *skeptical* programming model, which consists in acknowledging the unreliable nature of the software execution by being *skeptical* of the returned results. These results may be tested using known mathematical properties of the algorithms, such as theoretical bounds, provided that they are simple and cheap enough not to cause significant overhead. This helps detecting *some* (not all) errors, and guarantees that the remaining (undetected) errors are bounded.

Following this idea, we derived a priori bounds for the discrete solution of second-order elliptic partial differential equations. In [M4], we derive closed form expressions for the case of a conservative, second-order finite difference approximation of the diffusion equation with variable scalar diffusivity. The bounds depend on four parameters, namely:

- the minimum and maximum of the boundary conditions;
- the diffusivity field (and its gradient);
- the source field;

- the size of the domain.

The boundary conditions contribute to a first term in the bounds, which directly relates to the homogeneous PDE, for which the maximum principle applies, where only the boundary conditions are involved. The last three parameters influence a second term in the bounds, which we refer to as the *invariant* part of the bounds, because it does not depend on the boundary values.

The bounds are then incorporated into the resilient domain decomposition framework previously described in sections 2.1.1 and 2.1.2, in order to verify the admissibility of local PDE solutions. Specifically, the sampling stage (stage 4 of Fig. 2.1) is adapted to include the admissibility check, as described in algorithm 1. When a fault is detected, the corresponding sample is simply discarded, and the resilient regression stage (stage 5 of Fig. 2.1) will simply proceed with fewer (namely s_i), but better quality samples than the prescribed number of samples s_i^*. This approach is robust as long as not too many samples get discarded. The computations demonstrate that the bounds are able to detect most system faults, and thus considerably enhance the resilience and the overall performance of the solver. In particular, we demonstrate that this skeptical programming approach allows to reduce the value of the nominal sampling rate ρ (and so the overhead) compared to the original approach in [M2], described in sections 2.1.1 and 2.1.2.

2.1.4 Extension to uncertain elliptic PDEs

In [M7], we extend the deterministic approach of [M2] to uncertain (stochastic) elliptic PDEs. Specifically, we consider the following 1D ($d = 1$) problem:

$$\begin{cases} \mathcal{L}(\boldsymbol{\xi})u = f \text{ in } \mathcal{D} := (0,1), \\ u(x=0) = U_0, \quad u(x=1) = U_1, \end{cases} \tag{2.3}$$

where $\mathcal{L}$ is a linear, elliptic differential operator parameterized by a vector $\boldsymbol{\xi}$ of independent real-valued random variables defined on a probability space $(\Omega, \Sigma, \mathbb{P})$, with prescribed joint distribution $p_{\boldsymbol{\xi}}$. Our approach consists in learning polynomial chaos (PC) surrogates [23–26] of the affine coefficients of the stochastic boundary-to-boundary mappings $\mathcal{F}_i(\boldsymbol{\xi})$. The PC surrogates correspond to truncated spectral expansions of second-order random variables (or vectors, or fields) in a basis of orthonormal multivariate polynomials in $\boldsymbol{\xi}$.

As in the deterministic case described in section 2.1.1, for each subdomain $\mathcal{D}_i$, we rely on samples of the boundary data at the subdomains' inner boundaries, which are now sampled jointly with the random vector $\boldsymbol{\xi}$ according to $p_{\boldsymbol{\xi}}$. For each such input sample, the corresponding local problem is solved in $\mathcal{D}_i$, and we use the

Algorithm 1: Simplified algorithm of the sampling stage (stage 4 of Fig. 2.1), including the computation of the bounds and the admissibility check of the PDE samples. Adapted from [M4].

```
foreach subdomain D_i do
    // [SERVER] Pre-processing stage
    Compute the invariant parts of the bounds for D_i ;
    // [SERVER] Sample boundary conditions
    Sample s_i^* boundary conditions for D_i ;
    s_i ← s_i^* ;
    foreach sample do
        // [SERVER]
        Add contribution of the boundary conditions to the bounds ;
        Send task to a client ;
        // [CLIENT]
        Receive task from server ;
        Solve the local PDE in D_i using the received sample of boundary
          conditions ;
        Send task (with the solution) back to server ; ▷ Corruption may occur
        // [SERVER]
        Receive returning task from client ; ▷ Task is potentially corrupted
        if received solution does not lie between the bounds then
            Discard current sample ;
            s_i ← s_i − 1 ;
        end if
    end foreach
end foreach
```

resulting input-output sample pairs to infer the PC coefficients of the local boundary-to-boundary mapping in $\mathcal{D}_i$ through robust regression. Once we have constructed the PC surrogate of the mappings on all the subdomains, imposing compatibility conditions at the inner interfaces yields a stochastic fixed-point problem. We then seek PC approximations of the boundary data at the inner interfaces such that the residual of the fixed-point problem is orthogonal to space spanned by the PC basis, in a stochastic Galerkin fashion [25, 26]. This yields a deterministic system of linear equations $\mathbf{S}\boldsymbol{u}^{\Gamma} = \boldsymbol{b}^{\Gamma}$ whose structure is similar to that of the deterministic case described in [M2].

The higher-dimensionality of the regression problems associated with the number of terms in the PC expansions, which is controlled by the size of $\boldsymbol{\xi}$ (the number of random variables in $\boldsymbol{\xi}$) on the one hand, and by the expansion truncation strategy (related to the polynomial degree of the PC surrogates), introduces additional difficulties due to the risk of overfitting, especially in the presence of faults, for reasonable sampling rates ρ. To address this problem, we resort to a lasso-like

sparsity-promoting regularization of the regression problem [27]. To solve the regression problem, we proposed an *iteratively reweighted Tikhonov* algorithm, which extends the *iteratively reweighted least squares* algorithm introduced in section 2.1.1 to regularized regression problems. Finally, we derived a robust leave-one-out cross-validation procedure to select the regularization parameter, based on the median value of the leave-one-out residuals.

The computations performed in [M7] demonstrated that, provided sufficiently many samples are generated, the method effectively overcomes the occurrence of soft faults, even when considering unrealistically large rates of soft faults.

2.1.5 Energy consumption

The solver based on the approach described in sections 2.1.1 to 2.1.3 proved to be scalable and highly resilient to both soft and hard faults, even for high rates of fault occurrence, provided that the sampling rate ρ is adapted accordingly. Although algorithmic resilience may no longer be critically needed on today's emerging exascale systems, our approach addresses the more general question of unreliable systems. In particular, in order to reduce the energy consumption of a supercomputer, it is possible to decrease its supply voltage (and operating frequency accordingly), at the expense of an increased fault rate [28].

In [M6], we explore the application of such a voltage scaling strategy to our solver. Specifically, we decrease the supply voltage and operating frequency by the same factor $r < 1$, which results in a higher fault rate. To accommodate the increased fault rate, the sampling rate ρ needs to be increased accordingly. The combined effect of the decreased frequency and increased sampling rate results in an increase of the execution time (focusing on the sampling stage, i.e., stage 4 of Fig. 2.1) by a factor of ρ/r. Using the energy consumption model and analysis proposed in [28], we derive a closed-form expression for the ratio of the energy consumption $\tilde{E}$ in the voltage-reduced configuration to the energy consumption E for running the same code in the nominal voltage mode:

$$\frac{\tilde{E}}{E} = \rho \frac{(1/r)\hat{P}_{\text{ind}} + r^2}{\hat{P}_{\text{ind}} + 1}, \tag{2.4}$$

where $\hat{P}_{\text{ind}} := P_{\text{ind}}/(CV^2 f)$ is the normalized frequency-independent active power [28], C is the switch capacitance, V is the nominal supply voltage and f is the nominal operating frequency. This allowed us to identify (large) energy-saving ranges of r for various values of $\hat{P}_{\text{ind}}$ and ρ.

2.2 Domain decomposition methods for stochastic elliptic PDEs

In this section, we are interested in the numerical solution of stochastic elliptic PDEs of the form

$$\begin{cases} \nabla \cdot [\kappa(\boldsymbol{x}, \omega) \nabla u(\boldsymbol{x}, \omega)] = -f(\boldsymbol{x}), & \boldsymbol{x} \in \mathcal{D},\ \omega \in \Omega, \\ u(\boldsymbol{x}, \omega) = 0, & \boldsymbol{x} \in \partial\mathcal{D},\ \omega \in \Omega, \end{cases} \tag{2.5}$$

where $\kappa \colon \mathcal{D} \times \Omega \to \mathbb{R}$ is a second-order random field defined on a probability space $(\Omega, \Sigma, \mathbb{P})$. In the following, we focus on the specific case where κ is a log-normal stochastic field, whose logarithm is a zero-mean, stationary Gaussian process, with a prescribed covariance kernel for which we control the variance and the correlation length.

2.2.1 Discretization of random fields

In order to solve eq. (2.5) numerically, one first typically needs to be able to approximate the random field numerically. In our case, we are interested in the discretization of $\log(\kappa)$ both in the stochastic and in the spatial dimensions. The (truncated) Karhunen–Loève (KL) expansion [29] is one of the most common techniques for the approximation of random fields. An attractive feature of the KL representation of a random field is that it is optimal in the mean-square sense. Computing the KL expansion is, however, challenging as it involves the solution of a Fredholm integral problem of the second kind. Upon discretization, this typically requires solving a generalized eigenvalue problem [26, 30], which may be computationally demanding, especially for fine discretizations (i.e., involving many degrees of freedom).

In [M8], we developed a non-overlapping domain decomposition method to efficiently approximate the KL expansion of a stochastic process. The proposed method relies on a divide-and-conquer strategy based on a partitioning of the global domain $\mathcal{D}$ into non-overlapping subdomains $\{\mathcal{D}_i\}_{i=1}^N$ such that $\bar{\mathcal{D}} = \cup_{i=1}^N \bar{\mathcal{D}}_i$ and $\mathcal{D}_i \cap \mathcal{D}_{j \neq i} = \emptyset$. Specifically, the method consists of three stages:

i) solving a local generalized eigenvalue problem over each subdomain, resulting in local KL representations of the field;
ii) using the dominant eigenfunctions from the local expansions to assemble a reduced standard eigenvalue problem; and
iii) solving the reduced eigenvalue problem to obtain the desired (global) KL expansion.

The size of the reduced eigenvalue problem is generally much smaller than the origi-

nal KL problem, but it increases with the number of subdomains, so that a trade-off between the sizes of the local subproblems and that of the reduced eigenvalue problem needs to be found in order to achieve the best performance. We also present a rigorous error analysis and propose a truncation strategy for the local and global KL expansions so as to achieve a prescribed accuracy of the global KL representation. One of the key advantages of our method is that, for a given correlation length associated with the random field, fewer random variables can be used to locally parameterize the random field over the subdomains.

The method also allows us to efficiently distribute and parallelize most of the computations. In particular, step i involves solving *independent* problems, for which only local data is needed to construct the stiffness and mass matrices of each local generalized eigenvalue problem. Step ii, on the contrary, requires global communication. However, for large problems, the computational cost of steps i and iii should dominate, which is what we observed on 2D problems of moderate size (80k degrees of freedom). Finally, for step iii, one may resort to parallel eigenvalue solvers provided by libraries such as `SLEPc` [31].

2.2.2 Accelerated Monte Carlo sampling with local polynomial chaos expansions

Various numerical methods exist for solving eq. (2.5). Perhaps the most straightforward method is Monte Carlo (MC) simulation, which consists in solving many deterministic PDEs corresponding to realizations of the field κ. Another approach consists in constructing a surrogate model of the solution u, e.g., in the form of a polynomial chaos (PC) expansion or a Gaussian process (GP). Such approaches typically suffer from the so-called curse-of-dimensionality, expressing the fact that the construction cost of the surrogate increases quickly with the stochastic input dimension. In particular, assuming a global KL approximation of κ with K random variables, the number of terms in the (total-degree truncated) PC expansion of u increases exponentially with K and the polynomial degree.

Following the earlier observation that in subdomains, a limited number of random variables may be used to accurately parameterize the random field κ at the local level, we proposed (see [M9]) a PC-based domain decomposition method for the efficient solving of stochastic elliptic PDEs of the form (2.5). We rely on the same non-overlapping decomposition as in section 2.2.1, and we denote by $\Gamma^{\text{in}} := (\cup_{i=1}^{N} \partial\mathcal{D}_i) \cap \mathcal{D} = (\cup_{i=1}^{N} \partial\mathcal{D}_i) \backslash \partial\mathcal{D}$ the inner boundaries (i.e., the common interfaces). Upon finite element spatial discretization, a Schur complement approach (see, e.g., [3, Chap. 3]) leads to a stochastic Schur complement system

of the form $\mathbf{S}(\omega)\boldsymbol{u}^\Gamma(\omega) = \boldsymbol{b}^\Gamma(\omega)$, for $\omega \in \Omega$. The so-called Schur complement matrix $\mathbf{S}$ can be interpreted as the non-overlapping counterpart of the influence (boundary-to-boundary) matrix $\mathbf{S}$ introduced in section 2.1.1, and can be decomposed into a sum of local contributions, $\mathbf{S}(\omega) = \sum_{i=1}^N \mathbf{R}_i^\mathsf{T}\mathbf{S}_i(\omega)\mathbf{R}_i$, where $\mathbf{S}_i(\omega)$ will be referred to as the local (stochastic) Schur complement matrix associated with subdomain $\mathcal{D}_i$, and where $\mathbf{R}_i$ is a deterministic restriction matrix that relates global inner boundary nodes on $\Gamma^{\text{in}} := \cup_{i=1}^N \Gamma_i^{\text{in}}$ to the local inner boundary nodes on $\Gamma_i^{\text{in}} := \partial\mathcal{D}_i \cap \mathcal{D} = \partial\mathcal{D}_i \backslash \partial\mathcal{D}$. Likewise, $\boldsymbol{b}^\Gamma(\omega) = \sum_{i=1}^N \mathbf{R}_i^\mathsf{T}\boldsymbol{b}_i^\Gamma(\omega)$, where $\boldsymbol{b}_i^\Gamma(\omega)$ will be referred to as the local right-hand side of subdomain $\mathcal{D}_i$. In the following, the explicit dependence to ω may be dropped to simplify notations when no confusion is possible.

Our approach consists in constructing local PC surrogates $\tilde{\mathbf{S}}_i \approx \mathbf{S}_i$ and $\tilde{\boldsymbol{b}}_i^\Gamma \approx \boldsymbol{b}_i^\Gamma$,

$$\tilde{\mathbf{S}}_i(\omega) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}_i} [\hat{\mathbf{S}}_i]_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}_i(\omega)), \qquad \tilde{\boldsymbol{b}}_i^\Gamma = \sum_{\boldsymbol{\alpha}\in\mathcal{A}_i} [\hat{\boldsymbol{b}}_i^\Gamma]_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}_i(\omega)), \tag{2.6}$$

of the local Schur matrices and local right-hand sides. The key observation is that our approach relies on the local nature of $\mathbf{S}_i(\omega)$ on $\mathcal{D}_i$, in the sense that it only involves the local field $\kappa|_{\mathcal{D}_i}$, which in turn is locally parameterized using a limited number m_i of random variables, collected in the random vector $\boldsymbol{\xi}_i := (\xi_i^1, \ldots, \xi_i^{m_i})$. The local PC surrogates are constructed on each subdomain (column by column in the case of $\tilde{\mathbf{S}}_i$) by solving local, stochastic influence problems by means of a local stochastic Galerkin method, in a preprocessing stage. This can be done at reasonable cost because of the reduced dimensionality (small m_i), and efficiently parallelized because the local influence problems on different subdomains are independent. Then, in a sampling stage, the full set of random variables (i.e., on all subdomains) $\{\boldsymbol{\xi}_i\}_{i=1}^N$ is sampled in a Monte Carlo fashion according to their correlation structure. Indeed, although the random variables of a subdomain (i.e., the random variables that parameterize the local field in the subdomain) are, by construction, uncorrelated, the random variables of distinct subdomains are generally correlated, with correlation structure given in [M8, M9]. The corresponding samples of the approximate local Schur complement matrices $\tilde{\mathbf{S}}_i$ are then cheaply evaluated through their functional (PC) representation. Eventually, for each sample, the global approximate Schur system is solved using a preconditioned conjugate gradient (PCG) method. The global approximate Schur matrix is either fully assembled (substructuring approach) as $\tilde{\mathbf{S}}(\omega) = \sum_{i=1}^N \mathbf{R}_i^\mathsf{T}\tilde{\mathbf{S}}_i(\omega)\mathbf{R}_i$, or applied in a matrix-free manner using the local contributions $\tilde{\mathbf{S}}_i(\omega)$ in the PCG iterations. In both cases, the (deterministic) preconditioner $\mathbf{M}$ is based on the expectation of the approximate Schur complement matrix $\tilde{\mathbf{S}}$, specifically $\mathbf{M} = \sum_{i=1}^N \mathbf{R}_i^\mathsf{T}[\hat{\mathbf{S}}_i]_{\boldsymbol{\alpha}=\mathbf{0}}\mathbf{R}_i$.

We demonstrated the superiority of the method compared to direct MC sampling of the Schur system (i.e., without resorting to the PC surrogates and re-building the local Schur complements and right-hand sides for each sample). However, as evidenced in [M9], our approach suffers from several disadvantages:

i) The approximate Schur system (and thus its solution) is contaminated by model error coming from the truncated PC surrogates.

ii) In fact, the PC-based approximation $\tilde{\mathbf{S}}$ of the Schur complement matrix does not guarantee to yield realizations that are symmetric positive definite, which is a key property of $\mathbf{S}$ and a requirement for the PCG theory to hold.

iii) For cases where the random field κ has a large variance, the proposed preconditioner based on the expected Schur complement matrix may not be efficient for realizations that deviate significantly from the expected system.

iv) Due to the global nature of the proposed preconditioner, the sampling/solving stage is not expected to scale well in parallel.

We proposed alternative approaches to address these issues in subsequent work [M10, 32], as detailed in the next section.

2.2.3 Stochastic preconditioning approaches

To address issues i and iii described above, we proposed in [M10] an alternative approach consisting in using the PC-based approximate Schur complement matrix $\tilde{\mathbf{S}}$ to generate sample-dependent preconditioners for each sample of the actual Schur complement system. The fact that the PC approximation is now used as a preconditioner (instead of a surrogate/replacement of the actual system) addresses issue i, and the fact that the preconditioner is now sample-dependent (i.e., suited to each sampled system) addresses issue iii. To address issue ii, we proposed to rely on an orthogonal factorization of the local Schur complement matrices, namely $\mathbf{S}_i(\omega) = \mathbf{H}_i(\omega)^2$, with $\mathbf{H}_i(\omega) := \mathbf{Q}_i(\omega)\mathbf{\Lambda}_i(\omega)^{1/2}\mathbf{Q}_i(\omega)^\mathsf{T}$, where $\mathbf{\Lambda}_i(\omega)$ denotes the diagonal matrix containing the (positive) eigenvalues of $\mathbf{S}_i(\omega)$, and $\mathbf{Q}_i(\omega)$ denotes the matrix of the associated eigenvectors. Then, instead of constructing PC approximations of $\mathbf{S}_i$ directly, we construct PC approximations $\tilde{\mathbf{H}}_i(\omega) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}_i}[\hat{\mathbf{H}}_i]_{\boldsymbol{\alpha}}\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}_i(\omega))$ of the local factors $\mathbf{H}_i$, and define $\tilde{\mathbf{S}}_i = \tilde{\mathbf{H}}_i^2$. This prompts the use of a non-intrusive method for the construction of the PC surrogates, where a factorization of the local Schur complement matrix $\mathbf{S}_i = \mathbf{H}_i^2$ needs to be computed for each sample point of the experimental design used for the PC construction (e.g., quadrature points). The resulting approximate Schur complement matrix, $\tilde{\mathbf{S}}(\omega) = \sum_{i=1}^{N}\mathbf{R}_i^\mathsf{T}\tilde{\mathbf{H}}_i(\omega)^2\mathbf{R}_i$, is now guaranteed to be almost-surely non-negative. The method was shown to be efficient and robust over a wide range of values for the involved parameters (number of subdomains, PC and KL truncation, smoothness of the field).

The method described above addresses most of the issues of the method presented in section 2.2.2, but still relies on a global preconditioner, which penalizes the parallel scalability of the approach, and leaves issue iv unsolved. This question was further investigated in collaboration with João Reis during his PhD degree [32]. A scalable strategy was devised based on multipreconditioning techniques [33], specifically a Neumann-Neumann (NN) PC-based preconditioner with either a Nicolaides or a GenEO coarse space. The advantage of such a preconditioner in that its application is mainly local and only requires data exchange between neighboring subdomains. The NN preconditionner has the form $\mathbf{M}_{\text{inv}}(\omega) = \sum_{i=1}^{N} \mathbf{R}_i^{\mathsf{T}} \boldsymbol{\chi}_i \mathbf{S}_i^{\dagger}(\omega) \boldsymbol{\chi}_i \mathbf{R}_i$, where $\boldsymbol{\chi}_i$ is a diagonal matrix of non-negative entries corresponding to a discrete partition of unity such that $\sum_{i=1}^{N} \mathbf{R}_i^{\mathsf{T}} \boldsymbol{\chi}_i \mathbf{R}_i = \mathbf{I}$, and $\mathbf{S}_i^{\dagger}$ denotes a Moore-Penrose pseudoinverse of $\mathbf{S}_i$. Following the orthogonal factorization strategy described previously, we construct PC approximations $\tilde{\mathbf{H}}_i^{\dagger}(\omega)$ of the local factors $\mathbf{H}_i^{\dagger}(\omega) := \mathbf{Q}_i(\omega)[\boldsymbol{\Lambda}_i^{\dagger}(\omega)]^{1/2}\mathbf{Q}_i(\omega)^{\mathsf{T}}$ such that $\mathbf{S}_i^{\dagger} = [\mathbf{H}_i^{\dagger}]^2$, where $\boldsymbol{\Lambda}_i^{\dagger}$ is a diagonal matrix with entries corresponding to the inverse of the eigenvalues of $\mathbf{S}_i$ for positive eigenvalues, 0 otherwise. The resulting PC-based approximate preconditioner is given by $\tilde{\mathbf{M}}_{\text{inv}}(\omega) = \sum_{i=1}^{N} \mathbf{R}_i^{\mathsf{T}} \boldsymbol{\chi}_i [\tilde{\mathbf{H}}_i^{\dagger}(\omega)]^2 \boldsymbol{\chi}_i \mathbf{R}_i$. Preliminary results confirm the good performance of the method, in particular in terms of scalability, which addresses issue iv.

2.3 Summary

This chapter described several aspects of my research work on domain decomposition (DD) methods, covering the topics of resilience and the application to uncertain problems.

To achieve resilience, in an overlapping DD framework, we proposed to recast the original problem as an optimization problem of learning the boundary-to-boundary mappings on the subdomains. The optimization problem is then solved in a resilient manner, using sampling and robust regression techniques. The algorithm was implemented in a server-client framework within a *selective reliability* paradigm, which allows to sandbox certain data and compute regions that need to be highly reliable. The approach was further improved by incorporating a priori bounds on the solution of the local problems on the subdomains during the sampling stage, so that faulty solutions lying outside these bounds can be discarded, following a *skeptical programming* paradigm. Finally, we investigated energy consumption considerations using a simplified energy model. By decreasing the supply voltage of a supercomputer, its power consumption is reduced quadratically at the expense of a linearly reduced speed and a higher fault rate, which can be compensated in our resilient framework by increasing the sampling rate. This allowed us to identify energy-saving ranges of

supply voltage reduction factors.

The second aspect addresses the design of DD methods for uncertain problems. The first contribution consisted of the extension of the deterministic resilient approach to stochastic problems. We proposed an approach combining the use of polynomial chaos (PC) surrogate models with robust regression, sparsity-promoting regularization, and robust cross-validation. The second contribution concerned the derivation of non-overlapping DD methods for uncertain elliptic partial differential equations. First, we designed a DD approach for the discretization of random fields, based on local Karhunen–Loève expansions of the field on the subdomains. Second, we proposed to approximate the Schur complement of the partitioned problem using PC surrogates of the local Schur matrices on the subdomains. This preliminary idea was further refined by using such local PC approximations in a scalable DD preconditioner. As such, the surrogate models are no longer considered strictly as surrogates (i.e., replacements) of the true operators, but rather as accelerators, where their model error (bias) is actually acknowledged. This last point particularly resonates with the research avenues that I have been exploring recently and which constitute one of the main prospective directions for my future research activities (see chapters 3 and 6).

Chapter 3

Multifidelity methods for uncertainty quantification

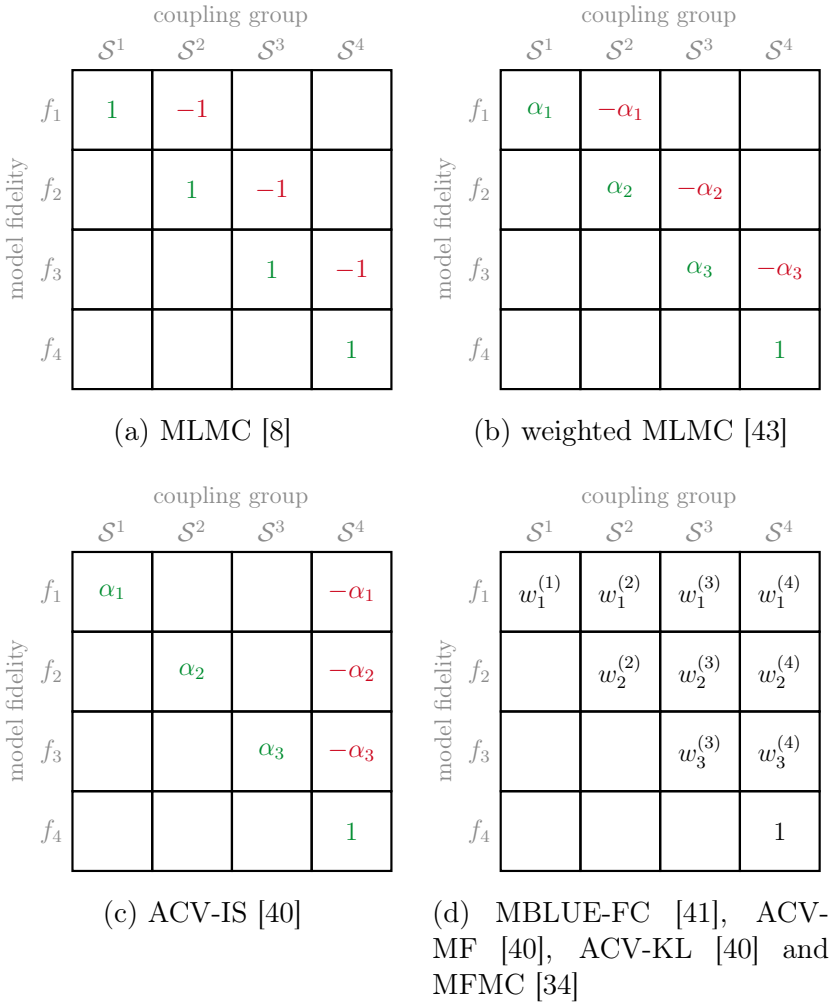
Multifidelity methods rely on a collection of numerical models of different fidelities (i.e., accuracy and associated computational cost) to reduce the overall cost of computing the solution of a problem, compared to using only the highest fidelity model. In the context of probabilistic uncertainty propagation, we focus here on multifidelity variance reduction techniques for MC-like sampling methods, which we will refer to as multifidelity MC (MFMC), although this name is used in the literature for a specific method [34, 35] that lies within this class of techniques. With that definition, MFMC includes broad classes of methods such as multilevel MC (MLMC) [7–9] multi-index MC (MIMC) [36], control variates (CV) [37, 38], control variates using estimated means (CVEM) [34, 35, 39, 40], and multilevel best linear unbiased estimators (MBLUE) [41, 42]. A unifying formulation for such methods was recently proposed in [41, 42], which may be written as

$$\hat{\theta}_K^{\mathrm{MF}} = \sum_{k=1}^{K} \sum_{\ell \in \mathcal{S}^k} w_\ell^{(k)} \hat{\theta}_\ell^{(k)}, \tag{3.1}$$

where $\hat{\theta}_K^{\mathrm{MF}}$ is the multifidelity estimator of the statistic θ we want to estimate, $\hat{\theta}_\ell^{(k)}$ is an *unbiased* estimator of θ on fidelity level ℓ based on an input sample $\zeta^{(k)}$ common to the coupling group $\mathcal{S}^k$, and $w_\ell^{(k)} \in \mathbb{R}$ is the associated weight. The members of samples $\zeta^{(k)}$ and $\zeta^{(k')}$ pertaining to two different coupling groups $\mathcal{S}^k$ and $\mathcal{S}^{k' \neq k}$ are independent, so that the terms of the outer sum are independent, while, for each of these terms, the estimators of the inner sum, thus pertaining to the same coupling group, are correlated. Finally, K represent the number of coupling groups.

Equation (3.1) can be compactly written as $\hat{\theta}_K^{\mathrm{MF}} = \langle \mathbf{W}, \hat{\mathbf{\Theta}} \rangle_{\mathrm{F}} = \mathrm{tr}(\mathbf{W}^{\mathsf{T}} \hat{\mathbf{\Theta}})$, where $\langle \cdot, \cdot \rangle_{\mathrm{F}}$ denotes the Frobenius inner product of matrices of the same size, $\mathbf{W}$ is the

(possibly sparse) matrix of coupling weights with entries $W_{ij} = w_i^{(j)}$, and $\hat{\mathbf{\Theta}}$ is the matrix of estimators with entries $\hat{\Theta}_{ij} = \hat{\theta}_i^{(j)}$. The structure of $\mathbf{W}$ alone makes it possible to distinguish different families of methods, as illustrated in fig. 3.1 (see also fig. 3.3 in section 3.3). Besides this distinction based on stochastic coupling groups, an additional segregation is usually made between general multifidelity methods and more specific multilevel methods. In the latter case, the low-fidelity models are structured in a hierarchy with a clear order in terms of cost and accuracy, while in the former case, no assumption is made about such an ordering, other than the fact that the low-fidelity models have lower cost and accuracy than the highest-fidelity model, as illustrated in fig. 3.2.



(a) MLMC [8]

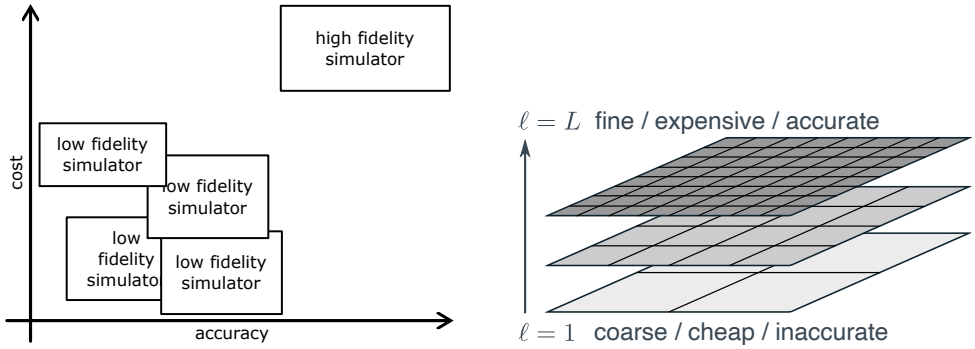
(b) weighted MLMC [43]

(c) ACV-IS [40]

(d) MBLUE-FC [41], ACV-MF [40], ACV-KL [40] and MFMC [34]

Figure 3.1: Illustration of the structure of the coupling weight matrices $\mathbf{W}$ associated with various MFMC methods, with $K = 4$. Inspired by [41, 42].

While MFMC methods were originally designed for the estimation of the expectation of a single scalar quantity, the MBLUE methodology was recently extended to the estimation of the expectation of several scalar quantities of interest [45]. The multilevel estimation of statistics of random fields is somewhat more challenging,



(a) Multifidelity methods (adapted from [44]).

(b) Multilevel methods.

Figure 3.2: Illustration of the structural difference between general multifidelity methods and more specific multilevel methods.

although discretized random fields may be seen as collections of scalar random variables so that the multi-output MBLUE methodology still applies. The difficulty lies in the fact that the linear combination in eq. (3.1) involves estimators that are themselves discretized random fields, which must be defined on the same (finest) level. Besides, it should be noted that certain UQ tasks, such as sensitivity analysis, may require the estimation of higher-order statistical moments than the expectation, which requires the original MFMC methods to be adapted. In this chapter, we present multilevel methods for statistical estimation and sensitivity analysis that address these limitations. Here the term multilevel relates to the fact that multifidelity originates from a hierarchy of levels similar to the MLMC framework (see fig. 3.2b). In section 3.1, we present an MLMC methodology for the estimation of Sobol' indices in the context of global sensitivity analysis. Then, in section 3.2, we present an extension of MLMC to discretized random fields, with an application to the estimation of normalization coefficients for diffusion-based covariance operators by (multilevel) randomization. Finally, when the exact statistics of the low-fidelity models are available, as is the case for certain surrogate models, strategies based on the original CV approach may be exploited. In section 3.3, we present a multilevel surrogate-based CV method for statistical estimation.

3.1 Multilevel global sensitivity analysis

In this and the following section, we consider an abstract numerical simulator

$$\begin{aligned} f \colon \mathcal{X} &\to \mathbb{R} \\ \boldsymbol{x} &\mapsto f(\boldsymbol{x}), \end{aligned} \tag{3.2}$$

where the p scalar input parameters, collected in $\boldsymbol{x} := (x_1, \ldots, x_p) \in \mathcal{X} \subset \mathbb{R}^p$, are uncertain. In a probabilistic UQ framework, we use random variables, defined on a probability space $(\Omega, \Sigma, \mathbb{P})$, to model the uncertainty in $\boldsymbol{x}$. Specifically, we define an $\mathcal{X}$-valued random vector $\boldsymbol{X}$ whose components $X_1 \ldots, X_p$ are *independent* random variables with prescribed probability distributions. We further assume here that $f(\boldsymbol{X})$ has a finite fourth-order moment, i.e., $\mathbb{E}[f(\boldsymbol{X})^4] < \infty$.

In [M11], we proposed an MLMC methodology and its mathematical analysis for the estimation of Sobol' indices S_u, for $\emptyset \neq u \subseteq \{1, \ldots, p\}$, based on their *pick-and-freeze* formulation [46]:

$$S_u = \frac{\mathbb{C}[f(\boldsymbol{X}_u \oplus \boldsymbol{X}_{\bar{u}}), f(\boldsymbol{X}_u \oplus \boldsymbol{X}'_{\bar{u}})]}{\mathbb{V}[f(\boldsymbol{X})]} - \sum_{\emptyset \neq v \subsetneq u} S_v, \tag{3.3}$$

where $\boldsymbol{X}_u := (X_i)_{i \in u}$, $\bar{u} := \{1, \ldots, p\} \backslash u$, $\boldsymbol{X}_u \oplus \boldsymbol{X}_{\bar{u}} = (X_1, \ldots, X_p) = \boldsymbol{X}$, and where $\boldsymbol{X}_u \oplus \boldsymbol{X}'_{\bar{u}}$ denotes the $\mathbb{R}^p$-valued random vector whose ith component is X_i if $i \in u$ and X'_i otherwise, where X_i and X'_i are i.i.d. Our multilevel methodology then lies in the MLMC estimation of the covariance term $\mathbb{C}[f(\boldsymbol{X}_u \oplus \boldsymbol{X}_{\bar{u}}), f(\boldsymbol{X}_u \oplus \boldsymbol{X}'_{\bar{u}})]$. To that end, we proposed an MLMC estimator of the covariance between two real-valued, fourth-order random variables Y and Z. The multilevel estimator is of the form $\hat{\mathrm{C}}_L^{\mathrm{ML}}[Y, Z] = \sum_{\ell=1}^{L} \hat{T}_\ell$, where $\hat{T}_1 := \hat{\mathrm{C}}_{n_1}^{(1)}[Y_1, Z_1]$ is a coarse, single-level estimator and $\hat{T}_{\ell>1} := \hat{\mathrm{C}}_{n_\ell}^{(\ell)}[Y_\ell, Z_\ell] - \hat{\mathrm{C}}_{n_\ell}^{(\ell)}[Y_{\ell-1}, Z_{\ell-1}]$ are multilevel correction estimators. Here, $\hat{\mathrm{C}}_{n_\ell}^{(\ell)}[\cdot, \cdot]$ denotes the single-level, unbiased, sample covariance estimator based on an input sample $\zeta^{(\ell)} := \{\boldsymbol{X}^{(\ell,1)}, \ldots, \boldsymbol{X}^{(\ell,n_\ell)}\}$ associated with the coupling group $\mathcal{S}^\ell$.

In addition, we provided upper bounds on the variance of the multilevel terms $\hat{T}_\ell$, for $\ell \geq 1$,

$$\mathbb{V}[\hat{T}_\ell] \leq \frac{1}{2} \frac{1}{n_\ell - 1} \left[\sqrt{\mathbb{M}^4[\delta_\ell^- Y] \, \mathbb{M}^4[\delta_\ell^+ Z]} + \sqrt{\mathbb{M}^4[\delta_\ell^- Z] \, \mathbb{M}^4[\delta_\ell^+ Y]} \right], \tag{3.4}$$

where $\mathbb{M}^4[Y] := \mathbb{E}[(Y - \mathbb{E}[Y])^4]$, and where $\delta_1^\pm Y := Y_1$ and $\delta_{\ell>1}^\pm Y := Y_\ell \pm Y_{\ell-1}$. These bounds allowed us to derive sufficient conditions, stated in [M11, theorem 2.5], to fit into the abstract MLMC framework of [47, theorem 4.1].

We also provided an adaptive MLMC algorithm aiming at reaching (heuristically) a minimal variance given a fixed computational budget. We conducted numerical experiments on an uncertain ordinary differential equation with known analytic solution and Sobol' indices. They confirmed the convergence properties of our estimator predicted by the theory.

3.2 Extension of MLMC to discretized random fields

In [M12], we proposed an extension of MLMC estimators to the case where

i) the numerical simulator $\boldsymbol{f}\colon \mathbb{R}^p \supset \mathcal{X} \to \mathcal{Y} \subset \mathbb{R}^q$ is now a vector-valued function;
ii) the input parameter $\boldsymbol{X}$ represents a discretized random field; and
iii) the output $\boldsymbol{Y} = \boldsymbol{f}(\boldsymbol{X})$ is also a discretized random field.

The main difficulty of such a setting lies in the inconsistent input and output vector dimensions across levels for a natural multilevel hierarchy of simulators $(\boldsymbol{f}_\ell)_{\ell=1}^L$ defined by $\boldsymbol{f}_\ell\colon \mathbb{R}^{p_\ell} \supset \mathcal{X}_\ell \to \mathcal{Y}_\ell \subset \mathbb{R}^{q_\ell}$, which prevent the MLMC telescoping sum from being used as is. We thus resorted to multigrid ingredients, in particular grid transfer operators, to extend the scalar MLMC estimators to the multilevel estimation of statistics of discretized random fields. Specifically, restriction operators $\mathbf{R}_L^\ell$ are used to transfer the fine input $\boldsymbol{X}$, defined on level L, to a grid on level ℓ (with $\mathbf{R}_L^L := \mathbf{I}_L$), and prolongation operators $\mathbf{P}_\ell^L$ are used to transfer the outputs defined on level ℓ to the finest grid on level L (with $\mathbf{P}_L^L := \mathbf{I}_L$). These transfer operators are defined as the composition of transfer operators between successive levels of the hierarchy,

$$\mathbf{P}_\ell^L = \mathbf{P}_{L-1}^L \cdots \mathbf{P}_{\ell+1}^{\ell+2} \mathbf{P}_\ell^{\ell+1} \text{ and } \mathbf{R}_L^\ell = \mathbf{R}_{\ell+1}^\ell \cdots \mathbf{R}_{L-1}^{L-2} \mathbf{R}_L^{L-1}, \quad \forall \ell = 1, \ldots, L-1. \tag{3.5}$$

The MLMC estimator $\hat{\boldsymbol{\theta}}_L^{\mathrm{ML}}$ of $\boldsymbol{\theta}_L := \mathbb{Q}[\boldsymbol{f}_L(\boldsymbol{X})]$ then reads $\hat{\boldsymbol{\theta}}_L^{\mathrm{ML}} = \sum_{\ell=1}^L \hat{\boldsymbol{T}}_\ell$, where $\hat{\boldsymbol{T}}_1$ and $\hat{\boldsymbol{T}}_{\ell>1}$ are unbiased estimators of $\boldsymbol{\theta}_1$ and $\boldsymbol{T}_{\ell>1} := \boldsymbol{\theta}_\ell - \boldsymbol{\theta}_{\ell-1}$, respectively, with $\boldsymbol{\theta}_\ell := \mathbb{Q}[\boldsymbol{Y}_\ell]$ and $\boldsymbol{Y}_\ell := \mathbf{P}_\ell^L \boldsymbol{f}_\ell(\mathbf{R}_L^\ell \boldsymbol{X})$. Each estimator $\hat{\boldsymbol{T}}_\ell$ is based on an input sample $\zeta^{(\ell)}$ of size n_ℓ associated with its coupling group $\mathcal{S}^\ell$, for $\ell = 1, \ldots, L$.

Following the typical Fourier analysis of multigrid methods (see, e.g., [48]), we conducted a spectral analysis of $\hat{\boldsymbol{\theta}}_L^{\mathrm{ML}}$, for the estimation of the expectation ($\mathbb{Q} = \mathbb{E}$). Similar to multigrid methods, the analysis revealed that the grid transfer operators introduce spurious high frequencies in the transferred discretized fields $\boldsymbol{Y}_\ell$. Nonetheless, owing to the telescoping correction mechanism of MLMC, these spurious high frequencies do not affect the bias of the multilevel estimators, so that $\mathbb{E}[\hat{\boldsymbol{\theta}}_L^{\mathrm{ML}}] = \boldsymbol{\theta}_L$. Instead, they do affect the estimator's variance, $\mathcal{V}(\hat{\boldsymbol{\theta}}_L^{\mathrm{ML}}) := \mathbb{E}[\|\hat{\boldsymbol{\theta}}_L^{\mathrm{ML}} - \mathbb{E}[\hat{\boldsymbol{\theta}}_L^{\mathrm{ML}}]\|_2^2]$. Using an orthogonal, Fourier-like basis $\mathbf{H}$, the estimator's variance can be decomposed into contributions associated with each basis vector $\boldsymbol{h}_k$, and, in turn, with the corresponding frequency, specifically $\mathcal{V}(\hat{\boldsymbol{\theta}}_L^{\mathrm{ML}}) = \sum_{k=1}^{q_L} \mathbb{V}[\boldsymbol{h}_k^\intercal \hat{\boldsymbol{\theta}}_L^{\mathrm{ML}}]$. This allowed us to show that the variance is deteriorated (i.e., increased) by contributions associated with high frequencies when the prolongation operators are based on low-order interpolation. Inspired by the smoothing iterations used in multigrid methods, we proposed to introduce a filtering step in the definition of the transfer operators between successive levels. More precisely, $\mathbf{P}_{\ell-1}^\ell$ and $\mathbf{R}_\ell^{\ell-1}$ are replaced with $\bar{\mathbf{P}}_{\ell-1}^\ell := \mathbf{S}_\ell \mathbf{P}_{\ell-1}^\ell$

and $\bar{\mathbf{R}}_\ell^{\ell-1} := \mathbf{R}_\ell^{\ell-1}\mathbf{S}_\ell$, respectively, where $\mathbf{S}_\ell$ denotes a low-pass filter on level ℓ, for $\ell = 2, \dots, L$. The filters are designed to damp the high-frequency components of fine signals before restricting them to a coarser grid, as well as the high-frequency components in signals prolongated from a coarser grid, where such frequencies cannot be represented.

We applied this filtered MLMC (FMLMC) approach to the multilevel estimation of normalization coefficients for diffusion-based covariance operators [49–51]. This approach relies on diffusion operators of the form $\mathbf{A}_\ell := \mathbf{V}_\ell^\mathsf{T}(\mathbf{I}_\ell - \boldsymbol{\Delta}_\ell)^J$, each related to the discretization of a heat (or, more generally, diffusion) equation on level $\ell = 1, \dots, L$, using an implicit, backward Euler (pseudo-time) integration scheme, with $J \in \mathbb{N}$ pseudo-time steps. Letting $\boldsymbol{X} \sim \mathcal{N}(\mathbf{0}_L, \mathbf{I}_L)$, we want to estimate $\boldsymbol{\theta}_L := \mathbb{V}[\mathbf{A}_L^{-1}\boldsymbol{X}] = \mathbb{E}[(\mathbf{A}_L^{-1}\boldsymbol{X}) \odot (\mathbf{A}_L^{-1}\boldsymbol{X})]$, where $\odot$ denotes the Hadamard (element-wise) product, which corresponds to the vector of the diagonal entries of $\mathbf{A}_L^{-1}\mathbf{A}_L^{-\mathsf{T}}$. Thus, the multilevel hierarchy of simulators $(\boldsymbol{f}_\ell)_{\ell=1}^L$ is defined by $\boldsymbol{f}_\ell \colon \boldsymbol{X}_\ell \mapsto (\mathbf{A}_\ell^{-1}\boldsymbol{X}_\ell) \odot (\mathbf{A}_\ell^{-1}\boldsymbol{X}_\ell)$. Our numerical experiments in 1D and 2D showed that the filtering step reduces the high-frequency contributions to the multilevel estimator's variance, resulting in smoother individual estimations of the discretized normalizing field. In addition, in instances where these high-frequency contributions are large, their filtering can even significantly improve the overall variance $\mathcal{V}(\hat{\boldsymbol{\theta}}_L^{\mathrm{ML}})$, thus exploiting the full power of MLMC. For a given computational budget, we obtained smooth estimators of the discretized normalizing fields with lower variance than the crude MC estimators and of (unfiltered) MLMC estimators.

3.3 Multilevel surrogate-based control variates

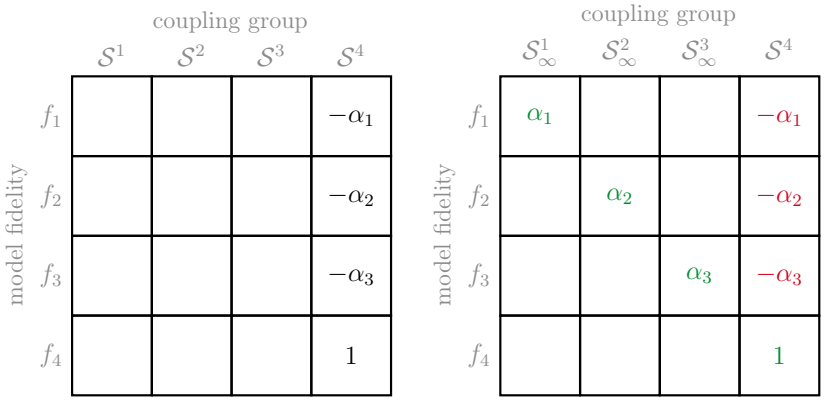
In [M13], we proposed a multilevel, surrogate-based CV estimator of some statistic θ. We start by recalling the standard (multiple) CV estimator $\hat{\theta}^{\mathrm{CV}}$ of θ, based on m control variates, which reads

$$\hat{\theta}^{\mathrm{CV}}(\boldsymbol{\alpha}) = \hat{\theta} - \hat{\boldsymbol{\nu}}^\mathsf{T}\boldsymbol{\alpha}, \quad \mathbb{E}[\hat{\boldsymbol{\nu}}] = 0, \tag{3.6}$$

where $\hat{\theta}$ is an unbiased estimator of θ, $\boldsymbol{\alpha} \in \mathbb{R}^m$ is the CV weight vector, and $\hat{\boldsymbol{\nu}}$ is a *zero-mean*, $\mathbb{R}^m$-valued random vector that is somehow correlated with $\hat{\theta}$. By construction, $\hat{\theta}^{\mathrm{CV}}$ is unbiased, and, in practice, one typically has $\hat{\boldsymbol{\nu}} = \hat{\boldsymbol{\tau}} - \boldsymbol{\tau}$, where $\boldsymbol{\tau}$ is a vector of statistics that are approximations of θ, and $\hat{\boldsymbol{\tau}}$ is an unbiased estimator of $\boldsymbol{\tau}$. The CV weight $\boldsymbol{\alpha}$ is optimized so that $\hat{\theta}^{\mathrm{CV}}(\boldsymbol{\alpha})$ has minimal variance. The optimal weight is given by $\boldsymbol{\alpha}_* = \boldsymbol{\Sigma}^{-1}\boldsymbol{c}$, where $\boldsymbol{\Sigma} := \mathbb{C}[\hat{\boldsymbol{\tau}}] \in \mathbb{R}^{m\times m}$ and $\boldsymbol{c} := \mathbb{C}[\hat{\boldsymbol{\tau}}, \hat{\theta}] \in \mathbb{R}^m$. The resulting optimal CV estimator, $\hat{\theta}^{\mathrm{CV}}(\boldsymbol{\alpha}_*)$, always has a lower variance than $\hat{\theta}$,

with a reduction factor of $1-R^2$, where $R^2 := \mathbb{V}[\hat{\theta}]^{-1}\boldsymbol{c}^\intercal\boldsymbol{\alpha}_* \in [0,1]$. In [M13], we show that introducing additional control variates always reduces (or, more precisely, never increases) the variance of the CV estimator. Note that, in practice, the statistics involved in the definition of $\boldsymbol{\alpha}_*$ are estimated using the same sample as $\hat{\theta}$ and $\hat{\boldsymbol{\tau}}$, thus deteriorating the theoretical variance reduction factor and introducing bias. The latter may be eliminated using splitting techniques [52].

In a multifidelity context, given a collection $\{f_1, \dots, f_K\}$ of models of different fidelities of the form (3.2), and f_K denoting (by convention) the high-fidelity model, we are interested in a statistical measure $\theta = \mathbb{Q}[f_K(\boldsymbol{X})]$ (e.g., $\theta = \mathbb{E}[f_K(\boldsymbol{X})]$), and we define $\boldsymbol{\tau} := (\tau_1, \dots, \tau_m)$, with $\tau_\ell = \mathbb{Q}[f_\ell(\boldsymbol{X})]$ for $\ell \leq m := K-1$. The resulting CV estimator can be written in the form of eq. (3.1), and an illustration of the resulting coupling weight matrix is given in fig. 3.3a. Note that this estimator involves a single coupling group. Assuming that $\hat{\boldsymbol{\tau}}$ is a consistent estimator of $\boldsymbol{\tau}$, we may define an asymptotic sampling version of eq. (3.6) by replacing $\boldsymbol{\tau}$ with its asymptotic version $\hat{\boldsymbol{\tau}}_\infty$, thus leading to the CV-AS estimator illustrated in fig. 3.3b.



(a) Classical CV (b) CV-AS

Figure 3.3: Illustration of the structure of the coupling weight matrices $\mathbf{W}$ associated with the CV method, with $K = m+1 = 4$. Inspired by [41, 42].

In practice, the true statistical measures $\tau_\ell = \mathbb{Q}[f_\ell(\boldsymbol{X})]$ may not be known, and thus need to be estimated. This has given rise to a family of MFMC methods exploiting this strategy (starting with the actual method referred to as "MFMC" in [35]), such as [34, 35, 39, 40] of the CVEM family, and, in a sense, also [41, 42] of the MBLUE family, which can be written in the form of eq. (3.1), as illustrated in fig. 3.1 (last two figures). In such instances, the sample allocation can be optimized considering the cost of estimating $\boldsymbol{\tau}$ using extra samples, and the deterioration of variance reduction introduced by this approximation. Alternatively, we propose a surrogate-based CV (SBCV) approach where we use surrogate models $f_{\ell \leq m} = \tilde{f}_\ell$, for which we know $\mathbb{Q}[\tilde{f}_\ell(\boldsymbol{X})]$ exactly, as low-fidelity models. PC surrogates are

ideal candidates, since their (exact) expected value and variance can be readily accessed through the PC coefficients [53–56]. Nonetheless, other candidates can be considered, in particular surrogate models based on a Taylor expansion of f around $\mathbb{E}[\boldsymbol{X}]$, as we suggested in [M13]. The advantage of such an approach is two-fold. First, the cost of evaluating the surrogate model is negligible compared to the cost of evaluating the high-fidelity model. Second, the exact (and free) access to $\boldsymbol{\tau}$ allows us not to deteriorate the variance reduction, contrary to the CVEM approaches mentioned previously. Even in the case of surrogate models for which $\boldsymbol{\tau}$ is generally not readily accessible, such as Gaussian processes, an accurate estimation of $\boldsymbol{\tau}$ can still be obtained for a negligible cost by sampling the surrogate model, corresponding to an ACV-IS [40] approach (see fig. 3.1c), whose asymptotic version is CV-AS (see fig. 3.3b). In any case, the resulting SBCV estimator is unbiased by construction, so that, in a sense, this surrogate-based approach can be seen as a bias reduction technique (w.r.t. the biased surrogate), in addition to the traditional variance reduction (w.r.t. the MC estimator) interpretation.

In [M13], we proposed a multilevel extension of the approach described above, resulting in the MLMC-MLCV estimator (here simply denoted by MLCV),

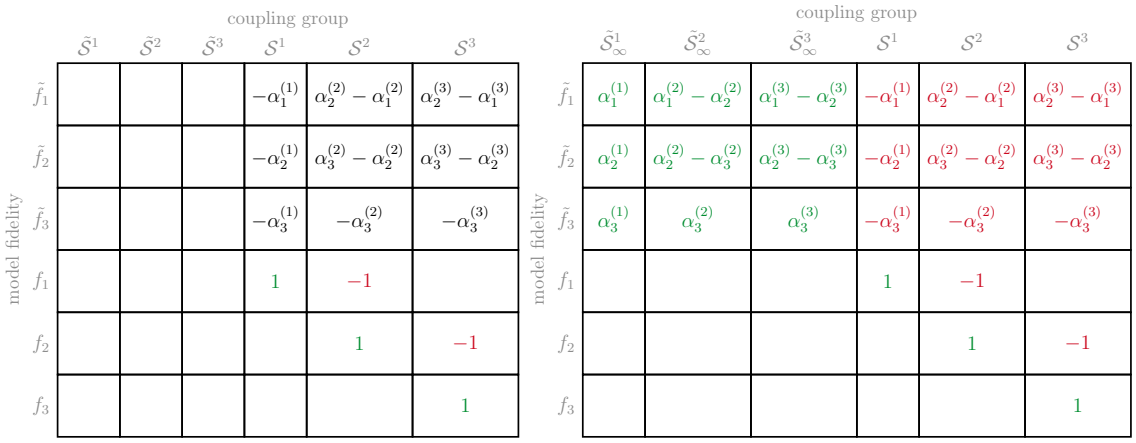
$$\hat{\theta}_L^{\mathrm{MLCV}}(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_L) = \hat{\theta}_1^{(1)} - \begin{bmatrix} \hat{\tau}_1^{(1)} - \tau_1 \\ \vdots \\ \hat{\tau}_L^{(1)} - \tau_L \end{bmatrix}^{\mathsf{T}} \boldsymbol{\alpha}^{(1)} + \sum_{\ell=2}^{L} \hat{T}_\ell^{(\ell)} - \begin{bmatrix} \hat{U}_2^{(\ell)} - U_2 \\ \vdots \\ \hat{U}_L^{(\ell)} - U_L \end{bmatrix}^{\mathsf{T}} \boldsymbol{\alpha}^{(\ell)}, \quad (3.7)$$

where

- $\hat{\theta}_1^{(1)}$ is an unbiased (coarse) estimator of $\theta_1 := \mathbb{Q}[f_1(\boldsymbol{X})]$, based on an input sample $\zeta^{(1)}$ of size n_1 associated with the coupling group $\mathcal{S}^1$;
- $\hat{\tau}_{\ell'}^{(1)}$ is an unbiased estimator of $\tau_{\ell'} := \mathbb{Q}[\tilde{f}_{\ell'}(\boldsymbol{X})] \approx \theta_{\ell'}$, based on the same input sample $\zeta^{(1)}$, for $\ell' > 1$;
- $\hat{T}_\ell^{(\ell)}$ is an unbiased estimator of the correction $T_\ell := \theta_\ell - \theta_{\ell-1}$, with $\theta_\ell := \mathbb{Q}[f_\ell(\boldsymbol{X})]$, based on an input sample $\zeta^{(\ell)}$ of size n_ℓ associated with the coupling group $\mathcal{S}^\ell$, for $\ell > 1$;
- $\hat{U}_{\ell'}^{(\ell)}$ is an unbiased estimator of $U_{\ell'} := \tau_{\ell'} - \tau_{\ell'-1} \approx T_{\ell'}$, based on the same input sample $\zeta^{(\ell)}$, for $\ell, \ell' > 1$.

This multilevel estimator proposes to reduce the variance of the MLMC estimator by using, for each term of the telescoping sum, multiple CVs based on surrogate models at *all* levels. Equation (3.7) can be written in the form of eq. (3.1), and an illustration of the coupling weight matrix, along with its asymptotic sampling counterpart, is given in fig. 3.4 in the specific case where $\hat{U}_{\ell'}^{(\ell)}$ is defined as $\hat{U}_{\ell'}^{(\ell)} := \hat{\tau}_{\ell'}^{(\ell)} - \hat{\tau}_{\ell'-1}^{(\ell)}$. The optimal CV weights are given by $\boldsymbol{\alpha}_*^{(\ell)} = \mathbb{C}[\hat{\boldsymbol{U}}^{(\ell)}]^{-1}\mathbb{C}[\hat{T}_\ell^{(\ell)}, \hat{\boldsymbol{U}}^{(\ell)}]$, where $\hat{T}_1^{(1)} := \hat{\theta}_1^{(1)}$,

$\hat{\boldsymbol{U}}^{(1)} := (\hat{\tau}_{\ell'}^{(1)})_{\ell'=1}^{L}$ and $\hat{\boldsymbol{U}}^{(\ell)} := (\hat{U}_{\ell'}^{(\ell)})_{\ell'=2}^{L}$ for $\ell > 1$. The resulting optimal estimator has variance $\sum_{\ell=1}^{L}(1 - R_\ell^2)\mathbb{V}[\hat{T}_\ell^{(\ell)}]$, with $R_\ell^2 := \mathbb{V}[\hat{T}_\ell^{(\ell)}]^{-1}\mathbb{C}[\hat{T}_\ell^{(\ell)}, \hat{\boldsymbol{U}}^{(\ell)}]^\intercal \boldsymbol{\alpha}_*^{(\ell)} \in [0, 1]$, which is smaller than the variance of the MLMC estimator given by $\sum_{\ell=1}^{L}\mathbb{V}[\hat{T}_\ell^{(\ell)}]$.



(a) Classical MLCV (b) MLCV-AS

Figure 3.4: Illustration of the structure of the coupling weight matrices $\mathbf{W}$ associated with the MLCV method, with $K = L = m = 3$ and $\hat{U}_{\ell'}^{(\ell)} := \hat{\tau}_{\ell'}^{(\ell)} - \hat{\tau}_{\ell'-1}^{(\ell)}$. Inspired by [41, 42].

We derived PC-based MLCV estimators for the estimation of the expected value ($\mathbb{Q} = \mathbb{E}$) and of the variance ($\mathbb{Q} = \mathbb{V}$) [M13]. The practical definition of $\hat{\boldsymbol{U}}^{(\ell)}$ was discussed, as well as several variants of the above strategy taking into account the construction cost of the surrogate models. The proposed estimators were successfully tested on the estimation of the expected value of the output of a discretized, uncertain heat equation. Our numerical experiments confirmed the significant variance reduction brought by the multilevel CVs, consistent with the theory.

3.4 Summary

In this chapter, we presented multilevel and/or multifidelity methods for the estimation of statistics of the output of expensive numerical simulators. The main contributions concern extensions of the well-established MLMC method to the estimation of covariance for sensitivity analysis (section 3.1) and to the estimation of statistics of discretized random fields (section 3.2), as well as the combination of MLMC with surrogate-based control variate estimation (section 3.3). The efficiency of these methods in terms of variance reduction was demonstrated theoretically and/or numerically. The proposed multilevel estimators all belong to the class of estimators defined by the unifying form eq. (3.1), so that the MBLUE approach of [41, 42] may be employed instead of MLMC, which should result in a greater variance reduction.

Chapter 4

Multigrid methods for the hybrid high-order discretization

This chapter is dedicated to the development of multigrid methods for the hybrid high-order (HHO) discretization [57–59] that were made in the context of Pierre Matalon's PhD degree [60]. The HHO discretization is hybrid in the sense that the discrete unknowns are located on the mesh and its skeleton. Specifically, the PDE is discretized (in variational form) using broken polynomial spaces defined on the mesh elements (or cells) and faces, which allows high-order representations of the solution. The main advantages of the HHO discretization lie in its native support of polyhedral elements, its optimal orders of convergence, and its reduced number of degrees of freedom (DoFs) through the static condensation process. Static condensation relies on the fact that the interior unknowns (i.e., cell unknowns) are not (directly) coupled with each other, instead they only interact through interface unknowns (i.e., face unknowns). As a consequence, in the resulting system of linear equations representing the discretized PDE, cell unknowns can be eliminated by a Schur complement technique, yielding a reduced system involving only face unknowns. This reduced system is referred to as the *statically condensed* system, *trace* system or *Schur complement* system. The specific structure of the trace system calls for dedicated solvers. In particular, for multigrid methods [48, 61–63], specific intergrid transfer operators need to be designed and appropriate smoothers must be employed.

In section 4.1 we present an h-multigrid method relying on a tailored prolongation operator. Then, section 4.2 introduces and compares different h-, p- and hp- coarsening strategies. Finally, in section 4.3 we propose an algebraic multigrid (AMG) method for the lowest polynomial order. The approaches are tested in 2D

and 3D on an elliptic diffusion problem of the form

$$\begin{cases} -\nabla \cdot (\mathbf{K}\nabla u) = f & \text{in } \mathcal{D} \subset \mathbb{R}^d,\ d \in \{2,3\}, \\ u = 0 & \text{on } \partial\mathcal{D}, \end{cases} \tag{4.1}$$

where $\mathbf{K}\colon \mathcal{D} \to \mathbb{R}^{d\times d}$ is a symmetric diffusion tensor field.

4.1 An h-multigrid method

4.1.1 Nested meshes

In [M15], we proposed an h-multigrid method for solving the trace system arising from the HHO discretization of model problem (4.1) on a hierarchy of nested meshes. A crucial assumption is that the hierarchy successively coarsens not only elements, but also faces, so that the smoother (which applies to the face unknowns) may perceive (and dampen) lower frequencies of the error on the faces. As mentioned previously, one of the key ingredient of the proposed multigrid method for the trace system lies in the definition of suitable intergrid transfer operators. In particular, because the system is defined on the mesh skeleton, it is not straightforward to define a face-based prolongation operator P that interpolates a face-defined correction from a coarse mesh skeleton to a finer one. The proposed prolongation operator, which we will refer to as the prolongation operator *by decondensation*, consists of the composition of two operations, which involves as an intermediary step the reconstruction of the correction over the elements. Specifically, we define the prolongation as $P := \Pi^{\ell}_{\ell-1} \circ \Theta_{\ell-1}$, where $\Theta_{\ell-1}$ will be referred to as the coarse potential reconstruction operator, and $\Pi^{\ell}_{\ell-1}$ will be referred to as the trace prolongation operator. The coarse potential reconstruction operator $\Theta_{\ell-1}$ is itself composed of two steps. First, it computes the interior correction defined on the coarse cells from the correction on the coarse faces (similar to how one would compute the solution of a PDE over subdomains from the boundary data in a domain decomposition framework). Then, using the hybrid information, i.e., both the cell- and face-defined correction (of degree k), it reconstructs a higher-degree approximation (of degree $k+1$) over the coarse cells, using the HHO potential reconstruction operator, which is a central component of the HHO methodology. Finally, the trace prolongation operator $\Pi^{\ell}_{\ell-1}$ is defined as an orthogonal projection that maps the corrections of degree $k+1$ defined on the coarse cells to a correction of degree k defined on the fine faces. The full prolongation process is illustrated in fig. 4.1. The restriction operator is defined as the adjoint of the prolongation operator, and block versions of fixed-point iterations are used as smoothers, whose block size correspond to the number of degrees

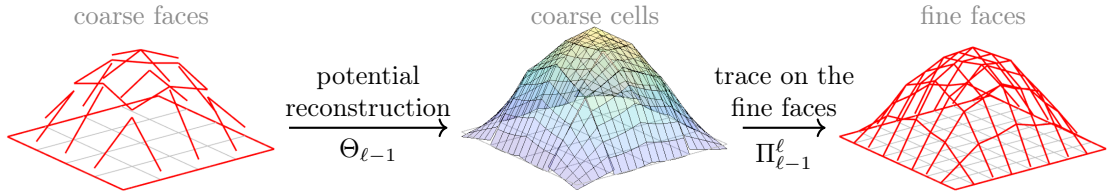


Figure 4.1: Prolongation from coarse to fine faces (edges in 2D). Adapted from [M15].

of freedom per face (i.e., the dimension of the polynomial basis on the faces).

The numerical experiments carried out in [M15] demonstrated the optimality of our multigrid solver, in the sense that the number of iterations required to reach convergence is independent of the mesh size, and its robustness with respect to the polynomial degree and to heterogeneity in the diffusion tensor.

4.1.2 Non-nested meshes

In [M16], we proposed an extension of our multigrid solver to non-nested mesh hierarchies. Again, a proper definition of the prolongation operator is at the heart of the proposed method. Specifically, we modify the definition of the prolongation to $P := \Pi_\ell \circ J^\ell_{\ell-1} \circ \Theta_{\ell-1}$, which includes an additional orthogonal projection step $J^\ell_{\ell-1}$ from the coarse cells to the (non-nested) fine cells. Note that this addition requires a slight modification of the trace operator, now denoted by Π_ℓ, which now maps the corrections of degree $k+1$ defined on the *fine* cells to a correction of degree k defined on the fine faces. The numerical evaluation of $J^\ell_{\ell-1}$ requires the computing of the geometric intersections between coarse and fine elements, which may be computationally demanding. To avoid this costly calculation, we propose an approximation of $J^\ell_{\ell-1}$ that does not require the computation of the geometric intersections. This approximate operator instead relies on the definition of an approximate intersection between a coarse cell T_c and a fine cell T_f, denoted by $T_\mathrm{c} \tilde{\cap} T_\mathrm{f}$, based on a subdivison of T_f. Each sub-element t_f of T_f belongs to the approximate intersection $T_\mathrm{c} \tilde{\cap} T_\mathrm{f}$ if and only if T_c is the closest coarse element to t_f. Otherwise, it will belong to the approximate intersection of T_f with another coarse element T'_c (which is the closest to it). Our numerical experiments showed that, in 2D, the solver based on approximate intersections behaved nearly identically to the solver based on exact intersections. In 3D, computing exact intersections was not affordable, but the solver based on approximate intersections exhibited a good convergence rate and scalability for moderate polynomial degrees, including for complex geometries requiring unstructured meshes.

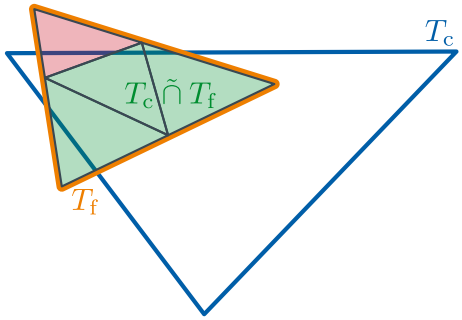


Figure 4.2: Illustration of the approximate intersection $T_\mathrm{c} \tilde{\cap} T_\mathrm{f}$ between a coarse cell T_c (in blue) and a fine cell T_f (in orange). The green area, consisting of a collection of sub-elements of T_f, represents $T_\mathrm{c} \tilde{\cap} T_\mathrm{f}$.

4.2 Multigrid coarsening strategies

In [M17], we introduced coarsening strategies for the definition of the multigrid hierarchy. In particular, we introduce p-coarsening, which consists in decreasing the polynomial degree of the HHO approximation. Then, the multigrid hierarchy can be designed using p-coarsening, h-coarsening (i.e., mesh coarsening), or a combination of both, referred to as hp-coarsening. When considering p-coarsening, the prolongation operator can either be defined as previously (prolongation by decondensation), and the restriction as its adjoint. Alternatively, the prolongation operator may be defined by natural injection and the restriction by orthogonal projection. In that case, efficient implementation can be achieved by constructing hierarchical, orthogonal polynomial bases on the faces. Different multigrid strategies, namely based on h-only, p-then-h, and hp-then-h multigrid hierarchies, are compared on a variety of 2D and 3D test problems. All the strategies exhibit fast convergence with a low, mesh-independent number of iterations to reach convergence. The strategies are then compared in terms of CPU time and equivalent work units for the setup stage and the iteration stage. The results show that the behavior of the different solvers is problem dependent. Nonetheless, we provide insights on how to choose a suitable strategy, depending on the problem, and depending on whether the problem needs to be solved only once or multiple times, e.g., for multiple right-hand sides.

4.3 An algebraic multigrid method for the lowest order

Geometric multigrid methods require geometric information about the PDE discretization, namely a hierarchy of meshes at different resolutions. This can be an obstacle to the use of such methods in an industrial context, despite their efficiency.

Algebraic multigrid (AMG) methods, on the other hand, do not require other information than the linear system itself. In [M18], we proposed an AMG methodology for the lowest order ($k = 0$) HHO discretization. The proposed AMG solver heavily relies on ingredients of the AGMG solver [64]. In particular, we rely on a pairwise aggregation strategy based on a strong negative coupling criterion, and we use the so-called K-cycle, which introduces Krylov acceleration into the multigrid recursive cycle. The resulting multigrid cycle is itself used as a variable preconditioner of an outer flexible Krylov method. The particularity of our method is that it uses the uncondensed HHO system (i.e., before static condensation), so that the connectivity information between elements and faces can be retrieved. From this information, an algebraic, element-based aggregation method can be devised to mimic the behavior of a geometric coarsening or semi-coarsening strategy. A crucial step of face collapsing is included so that the faces are coarsened. Inspired by plain aggregation techniques, we define (highly) sparse restriction and prolongation matrices that are then used to construct coarse uncondensed systems from fine ones in a Galerkin manner. These algebraic intergrid transfer matrices are only used for the construction of the hierarchy of uncondensed systems, but not inside the multigrid cycle. Instead, we derive algebraic versions of the intergrid transfer operators designed previously for the geometric solver described in section 4.1, in particular the prolongation operator by decondensation, which make use of the hierarchy of uncondensed systems. The multigrid iterations operate on the condensed systems, whose hierarchy is also constructed in a Galerkin manner, but with these newly defined operators. Our numerical experiments demonstrated the good performance of the proposed AMG solver, comparable to AGMG [64] on a variety of test problems, but with a better robustness to orthotropic anisotropy.

4.4 Summary

In this chapter, we presented contributions to the design of multigrid solvers for hybrid high-order (HHO) discretizations. The core contribution concerned the definition of appropriate intergrid transfer operators based on decondensation, and block smoothers, which lead to the design of a robust and scalable geometric multigrid method. Originally designed for nested meshes, our solver was then adapted to non-nested mesh hierarchies, by incorporating an additional projection operator. In practice, for computational reasons, this projection operator is approximated through the definition of approximate intersections between (non-nested) coarse and fine elements, thus alleviating the heavy calculation of all exact geometric intersections. We also investigated various coarsening strategies, mixing h-coarsening (i.e.,

increasing the size of the mesh elements) and p-coarsening (i.e., deteriorating the polynomial degree of approximation). Finally, we proposed an algebraic multigrid preconditioner for lowest-order HHO discretizations, relying on the uncondensed system to extract coupling information about the degrees of freedom. One natural continuation of this work would concern HHO discretizations of other partial differential equations (PDEs), such as the Stokes or the Navier-Stokes equations, although the design of efficient multigrid algorithms is already a challenge for more traditional discretizations of such PDEs.

Chapter 5

Sequences of linear systems

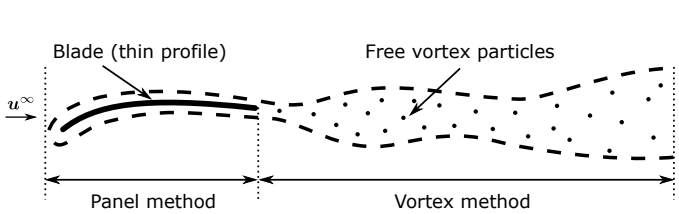
This chapter is dedicated to the design of iterative solver strategies for solving sequences of linear systems with multiple left- and right-hand sides. Such sequences may arise, for example, from unsteady computational fluid dynamics (CFD) simulations (typically with implicit time-stepping schemes), in non-linear optimization problems, or when sampling discretized stochastic PDEs. In section 5.1, we present an iterative solver approach for solving sequences of dense linear systems arising from unsteady CFD simulations of marine current turbine farms, based on a Krylov iterative solver with an appropriate preconditioner. In section 5.2, we discuss spectral recycling strategies for solving sequences of correlated systems corresponding to samples of discretized stochastic PDEs.

5.1 Preconditioning BEM systems for marine current turbine farms

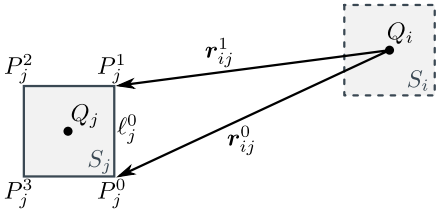
During my PhD [M20], we developed an iterative solver approach for the modeling of marine current turbine farms, based on a vortex particle method [M21]. The set-up consists of an exterior fluid domain with moving boundaries S, assumed non-deformable, representing the surfaces of the blades and hubs of the $N \geq 1$ turbines. The flow is governed by the incompressible Navier-Stokes equations, which may be written in their velocity-vorticity formulation as

$$\begin{cases} \dfrac{\mathrm{D}\boldsymbol{\omega}}{\mathrm{D}t} = (\boldsymbol{\omega} \cdot \nabla \boldsymbol{u}) + \boldsymbol{u}\Delta\boldsymbol{\omega}, \quad \boldsymbol{\omega} = \nabla \times \boldsymbol{u}, \\ \nabla \cdot \boldsymbol{u} = 0, \end{cases} \tag{5.1}$$

where $\times$ denotes the cross product and $\nabla\times$ denotes the curl operator. The velocity field $\boldsymbol{u}$ is decomposed as $\boldsymbol{u} = \boldsymbol{u}^{\infty} + \boldsymbol{u}^{\phi} + \boldsymbol{u}^{\psi}$ according to the Helmholtz decom-



(a) Cross-sectional illustration of the discretization method, showing the location of the panel method and of the particle method.



(b) Notations for the interaction between two panels S_i and S_j.

Figure 5.1: Illustration of the particle and panel methods, adapted from [M21].

position, where $\boldsymbol{u}^\infty$ is the inflow, $\boldsymbol{u}^\phi$ is the potential velocity field and $\boldsymbol{u}^\psi$ is the rotational part of the velocity field. The potential velocity may be expressed as

$$\boldsymbol{u}^\phi(\boldsymbol{x}) = \frac{1}{4\pi}\nabla \iint_S \mu(\boldsymbol{y}) \frac{(\boldsymbol{y}-\boldsymbol{x})\cdot \boldsymbol{n}(\boldsymbol{y})}{|\boldsymbol{y}-\boldsymbol{x}|^3}\, \mathrm{d}S(\boldsymbol{y}), \tag{5.2}$$

where μ represents a distribution of normal dipoles on the surface S of the turbine blades. The discretization method consists of two parts: first, the moving surfaces S are discretised using an integral panel method, otherwise known as the boundary element method (BEM); and second, the fluid domain is discretized into free vortex particles. A sketch of the method is depicted in fig. 5.1a, for a single blade represented as a thin profile.

The boundaries S are discretized using quadrangular panels $\{S_i\}_{i=1}^n$, over each of which the dipole distribution is assumed to be constant. The dipole distribution is determined so as to enforce boundary conditions for the flow velocity $\boldsymbol{u}$ on the panels. This involves solving, at each time step t, a linear system of the form $\mathbf{A}(t)\boldsymbol{\mu}(t) = \boldsymbol{b}(t)$, where the elements μ_i of $\boldsymbol{\mu} \in \mathbb{R}^n$ correspond to the piecewise constant values of the dipole distribution on the panels. The entries A_{ij} of $\mathbf{A}$ are defined as

$$A_{ij} = \frac{\boldsymbol{n}_i}{4\pi} \cdot \sum_{k=0}^{3} \int_{\ell_j^k} \frac{\boldsymbol{y}-\boldsymbol{x}_i}{|\boldsymbol{y}-\boldsymbol{x}_i|^3} \times \mathrm{d}\boldsymbol{\ell}(\boldsymbol{y}) \tag{5.3}$$

$$= \frac{\boldsymbol{n}_i}{4\pi} \cdot \sum_{k=0}^{3} (|\boldsymbol{r}_{ij}^k| + |\boldsymbol{r}_{ij}^{k+1}|) \left(1 - \frac{\boldsymbol{r}_{ij}^k \cdot \boldsymbol{r}_{ij}^{k+1}}{|\boldsymbol{r}_{ij}^k|\,|\boldsymbol{r}_{ij}^{k+1}|}\right) \frac{\boldsymbol{r}_{ij}^k \times \boldsymbol{r}_{ij}^{k+1}}{|\boldsymbol{r}_{ij}^k \times \boldsymbol{r}_{ij}^{k+1}|^2}, \tag{5.4}$$

where the notations are depicted in fig. 5.1b, $k+1$ is taken modulo 4, and where $\boldsymbol{x}_i$ represents the location of Q_i, i.e., the center of panel S_i. We see immediately that $\mathbf{A}$ is dense, and that it is generally non-symmetric. In fact, our experiments indicate that it is generally not positive definite either.

When a single ($N = 1$) turbine is considered, the matrix $\mathbf{A}$ represents the influ-

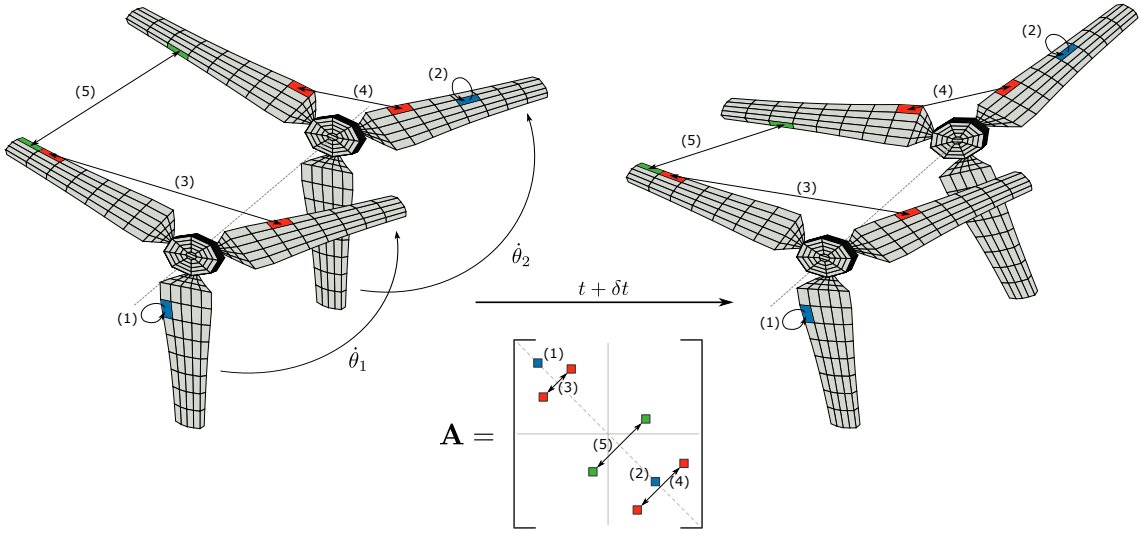


Figure 5.2: Illustration of two rotating turbines in interaction, at two successive time steps t and $t + \delta t$, and of the consequence on the interaction between panels. Reproduced from [M21].

ence of the turbine on itself. Because the turbine is assumed to be non-deformable, the rotation does not change the coefficients of $\mathbf{A}$ [M21]. In such a case, because the system is small enough ($n \approx 2000$ at most), we may rely on an LU factorization of $\mathbf{A}$ at the beginning at the simulation, and use it to subsequently solve the linear system $\mathbf{A}\boldsymbol{\mu}(t) = \boldsymbol{b}(t)$ at each time step. However, this property no longer holds when considering multiple ($N > 1$) turbines rotating at different velocities, as illustrated in fig. 5.2. The matrix $\mathbf{A}$ then has a natural block structure, as shown in eq. (5.5), where each block represents the influence of on turbine on another (or the same, for diagonal blocks) one. The diagonal blocks, which represent the influence of a turbine on itself, are stationary (i.e., their coefficients do not change with time). The entries of the off-diagonal blocks, corresponding to the interaction between panels belonging to different turbines, change with time but have a small magnitude, due to the quadratic decay of the interactions with the distance, see eq. (5.3). We thus resort to an iterative Krylov solver, namely Bi-CGSTAB [65], to solve the linear systems at each time step, using an appropriate preconditioner. The preconditioner $\mathbf{M}$ is obtained by replacing the non-stationary blocks, i.e., the off-diagonal blocks, of the influence matrix $\mathbf{A}$ by zero blocks,

$$\mathbf{A}(t) = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12}(t) & \cdots & \mathbf{A}_{1N}(t) \\ \mathbf{A}_{21}(t) & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2N}(t) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{N1}(t) & \mathbf{A}_{N2}(t) & \cdots & \mathbf{A}_{NN} \end{bmatrix}, \quad \mathbf{M} := \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{NN} \end{bmatrix}, \quad (5.5)$$

which corresponds to a block-Jacobi preconditioner. The individual LU factoriza-

tions of the diagonal blocks $\mathbf{A}_{ii}$ may be computed at the beginning of the simulation, and subsequently used to precondition the Bi-CGSTAB iterations. This strategy was tested on a 10-turbine farm, discretized with a total of $n \approx 20{,}000$ panels. Our experiments indicate that the iterative approach brings a substantial time save as compared to a naive use of a direct method at each time step, or of an unpreconditioned Bi-CGSTAB iterative method. In particular, this allows the cost of emitting particles at the trailing edge of the blades (transition between the panel method and the vortex method in fig. 5.1a), which includes the linear system solve, to remain negligible compared to the cost of advecting the particles in the fluid domain.

5.2 Spectral recycling techniques

In [M19, 66], we investigated the use of Krylov subspace spectral recycling approaches based on deflation techniques for solving sequences of linear systems with multiple left- and right-hand sides, $\{\mathbf{A}_t\boldsymbol{x}_t = \boldsymbol{b}_t\}_{t\geq 1}$. In that context, deflation techniques are used to recycle approximate spectral information about the left-hand side $\mathbf{A}_t$ from one system to the next. Such techniques have proved to be very effective for sequences with multiple right-hand sides, but with constant left-hand side [67], i.e., sequences of the form $\{\mathbf{A}\boldsymbol{x}_t = \boldsymbol{b}_t\}_{t\geq 1}$. The reason is that the spectral information about the constant matrix $\mathbf{A}$ is continually refined from one system to the next. When the matrix $\mathbf{A}_t$ changes, however, it becomes less clear whether recycling spectral information from $\mathbf{A}_t$ may be beneficial for the subsequent system $\mathbf{A}_{t+1}\boldsymbol{x}_{t+1} = \boldsymbol{b}_{t+1}$. As a natural first step, we investigate if this approach may still be effective for sequences where $\mathbf{A}_t$ does not "change much" from one system to the next. Such a situation occurs, for instance, when sampling the random field κ of eq. (2.5) according to a Markov chain, in the context of Bayesian inference. As a consequence, the sample matrices $\{\mathbf{A}_t\}_{t\geq 1}$, corresponding to the discretization of the PDE with the sample random fields $\{\kappa_t\}_{t\geq 1}$, are correlated. We set up an illustrative test case where κ is a log-normal random field, whose logarithm is a Gaussian field approximated by a truncated KL expansion. The resulting stochastic coordinates $\boldsymbol{\xi}$ of the parameterized field $\log \kappa(\boldsymbol{\xi})$ are then sampled by Markov chain Monte Carlo (MCMC) using a Gaussian proposal distribution with variance ϑ^2, which controls the correlation between subsequent samples. Our numerical experiments indicate that, in the absence of a (first-level) preconditioner, deflation accelerates the convergence of the conjugate gradient (CG) solver from a certain point in the sample chain as compared to the plain CG without deflation. We then tested deflated CG in combination with a first-level block-Jacobi preconditioner (Def-PCG-bJ), and we compared it to a preconditioned CG without deflation with an AMG preconditioner

(PCG-AMG), which is the state-of-the-art preconditioner for such SPD systems arising from discretized elliptic PDEs. Although block-Jacobi preconditioning is cheaper and (on its own) less efficient for this class of problems, our experiments indicate that its combination with deflation, resulting in the Def-PCG-bJ solver, makes it possible to reach similar convergence rates as PCG-AMG. This constitutes encouraging results for discretizations for which an AMG preconditioner is not readily accessible (e.g., HHO discretizations, see section 4.3), or for other types of PDEs (e.g., the Helmholtz equation) for which there exist no state-of-the-art preconditioner, at least not as versatile and efficient as AMG.

5.3 Summary

In this chapter, I presented two contributions for speeding-up the solving of sequences of linear systems in specific contexts. The first is based on a block Jacobi preconditioner for systems arising from a BEM discretization of interacting turbines, taking advantage of the isometry-invariance of the diagonal blocks corresponding to individual, rigid turbines. The second contribution is an extension of deflation techniques, originally designed for systems with multiple right-hand sides, to sequences of systems with multiple left-hand sides, which proved to be efficient when successive operators are correlated. A possible avenue for relaxing the constraints of similarity between successive systems would be to resort to randomized algorithms [68–70] to approximate eigen-pairs of the *current* operator, as suggested in [71] using limited memory preconditioning [72], which is a preconditioning technique similar to deflation.

Chapter 6

Outlook and future work

My future research activities will primarily focus on multilevel and multifidelity methods for UQ and optimization under uncertainty (OUU) as well as advanced preconditioning techniques for (sequences of) systems of linear equations. In particular, one of the ideas I would like to explore further in the coming years is the use of surrogate models based on neural networks (NNs) as accelerators in hybrid machine learning frameworks for the propagation of uncertainties or for the solution of systems of linear equations. Such surrogate models fit well into the framework of control variates for multifidelity estimation (see chapter 3) or the framework of NN preconditioners for flexible subspace iteration solvers (see section 6.3 below). In this chapter, I detail three medium-term research perspectives. In section 6.1, I present ideas for combining (F)MLMC with multigrid techniques for the normalization problem introduced in section 3.2. Then, in section 6.2, I discuss avenues and obstacles for the design of multifidelity ensemble-variational data assimilation algorithms for large-scale applications in geosciences, e.g, for numerical weather prediction systems. Finally, in section 6.3, I introduce a prospective hybrid machine learning and numerical linear algebra solver for systems of linear equations arising from the discretization of PDEs, based on a non-linear, NN-based preconditioner for flexible subspace iteration solvers.

6.1 Combining MLMC and multigrid techniques

A natural extension of the work presented in section 3.2 would be the combination of FMLMC with a multigrid method for solving the discretized diffusion problems. In particular, the full multigrid (FMG) cycle [63, section 2.6] seems appropriate, as suggested by [73, 74] in a different context. From section 3.2, we observe that evaluations of $\boldsymbol{f}_\ell$ require the solution of a system of linear equations, namely $\boldsymbol{f}_\ell(\boldsymbol{X}_\ell) = \boldsymbol{g}_\ell(\mathbf{A}_\ell^{-1}\boldsymbol{X}_\ell) = (\mathbf{A}_\ell^{-1}\boldsymbol{X}_\ell) \odot (\mathbf{A}_\ell^{-1}\boldsymbol{X}_\ell)$. The FMLMC estimator of $\boldsymbol{\theta}_L =$

$\mathbb{E}[(\mathbf{A}_L^{-1}\boldsymbol{X}) \odot (\mathbf{A}_L^{-1}\boldsymbol{X})]$ involves sampling the quantities $\boldsymbol{Y}_\ell = \bar{\mathbf{P}}_\ell^L \boldsymbol{g}_\ell(\mathbf{A}_\ell^{-1}(\bar{\mathbf{R}}_L^\ell \boldsymbol{X}))$, with filtered grid transfer operators defined as in eq. (3.5) but with the filtered versions of the inter-level operators, $\bar{\mathbf{P}}_{\ell-1}^\ell$ and $\bar{\mathbf{R}}_\ell^{\ell-1}$. The (approximate) solving of the systems of linear equations could be obtained by FMG, and the samples of $\boldsymbol{Y}_\ell$ could be collected after each V-cycle on level ℓ, as illustrated in fig. 6.1, inspired by [73]. Specifically, on level $\ell > 1$, an approximation $\boldsymbol{Z}_\ell$ of $\mathbf{A}_\ell^{-1}(\bar{\mathbf{R}}_L^\ell \boldsymbol{X})$ is obtained by $\texttt{V-cycle}_\ell(\boldsymbol{Z}_\ell^0, \bar{\mathbf{R}}_L^\ell \boldsymbol{X})$ from an initial guess $\boldsymbol{Z}_\ell^0$ corresponding to the interpolation of the approximate solution $\boldsymbol{Z}_{\ell-1}$ at the previous (coarser) level onto level ℓ using the FMG interpolation operator $\boldsymbol{\Pi}_{\ell-1}^\ell$ (see, e.g., [63, section 2.6]), i.e., $\boldsymbol{Z}_\ell^0 = \boldsymbol{\Pi}_{\ell-1}^\ell \boldsymbol{Z}_{\ell-1}$. On the coarsest level $\ell = 1$, $\boldsymbol{Z}_1$ may be obtained using a direct solver. The approximate solutions $\boldsymbol{Z}_\ell$ are then squared element-wise through $\boldsymbol{g}_\ell$ and prolongated to the finest grid through $\bar{\mathbf{P}}_\ell^L$ to obtain $\boldsymbol{Y}_\ell$. A multilevel sampling strategy similar to that proposed in [73] may be employed. Sampling via FMG up to level L not only provides an n_L-sample of $\boldsymbol{Y}_L$, but also of $\boldsymbol{Y}_{L-1}, \dots, \boldsymbol{Y}_1$ on the coarser levels. To obtain an n_{L-1}-sample of $\boldsymbol{Y}_{L-1}$, one may simply complement the previously obtained n_L-sample with an $(n_{L-1} - n_L)$-sample via FMG up to level $L-1$, and so on.

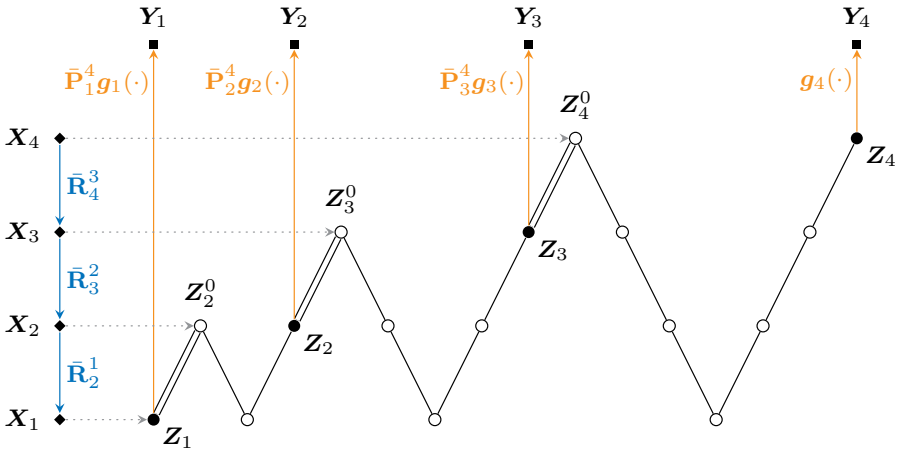


Figure 6.1: Sketch of a prospective FMG-MLMC algorithm for the estimation of $\boldsymbol{\theta}_L = \mathbb{E}[(\mathbf{A}_L^{-1}\boldsymbol{X}) \odot (\mathbf{A}_L^{-1}\boldsymbol{X})]$ for the problem described in section 3.2, here with $L = 4$ and $\boldsymbol{X}_4 := \boldsymbol{X}$. The double line ($=$) represents the FMG interpolation $\boldsymbol{\Pi}_\ell^{\ell+1}$ and $\boldsymbol{g}_\ell \colon \boldsymbol{Z}_\ell \mapsto \boldsymbol{Z}_\ell \odot \boldsymbol{Z}_\ell$.

In fact, for large-scale applications of the problem described in section 3.2, the operator $\mathbf{A}_\ell = \mathbf{V}_\ell^\mathsf{T}(\mathbf{I}_\ell - \boldsymbol{\Delta}_\ell)^J$ is neither explicitly stored, nor directly accessible. Instead, only matrix-free representations of $\mathbf{U}_\ell := (\mathbf{I}_\ell - \boldsymbol{\Delta}_\ell)$ and $\mathbf{V}_\ell$ are available. The latter is designed such that $\mathbf{V}_\ell \mathbf{V}_\ell^\mathsf{T} = \mathbf{W}_\ell$ is a symmetric and positive definite Gram matrix related to the discretization of the diffusion equation so that $\boldsymbol{\Delta}_\ell$ is self-adjoint w.r.t. the Euclidean inner product weighted by $\mathbf{W}_\ell$, i.e., $\mathbf{W}_\ell \boldsymbol{\Delta}_\ell = \boldsymbol{\Delta}_\ell^\mathsf{T} \mathbf{W}_\ell$. As a result, $\mathbf{W}_\ell \boldsymbol{\Delta}_\ell$ is symmetric, and so is $\tilde{\mathbf{U}}_\ell := \mathbf{W}_\ell \mathbf{U}_\ell$. The solution $\boldsymbol{Z}_\ell := \boldsymbol{Z}_\ell^{(J)}$ of $\mathbf{A}_\ell \boldsymbol{Z}_\ell = \boldsymbol{X}_\ell$, with $\boldsymbol{X}_\ell := \bar{\mathbf{R}}_L^\ell \boldsymbol{X}$, can thus be obtained by solving the sequence of

symmetric systems

$$\tilde{\mathbf{U}}_\ell \boldsymbol{Z}_\ell^{(1)} = \mathbf{V}_\ell \boldsymbol{X}_\ell, \quad \tilde{\mathbf{U}}_\ell \boldsymbol{Z}_\ell^{(2)} = \mathbf{W}_\ell \boldsymbol{Z}_\ell^{(1)}, \quad \ldots, \quad \tilde{\mathbf{U}}_\ell \boldsymbol{Z}_\ell^{(J)} = \mathbf{W}_\ell \boldsymbol{Z}_\ell^{(J-1)}. \tag{6.1}$$

This raises interesting questions as to how the FMG-MLMC strategy described above may be applied to such a sequence of systems. A straightforward approach would be to apply FMG successively to each of the systems in eq. (6.1). More involved strategies could possibly be adapted from parallel-in-time multigrid methods (see, e.g., [75]) that are used for time-dependent problems disctretized with implicit time-stepping methods, resulting in sequences of systems similar to eq. (6.1). Alternative multigrid ideas might be devised by considering the algebraic form of the operator $\mathbf{A}_\ell$ involving powers of the discretized shifted diffusion operator $\mathbf{I}_\ell - \boldsymbol{\Delta}_\ell$.

6.2 Multifidelity ensemble-variational data assimilation

Another extension of the work presented in section 3.2 would be the multilevel, or multifidelity, estimation of background error covariance matrices using ensembles of data assimilation (EDAs) [76–79] for ensemble-variational data assimilation [80–82]. In such a framework, independent low-fidelity EDAs are used for the multifidelity estimation of a background error covariance matrix $\hat{\mathbf{B}}_{L-1}^{\mathrm{ML}}$, which is used in a high-fidelity, deterministic variational data assimilation method. This process is illustrated in fig. 6.2, with a multifidelity structure corresponding to that of (possibly weighted) MLMC (see fig. 3.1). In the more general multifidelity framework presented in chapter 3, the multifidelity estimator $\hat{\mathbf{B}}_{L-1}^{\mathrm{ML}}$ would read $\hat{\mathbf{B}}_{L-1}^{\mathrm{ML}} = \sum_{k=1}^{L-1} \sum_{\ell \in \mathcal{S}^k} w_\ell^{(k)} \hat{\mathbf{B}}_\ell^{(k)}$, where $\hat{\mathbf{B}}_\ell^{(k)} = \bar{\mathbf{P}}_\ell^L \mathbf{E}_\ell^{(k)} (\bar{\mathbf{P}}_\ell^L \mathbf{E}_\ell^{(k)})^\mathsf{T}$ are single-level estimators of the covariance matrix using the ensemble matrices $\mathbf{E}_\ell^{(k)}$ of fidelity level ℓ associated with the coupling group $\mathcal{S}^k$, appropriately prolongated to the high-fidelity level L. Specifically, the ensemble matrices $\mathbf{E}_\ell^{(k)}$ are defined by $\mathbf{E}_\ell^{(k)} = (n_k - 1)^{-1/2} [\boldsymbol{\varepsilon}_\ell^{(k,1)} \ \cdots \ \boldsymbol{\varepsilon}_\ell^{(k,n_k)}]$, with $\boldsymbol{\varepsilon}_\ell^{(k,i)} := \boldsymbol{x}_\ell^{\mathrm{b},(k,i)} - n_k^{-1} \sum_{j=1}^{n_k} \boldsymbol{x}_\ell^{\mathrm{b},(k,j)}$. In [M14], we proposed multivariate extensions of the original MBLUE framework proposed by [41, 42], with various flavors of MBLUE-like covariance matrix estimators with scalar, vector, or matrix weights.

Preliminary experiments, conducted on two-layer quasi-geostrophic model [84, 85], revealed a number of potential difficulties related to the multifidelity covariance matrix estimator described above. First, contrary to the classical, single-level MC estimator, a decomposition of the form $\hat{\mathbf{B}} = \mathbf{E}\mathbf{E}^\mathsf{T}$ is not readily available. Such a decomposition is used in some applications to generate a discretized correlated

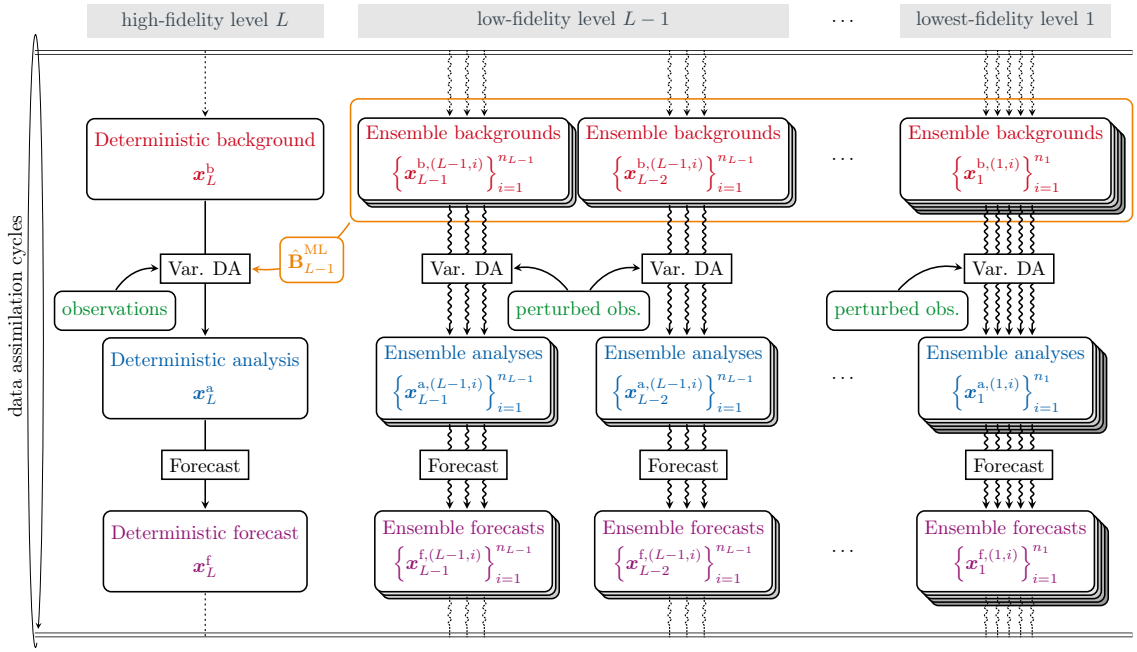


Figure 6.2: Sketch of a prospective multilevel EDA algorithm, inspired by [83].

Gaussian random field $\mathbf{E}\boldsymbol{\eta}$, with $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, having the covariance structure defined by $\hat{\mathbf{B}}$. Second, because of the linear combination of single-level covariance matrix estimators involved in the multilevel definition, realizations of $\hat{\mathbf{B}}_{L-1}^{\mathrm{ML}}$ are not guaranteed to be positive definite when using finite (and, *a fortiori*, small) ensemble sizes. This may cause problems in the optimization stage of the variational data assimilation, for which conjugate-gradient-like solvers, requiring positive definiteness, are typically used. Again, this is a drawback compared to the single-level MC estimator of the form $\hat{\mathbf{B}} = \mathbf{E}\mathbf{E}^{\mathsf{T}}$, for which positive semi-definiteness is guaranteed by construction. However, $\hat{\mathbf{B}}$ is usually rank-deficient in large-scale applications, because the ensemble size is typically orders of magnitude smaller than the number of degrees of freedom of the discretized system. The multilevel estimator does have a larger rank, because larger ensemble sizes are used on the lower-fidelity levels. A third potential difficulty of using the multilevel estimator in ensemble-variational data assimilation is to maintain the correlation of the ensembles within the coupling groups throughout the assimilation cycles. In the prospective framework depicted in fig. 6.2, the stochastic coupling originates from the perturbed observations alone, which may not ensure sufficient correlation for the multilevel estimator's variance to be significantly reduced.

As mentioned previously, a well-known issue of single-level MC estimates of covariance matrices is that they are singular for large-scale applications, which poses problems for the minimization algorithms used in variational data assimilation. A

common way to alleviate this problem is to resort to localization, which, in addition to reducing the sampling noise (variance) in regions "far away" from the diagonal, does ensure the strict positive definiteness of the localized matrix. Localizing, however, introduces bias into the otherwise unbiased estimator. An interesting questions lies in how to apply localization to the multilevel estimator. A natural way would consist in localizing the final estimator $\hat{\mathbf{B}}_{L-1}^{\mathrm{ML}}$ on level $L-1$. Alternatively, one could localize the prolongated single-level estimators $\hat{\mathbf{B}}_{\ell}^{(k)}$, or the unprolongated ones, $\mathbf{E}_{\ell}^{(k)}(\mathbf{E}_{\ell}^{(k)})^{\mathsf{T}}$, individually on each fidelity level. In any case, both approaches would also introduce bias into the otherwise unbiased multilevel estimator. This bias may be controlled through the optimal localization approach developed in [86–88]. Incidentally, this approach is very close to MBLUE, in that it seeks the localization operator that minimizes the mean square error (MSE) of the localized estimator. In a multifidelity framework, one may then try to find the optimal level-dependent localization operators that minimize the overall MSE. In the same spirit, it may also be possible to enforce the positive-definiteness (with a certain, high probability) of the resulting estimator through the addition of probabilistic constraints to the optimization problem, or possibly by imposing it by design, as was recently proposed in [89, 90].

6.3 Neural network preconditioners for flexible subspace methods

In this section, I present avenues for hybridizing machine learning techniques and numerical linear algebra solvers, specifically by using a trained neural network (NN) as a non-linear preconditioner for a flexible subspace iteration method. Let $\mathcal{N}_\theta$ denote a neural network that has been trained so that $\mathcal{N}_\theta(\boldsymbol{b})$ provides a good approximation to the solution of a linear system $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ corresponding to the discretization of a PDE of the form $\mathcal{L}u = f$. Provided that the trained network generalizes to a variety of plausible discretized sources $\boldsymbol{b}$, the mapping $\mathcal{N}_\theta \colon \boldsymbol{v} \mapsto \mathcal{N}_\theta(\boldsymbol{v}) \approx \mathbf{A}^{-1}\boldsymbol{v}$ can then be seen as a non-linear approximation of $\mathbf{A}^{-1}$, i.e., loosely speaking, $\mathcal{N}_\theta \approx \mathbf{A}^{-1}$. Thus, $\mathcal{N}_\theta$ may be used as a non-linear preconditioner for a flexible subspace iteration method such as the flexible generalized minimum residual (FGMRES) method or the flexible full orthogonalization method (FFOM) [91, 92]. Given an initial guess $\boldsymbol{x}_0$, at each iteration of the flexible solver, the iterate $\boldsymbol{x}_m$ is sought in the affine subspace $\boldsymbol{x}_0 + \operatorname{span}\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m\}$, with $\boldsymbol{z}_j = \mathcal{N}_\theta(\boldsymbol{v}_j)$ instead of $\boldsymbol{z}_j = \mathbf{M}^{-1}\boldsymbol{v}_j$ for a constant, linear, right preconditioner $\mathbf{M}$. A simplified NN-preconditioned FGMRES algorithm, adapted from [92, Algorithm 9.6], is provided in algorithm 2, where the

key step corresponding to the NN preconditioning is performed on line 3.

Algorithm 2: Simplified NN-preconditioned FGMRES algorithm.

1 $\boldsymbol{r}_0 = \boldsymbol{b} - \mathbf{A}\boldsymbol{x}_0$, $\beta = \|\boldsymbol{r}_0\|_2$, $\boldsymbol{v}_1 = \boldsymbol{r}_0/\beta$
2 **for** $j = 1, \ldots, m$ **do**
3 $\quad \boldsymbol{z}_j = \mathcal{N}_\theta(\boldsymbol{v}_j)$, $\boldsymbol{w} := \mathbf{A}\boldsymbol{z}_j$
4 $\quad$ **for** $i = 1, \ldots, j$ **do**
5 $\quad\quad h_{i,j} = \langle \boldsymbol{w}, \boldsymbol{v}_i \rangle$, $\boldsymbol{w} \leftarrow \boldsymbol{w} - h_{i,j}\boldsymbol{v}_i$
6 $\quad$ **end for**
7 $\quad h_{j+1,j} = \|\boldsymbol{w}\|_2$, $\boldsymbol{v}_{j+1} = \boldsymbol{w}/h_{j+1,j}$
8 **end for**
9 $\bar{\mathbf{H}}_m = [h_{i,j}]_{1 \le j \le m, 1 \le i \le j+1}$
10 $\boldsymbol{y}_m = \arg\min_{\boldsymbol{y}} \|\beta \boldsymbol{e}_1 - \bar{\mathbf{H}}_m \boldsymbol{y}\|_2$, $\boldsymbol{x}_m = \boldsymbol{x}_0 + \begin{bmatrix} \boldsymbol{z}_1 & \cdots & \boldsymbol{z}_m \end{bmatrix} \boldsymbol{y}_m$

A natural candidate NN architecture is that of neural operators [93], and more particularly Fourier neural operators (FNOs) [94], which are designed to learn an entire family of PDEs. Their principal advantage lies in the fact that a single trained neural operator can be used for different discretizations. A promising alternative is the recurrent NN based on a U-net architecture proposed in [95], which is trained by unsupervised learning inside non-linear fixed-point iterations. We performed preliminary experiments on discretized Helmholtz equations using both architectures [96, 97], which demonstrated the validity of the NN preconditioning approach described above, although the generalization to different meshes seems to be challenging. An interesting follow-up would be to devise an active learning strategy to train the NN-based preconditioner on-the-fly during the solving of sequences of linear systems corresponding to a massive number of right- and/or left-hand sides.

Co-authored publications

[M1] E. Agullo, M. Altenbernd, H. Anzt, L. Bautista-Gomez, T. Benacchio, L. Bonaventura, H.-J. Bungartz, S. Chatterjee, F. M. Ciorba, N. DeBardeleben, D. Drzisga, S. Eibl, C. Engelmann, W. N. Gansterer, L. Giraud, D. Göddeke, M. Heisig, F. Jézéquel, N. Kohl, X. S. Li, R. Lion, M. Mehl, P. Mycek, M. Obersteiner, E. S. Quintana-Ortí, F. Rizzi, U. Rüde, M. Schulz, F. Fung, R. Speck, L. Stals, K. Teranishi, S. Thibault, D. Thönnes, A. Wagner, and B. Wohlmuth. "Resiliency in Numerical Algorithm Design for Extreme Scale Simulations". In: *The International Journal of High Performance Computing Applications* (2021), p. 10943420211055188. ISSN: 1094-3420. DOI: 10.1177/10943420211055188. arXiv: 2010.13342.

[M2] K. Sargsyan, F. Rizzi, P. Mycek, C. Safta, K. Morris, H. Najm, O. Le Maître, O. Knio, and B. Debusschere. "Fault Resilient Domain Decomposition Preconditioner for PDEs". In: *SIAM Journal on Scientific Computing* 37.5 (2015), A2317–A2345. ISSN: 1064-8275. DOI: 10.1137/15M1014474. <Open Access>.

[M3] F. Rizzi, K. Morris, K. Sargsyan, P. Mycek, C. Safta, B. Debusschere, O. Le Maître, and O. Knio. "ULFM-MPI Implementation of a Resilient Task-Based Partial Differential Equations Preconditioner". In: *Proceedings of the ACM Workshop on Fault-Tolerance for HPC at Extreme Scale*. HPDC'16: The 25th International Symposium on High-Performance Parallel and Distributed Computing. Kyoto, Japan: ACM, 2016, pp. 19–26. ISBN: 978-1-4503-4349-7. DOI: 10.1145/2909428.2909429. <Open Access>.

[M4] P. Mycek, F. Rizzi, O. L. Maître, K. Sargsyan, K. Morris, C. Safta, B. Debusschere, and O. Knio. "Discrete A Priori Bounds for the Detection of Corrupted PDE Solutions in Exascale Computations". In: *SIAM Journal on Scientific Computing* 39.1 (2017), pp. C1–C28. ISSN: 1064-8275. DOI: 10.1137/15M1051786. <ResearchGate>.

[M5] F. Rizzi, K. Morris, K. Sargsyan, P. Mycek, C. Safta, O. Le Maître, O. Knio, and B. Debusschere. "Partial Differential Equations Preconditioner Resilient to Soft and Hard Faults". In: *The International Journal of High Performance Computing Applications* 32.5 (2018), pp. 658–673. ISSN: 1094-3420. DOI: 10.1177/1094342016684975. <Open Access>.

[M6] F. Rizzi, K. Morris, K. Sargsyan, P. Mycek, C. Safta, O. Le Maître, O. M. Knio, and B. J. Debusschere. "Exploring the Interplay of Resilience and Energy Consumption for a Task-Based Partial Differential Equations Preconditioner". In: *Parallel Computing*. Parallel Programming for Resilience and Energy Efficiency 73 (2018), pp. 16–27. ISSN: 0167-8191. DOI: 10.1016/j.parco.2017.05.005. <Open Access>.

[M7] P. Mycek, A. Contreras, O. Le Maître, K. Sargsyan, F. Rizzi, K. Morris, C. Safta, B. Debusschere, and O. Knio. "A Resilient Domain Decomposition Polynomial Chaos Solver for Uncertain Elliptic PDEs". In: *Computer Physics Communications* 216 (2017), pp. 18–34. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2017.02.015. <ResearchGate>.

[M8] A. A. Contreras, P. Mycek, O. P. Le Maître, F. Rizzi, B. Debusschere, and O. M. Knio. "Parallel Domain Decomposition Strategies for Stochastic Elliptic Equations. Part A: Local Karhunen–Loève Representations". In: *SIAM Journal on Scientific Computing* 40.4 (2018), pp. C520–C546. ISSN: 1064-8275. DOI: 10.1137/17M1132185. HAL: hal-02327682.

[M9] A. A. Contreras, P. Mycek, O. P. Le Maître, F. Rizzi, B. Debusschere, and O. M. Knio. "Parallel Domain Decomposition Strategies for Stochastic Elliptic Equations. Part B: Accelerated Monte Carlo Sampling with Local PC Expansions". In: *SIAM Journal on Scientific Computing* 40.4 (2018), pp. C547–C580. ISSN: 1064-8275. DOI: 10.1137/17M1132197. HAL: hal-02327678.

[M10] J. F. Reis, O. P. Le Maître, P. M. Congedo, and P. Mycek. "Stochastic Preconditioning of Domain Decomposition Methods for Elliptic Equations with Random Coefficients". In: *Computer Methods in Applied Mechanics and Engineering* 381 (2021), p. 113845. ISSN: 0045-7825. DOI: 10.1016/j.cma.2021.113845. HAL: hal-03201297.

[M11] P. Mycek and M. De Lozzo. "Multilevel Monte Carlo Covariance Estimation for the Computation of Sobol' Indices". In: *SIAM/ASA Journal on Uncertainty Quantification* 7.4 (2019), pp. 1323–1348. ISSN: 2166-2525. DOI: 10.1137/18M1216389. HAL: hal-01894503v2.

[M12] J. Briant, P. Mycek, M. Destouches, O. Goux, S. Gratton, S. Gürol, E. Simon, and A. T. Weaver. *A Filtered Multilevel Monte Carlo Method for Estimating the Expectation of Discretized Random Fields*. Submitted to SIAM/ASA Journal on Uncertainty Quantification. 2023. DOI: 10.48550/arXiv.2311.06069. arXiv: 2311.06069.

[M13] M. R. El Amri, P. Mycek, S. Ricci, and M. De Lozzo. *Multilevel Surrogate-based Control Variates*. Submitted to SIAM/ASA Journal on Uncertainty Quantification. 2023. DOI: 10.48550/arXiv.2306.10800. arXiv: 2306.10800.

[M14] M. Destouches, P. Mycek, and S. Gürol. *Multivariate Extensions of the Multilevel Best Linear Unbiased Estimator for Ensemble-Variational Data Assimilation*. 2023. DOI: 10.48550/arXiv.2306.07017. arXiv: 2306.07017.

[M15] D. A. Di Pietro, F. Hülsemann, P. Matalon, P. Mycek, U. Rüde, and D. Ruiz. "An h-Multigrid Method for Hybrid High-Order Discretizations". In: *SIAM Journal on Scientific Computing* 43.5 (2021), S839–S861. ISSN: 1064-8275. DOI: 10.1137/20M1342471. HAL: hal-02434411v3.

[M16] D. A. Di Pietro, F. Hülsemann, P. Matalon, P. Mycek, U. Rüde, and D. Ruiz. "Towards Robust, Fast Solutions of Elliptic Equations on Complex Domains through Hybrid High-Order Discretizations and Non-Nested Multigrid Methods". In: *International Journal for Numerical Methods in Engineering* 122.22 (2021), pp. 6576–6595. ISSN: 1097-0207. DOI: 10.1002/nme.6803. HAL: hal-03163476v2.

[M17] D. A. Di Pietro, P. Matalon, P. Mycek, and U. Rüde. "High-Order Multigrid Strategies for Hybrid High-Order Discretizations of Elliptic Equations". In: *Numerical Linear Algebra with Applications* (2022), e2456. ISSN: 1099-1506. DOI: 10.1002/nla.2456. HAL: hal-03531293v2.

[M18] D. A. Di Pietro, F. Hülsemann, P. Matalon, P. Mycek, and U. Rüde. "Algebraic Multigrid Preconditioner for Statically Condensed Systems Arising from Lowest-Order Hybrid Discretizations". In: *SIAM Journal on Scientific Computing* 45.3 (2023), S329–S350. ISSN: 1064-8275. DOI: 10.1137/21M1429849. HAL: hal-03272468v2.

[M19] N. Venkovic, P. Mycek, L. Giraud, and O. Le Maitre. *Recycling Krylov Subspace Strategies for Sequences of Sampled Stochastic Elliptic Equations*. Research Report RR-9425. Inria Bordeaux - Sud Ouest, 2021. URL: https://hal.science/hal-03366966. HAL: hal-03366966.

[M20] P. Mycek. "Étude numérique et expérimentale du comportement d'hydroliennes". PhD thesis. Université du Havre, 2013. URL: https://theses.fr/en/2013LEHA0010.

[M21] P. Mycek, G. Pinon, C. Lothodé, A. Dezotti, and C. Carlier. "Iterative Solver Approach for Turbine Interactions: Application to Wind or Marine Current Turbine Farms". In: *Applied Mathematical Modelling* 41 (2017), pp. 331–349. ISSN: 0307-904X. DOI: 10.1016/j.apm.2016.08.027. HAL: hal-01919332v1.

Bibliography

[1] A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential Equations*. Numerical Mathematics and Scientific Computation. Oxford, New York: Oxford University Press, 1999. ISBN: 978-0-19-850178-7. URL: https://global.oup.com/academic/product/domain-decomposition-methods-for-partial-differential-equations-9780198501787.

[2] A. Toselli and O. Widlund. *Domain Decomposition Methods - Algorithms and Theory*. Springer Series in Computational Mathematics. Berlin Heidelberg: Springer-Verlag, 2005. ISBN: 978-3-540-20696-5. DOI: 10.1007/b137868.

[3] T. P. A. Mathew. *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*. Ed. by T. J. Barth, M. Griebel, D. E. Keyes, R. M. Nieminen, D. Roose, and T. Schlick. Lecture Notes in Computational Science and Engineering 61. Berlin: Springer, 2008. ISBN: 978-3-540-77205-7. DOI: 10.1007/978-3-540-77209-5.

[4] V. Dolean, P. Jolivet, and F. Nataf. *An Introduction to Domain Decomposition Methods: Algorithms, Theory, and Parallel Implementation*. Philadelphia: Society for Industrial and Applied Mathematics, 2015. ISBN: 978-1-61197-405-8. DOI: 10.1137/1.9781611974065.

[5] M. Snir, R. W. Wisniewski, J. A. Abraham, S. V. Adve, S. Bagchi, P. Balaji, J. Belak, P. Bose, F. Cappello, B. Carlson, A. A. Chien, P. Coteus, N. A. DeBardeleben, P. C. Diniz, C. Engelmann, M. Erez, S. Fazzari, A. Geist, R. Gupta, F. Johnson, S. Krishnamoorthy, S. Leyffer, D. Liberty, S. Mitra, T. Munson, R. Schreiber, J. Stearley, and E. V. Hensbergen. "Addressing Failures in Exascale Computing". In: *The International Journal of High Performance Computing Applications* 28.2 (2014), pp. 129–173. ISSN: 1094-3420, 1741-2846. DOI: 10.1177/1094342014522573.

[6] M. A. Heroux. *Toward Resilient Algorithms and Applications*. 2014. DOI: 10.48550/arXiv.1402.3809. arXiv: 1402.3809.

[7] S. Heinrich. "Monte Carlo Complexity of Global Solution of Integral Equations". In: *Journal of Complexity* 14.2 (1998), pp. 151–175. ISSN: 0885064X. DOI: 10.1006/jcom.1998.0471.

[8] M. B. Giles. "Multilevel Monte Carlo Path Simulation". In: *Operations Research* 56.3 (2008), pp. 607–617. ISSN: 0030-364X, 1526-5463. DOI: 10.1287/opre.1070.0496.

[9] M. B. Giles. "Multilevel Monte Carlo Methods". In: *Acta Numerica* 24 (2015), pp. 259–328. DOI: 10.1017/S096249291500001X.

[10] A. Sarkar, N. Benabbou, and R. Ghanem. "Domain Decomposition of Stochastic PDEs: Theoretical Formulations". In: *International Journal for Numerical Methods in Engineering* 77.5 (2009), pp. 689–701. ISSN: 1097-0207. DOI: 10.1002/nme.2431.

[11] Y. Chen, J. Jakeman, C. Gittelson, and D. Xiu. "Local Polynomial Chaos Expansion for Linear Differential Equations with High Dimensional Random Inputs". In: *SIAM Journal on Scientific Computing* 37.1 (2015), A79–A102. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/140970100.

[12] S. Pranesh and D. Ghosh. "Addressing the Curse of Dimensionality in SSFEM Using the Dependence of Eigenvalues in KL Expansion on Domain Size". In: *Computer Methods in Applied Mechanics and Engineering* 311 (2016), pp. 457–475. ISSN: 0045-7825. DOI: 10.1016/j.cma.2016.08.023.

[13] D. Zhang, H. Babaee, and G. E. Karniadakis. "Stochastic Domain Decomposition via Moment Minimization". In: *SIAM Journal on Scientific Computing* 40.4 (2018), A2152–A2173. ISSN: 1064-8275. DOI: 10.1137/17M1160756.

[14] L. Chen, Y. Chen, Q. Li, and Z. Zhang. *Stochastic Domain Decomposition Based on Variable-Separation Method*. 2023. DOI: 10.48550/arXiv.2304.05708. arXiv: 2304.05708.

[15] M. Dryja and O. Widlund. *An Additive Variant of the Schwarz Alternating Method for the Case of Many Subregions*. Technical report 339. Division of Computer Science, Courant Institute, New York University, 1987.

[16] S. Portnoy and R. Koenker. "The Gaussian Hare and the Laplacian Tortoise: Computability of Squared-Error versus Absolute-Error Estimators". In: *Statistical Science* 12.4 (1997). ISSN: 0883-4237. DOI: 10.1214/ss/1030037960.

[17] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk. "Iteratively Reweighted Least Squares Minimization for Sparse Recovery". In: *Communications on Pure and Applied Mathematics* 63.1 (2010), pp. 1–38. ISSN: 1097-0312. DOI: 10.1002/cpa.20303.

[18] P. G. Bridges, K. B. Ferreira, M. A. Heroux, and M. Hoemmen. *Fault-Tolerant Linear Solvers via Selective Reliability*. 2012. DOI: 10.48550/arXiv.1206.1390. arXiv: 1206.1390.

[19] W. Bland, A. Bouteiller, T. Herault, G. Bosilca, and J. Dongarra. "Post-Failure Recovery of MPI Communication Capability: Design and Rationale". In: *The International Journal of High Performance Computing Applications* 27.3 (2013), pp. 244–254. ISSN: 1094-3420. DOI: 10.1177/1094342013488238.

[20] M.-L. Li, P. Ramachandran, S. K. Sahoo, S. V. Adve, V. S. Adve, and Y. Zhou. "Understanding the Propagation of Hard Errors to Software and Implications for Resilient System Design". In: *ACM SIGOPS Operating Systems Review* 42.2 (2008), pp. 265–276. ISSN: 0163-5980. DOI: 10.1145/1353535.1346315.

[21] C. Engelmann and T. Naughton. "Toward a Performance/Resilience Tool for Hardware/Software Co-design of High-Performance Computing Systems". In: *2013 42nd International Conference on Parallel Processing*. 2013 42nd International Conference on Parallel Processing. 2013, pp. 960–969. DOI: 10.1109/ICPP.2013.114.

[22] J. Elliott. "Resilient Iterative Linear Solvers Running Through Errors". PhD thesis. North Carolina State University, 2015. URL: https://repository.lib.ncsu.edu/handle/1840.16/10701.

[23] N. Wiener. "The Homogeneous Chaos". In: *American Journal of Mathematics* 60.4 (1938), pp. 897–936. ISSN: 0002-9327. DOI: 10.2307/2371268.

[24] R. H. Cameron and W. T. Martin. "The Orthogonal Development of Non-Linear Functionals in Series of Fourier-Hermite Functionals". In: *Annals of Mathematics* 48.2 (1947), pp. 385–392. ISSN: 0003-486X. DOI: 10.2307/1969178. JSTOR: 1969178.

[25] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Revised Edition. Dover publications, 2003. ISBN: 978-0-486-42818-5.

[26] O. P. Le Maître and O. M. Knio. *Spectral Methods for Uncertainty Quantification*. Scientific Computation. Dordrecht: Springer Netherlands, 2010. ISBN: 978-90-481-3519-6. DOI: 10.1007/978-90-481-3520-2.

[27] R. Tibshirani. "Regression Shrinkage and Selection via the Lasso". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 58.1 (1996), pp. 267–288. ISSN: 2517-6161. DOI: 10.1111/j.2517-6161.1996.tb02080.x.

[28] D. Zhu, R. Melhem, and D. Mosse. "The Effects of Energy Management on Reliability in Real-Time Embedded Systems". In: *IEEE/ACM International Conference on Computer Aided Design, 2004. ICCAD-2004.* IEEE/ACM International Conference on Computer Aided Design, 2004. ICCAD-2004. 2004, pp. 35–40. DOI: 10.1109/ICCAD.2004.1382539.

[29] M. Kac and A. J. F. Siegert. "An Explicit Representation of a Stationary Gaussian Process". In: *The Annals of Mathematical Statistics* 18.3 (1947), pp. 438–442. ISSN: 0003-4851, 2168-8990. DOI: 10.1214/aoms/1177730391.

[30] W. Betz, I. Papaioannou, and D. Straub. "Numerical Methods for the Discretization of Random Fields by Means of the Karhunen–Loève Expansion". In: *Computer Methods in Applied Mechanics and Engineering* 271 (2014), pp. 109–129. ISSN: 0045-7825. DOI: 10.1016/j.cma.2013.12.010.

[31] V. Hernandez, J. E. Roman, and V. Vidal. "SLEPc: A Scalable and Flexible Toolkit for the Solution of Eigenvalue Problems". In: *ACM Transactions on Mathematical Software* 31.3 (2005), pp. 351–362. ISSN: 0098-3500. DOI: 10.1145/1089014.1089019.

[32] J. F. Reis. "Preconditioning of Domain Decomposition Methods for Stochastic Elliptic Equations". PhD thesis. Institut polytechnique de Paris, 2021. URL: http://www.theses.fr/en/2021IPPAX059.

[33] N. Spillane. "An Adaptive MultiPreconditioned Conjugate Gradient Algorithm". In: *SIAM Journal on Scientific Computing* 38.3 (2016), A1896–A1918. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/15M1028534.

[34] L. W. T. Ng and K. E. Willcox. "Multifidelity Approaches for Optimization under Uncertainty". In: *International Journal for Numerical Methods in Engineering* 100.10 (2014), pp. 746–772. ISSN: 00295981. DOI: 10.1002/nme.4761.

[35] B. Peherstorfer, K. Willcox, and M. Gunzburger. "Optimal Model Management for Multifidelity Monte Carlo Estimation". In: *SIAM Journal on Scientific Computing* 38.5 (2016), A3163–A3194. ISSN: 1064-8275. DOI: 10.1137/15M1046472.

[36] A.-L. Haji-Ali, F. Nobile, and R. Tempone. "Multi-Index Monte Carlo: When Sparsity Meets Sampling". In: *Numerische Mathematik* 132.4 (2016), pp. 767–806. ISSN: 0029-599X, 0945-3245. DOI: 10.1007/s00211-015-0734-5.

[37] S. S. Lavenberg, T. L. Moeller, and P. D. Welch. "The Application of Control Variables to the Simulation of Closed Queueing Networks". In: *Proceedings of the 9th Conference on Winter Simulation - Volume 1*. WSC '77. Gaithersburg, Maryland, USA: Winter Simulation Conference, 1977, pp. 152–154. URL: https://dl.acm.org/doi/abs/10.5555/800251.807534.

[38] C. Lemieux. "Control Variates". In: *Wiley StatsRef: Statistics Reference Online*. American Cancer Society, 2017, pp. 1–8. ISBN: 978-1-118-44511-2. DOI: 10.1002/9781118445112.stat07947.

[39] R. Pasupathy, B. W. Schmeiser, M. R. Taaffe, and J. Wang. "Control-Variate Estimation Using Estimated Control Means". In: *IIE Transactions* 44.5 (2012), pp. 381–385. ISSN: 0740-817X, 1545-8830. DOI: 10.1080/0740817X.2011.610430.

[40] A. A. Gorodetsky, G. Geraci, M. S. Eldred, and J. D. Jakeman. "A Generalized Approximate Control Variate Framework for Multifidelity Uncertainty Quantification". In: *Journal of Computational Physics* 408 (2020), p. 109257. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2020.109257.

[41] D. Schaden and E. Ullmann. "On Multilevel Best Linear Unbiased Estimators". In: *SIAM/ASA Journal on Uncertainty Quantification* 8.2 (2020), pp. 601–635. DOI: 10.1137/19M1263534.

[42] D. Schaden and E. Ullmann. "Asymptotic Analysis of Multilevel Best Linear Unbiased Estimators". In: *SIAM/ASA Journal on Uncertainty Quantification* 9.3 (2021), pp. 953–978. DOI: 10.1137/20M1321607.

[43] J. Šukys, U. Rasthofer, F. Wermelinger, P. Hadjidoukas, and P. Koumoutsakos. *Optimal Fidelity Multi-Level Monte Carlo for Quantification of Uncertainty in Simulations of Cloud Cavitation Collapse*. 2017. DOI: 10.48550/arXiv.1705.04374. arXiv: 1705.04374.

[44] B. Peherstorfer, K. Willcox, and M. Gunzburger. "Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization". In: *SIAM Review* 60.3 (2018), pp. 550–591. ISSN: 0036-1445. DOI: 10.1137/16M1082469.

[45] M. Croci, K. E. Willcox, and S. J. Wright. *Multi-Output Multilevel Best Linear Unbiased Estimators via Semidefinite Programming*. 2023. DOI: 10.48550/arXiv.2301.07831. arXiv: 2301.07831.

[46] A. Janon, T. Klein, A. Lagnoux, M. Nodet, and C. Prieur. "Asymptotic Normality and Efficiency of Two Sobol Index Estimators". In: *ESAIM: Probability and Statistics* 18 (2014), pp. 342–364. ISSN: 1292-8100, 1262-3318. DOI: 10.1051/ps/2013040.

[47] A. L. Teckentrup. "Multilevel Monte Carlo Methods and Uncertainty Quantification". PhD thesis. University of Bath, 2013. URL: https://researchportal.bath.ac.uk/en/studentTheses/multilevel-monte-carlo-methods-and-uncertainty-quantication.

[48] W. L. Briggs, V. E. Henson, and S. F. McCormick. *A Multigrid Tutorial*. 2nd ed. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics, 2000. 196 pp. ISBN: 978-0-89871-462-3. DOI: 10.1137/1.9780898719505.

[49] A. Weaver and P. Courtier. "Correlation Modelling on the Sphere Using a Generalized Diffusion Equation". In: *Quarterly Journal of the Royal Meteorological Society* 127.575 (2001), pp. 1815–1846. ISSN: 1477-870X. DOI: 10.1002/qj.49712757518.

[50] I. Mirouze and A. T. Weaver. "Representation of Correlation Functions in Variational Assimilation Using an Implicit Diffusion Operator". In: *Quarterly Journal of the Royal Meteorological Society* 136.651 (2010), pp. 1421–1443. ISSN: 1477-870X. DOI: 10.1002/qj.643.

[51] A. T. Weaver, M. Chrust, B. Ménétrier, and A. Piacentini. "An Evaluation of Methods for Normalizing Diffusion-based Covariance Operators in Variational Data Assimilation". In: *Quarterly Journal of the Royal Meteorological Society* 147.734 (2021), pp. 289–320. ISSN: 0035-9009, 1477-870X. DOI: 10.1002/qj.3918.

[52] B. L. Nelson. "Control Variate Remedies". In: *Operations Research* 38.6 (1990), pp. 974–992. ISSN: 0030-364X, 1526-5463. DOI: 10.1287/opre.38.6.974.

[53] S. Garg, N. Sood, and C. D. Sarris. "Uncertainty quantification of ray-tracing based wireless propagation models with a Control Variate-Polynomial Chaos Expansion method". In: *2015 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting*. 2015, pp. 1776–1777. DOI: 10.1109/APS.2015.7305277.

[54] Z. Gu and C. D. Sarris. "Multi-parametric uncertainty quantification with a hybrid Monte-Carlo / polynomial chaos expansion FDTD method". In: *2015 IEEE MTT-S International Microwave Symposium*. 2015, pp. 1–3. DOI: 10.1109/MWSYM.2015.7166881.

[55] J. Fox and G. Ökten. "Polynomial Chaos as a Control Variate Method". In: *SIAM Journal on Scientific Computing* 43.3 (2021), A2268–A2294. DOI: 10.1137/20M1336515.

[56] H. Yang, Y. Fujii, K. W. Wang, and A. A. Gorodetsky. *Control Variate Polynomial Chaos: Optimal Fusion of Sampling and Surrogates for Multifidelity Uncertainty Quantification*. 2022. DOI: 10.48550/arXiv.2201.10745. arXiv: 2201.10745.

[57] D. A. Di Pietro, A. Ern, and S. Lemaire. "An Arbitrary-Order and Compact-Stencil Discretization of Diffusion on General Meshes Based on Local Reconstruction Operators". In: *Computational Methods in Applied Mathematics* 14.4 (2014). ISSN: 1609-4840, 1609-9389. DOI: 10.1515/cmam-2014-0018.

[58] D. A. Di Pietro and A. Ern. "A Hybrid High-Order Locking-Free Method for Linear Elasticity on General Meshes". In: *Computer Methods in Applied Mechanics and Engineering* 283 (2015), pp. 1–21. ISSN: 00457825. DOI: 10.1016/j.cma.2014.09.009.

[59] D. A. Di Pietro and J. Droniou. *The Hybrid High-Order Method for Polytopal Meshes*. Modeling, Simulation and Applications Series. Springer International Publishing, 2020. DOI: 10.1007/978-3-030-37203-3.

[60] P. Matalon. "Fast Solvers for Robust Discretizations in Computational Fluid Dynamics". PhD thesis. Université Montpellier ; Friedrich-Alexander-Universität Erlangen-Nürnberg, 2021. URL: https://theses.fr/en/2021MONTS051.

[61] A. Brandt. "Guide to Multigrid Development". In: *Multigrid Methods*. Ed. by W. Hackbusch and U. Trottenberg. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer, 1982, pp. 220–312. ISBN: 978-3-540-39544-7. DOI: 10.1007/BFb0069930.

[62] W. Hackbusch. *Multi-Grid Methods and Applications*. Berlin Heidelberg: Springer-Verlag, 1985. ISBN: 978-3-540-12761-1. DOI: 10.1007/978-3-662-02427-0.

[63] U. Trottenberg, C. W. Oosterlee, and A. Schuller. *Multigrid*. Elsevier, 2000. 648 pp. ISBN: 978-0-08-047956-9. URL: https://www.elsevier.com/books/multigrid/trottenberg/978-0-08-047956-9.

[64] Y. Notay. "An Aggregation-Based Algebraic Multigrid Method". In: *Electronic Transactions on Numerical Analysis* 37 (2010), pp. 123–146. URL: https://etna.ricam.oeaw.ac.at/volumes/2001-2010/vol37/abstract.php?vol=37&pages=123-146.

[65] H. A. van der Vorst. "Bi-CGSTAB: A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems". In: *SIAM Journal on Scientific and Statistical Computing* 13.2 (1992), pp. 631–644. ISSN: 0196-5204. DOI: 10.1137/0913035.

[66] N. Venkovic. "Preconditioning Strategies for Stochastic Elliptic Partial Differential Equations". PhD thesis. Université de Bordeaux, 2023. URL: http://www.theses.fr/en/s365259.

[67] Y. Saad, M. Yeung, J. Erhel, and F. Guyomarc'h. "A Deflated Version of the Conjugate Gradient Algorithm". In: *SIAM Journal on Scientific Computing* 21.5 (2000), pp. 1909–1926. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/S1064829598339761.

[68] E. Liberty, F. Woolfe, P.-G. Martinsson, V. Rokhlin, and M. Tygert. "Randomized Algorithms for the Low-Rank Approximation of Matrices". In: *Proceedings of the National Academy of Sciences* 104.51 (2007), pp. 20167–20172. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0709640104.

[69] P.-G. Martinsson. *Randomized Methods for Matrix Computations*. 2019. DOI: 10.48550/arXiv.1607.01649. arXiv: 1607.01649.

[70] P.-G. Martinsson and J. A. Tropp. "Randomized Numerical Linear Algebra: Foundations and Algorithms". In: *Acta Numerica* 29 (2020), pp. 403–572. ISSN: 0962-4929, 1474-0508. DOI: 10.1017/S0962492920000021.

[71] A. Scotto Di Perrotolo. "Randomized Numerical Linear Algebra Methods with Application to Data Assimilation." PhD thesis. Toulouse, ISAE, 2022. URL: https://www.theses.fr/en/2022ESAE0039.

[72] S. Gratton, A. Sartenaer, and J. Tshimanga. "On A Class of Limited Memory Preconditioners For Large Scale Linear Systems With Multiple Right-Hand Sides". In: *SIAM Journal on Optimization* 21.3 (2011), pp. 912–935. ISSN: 1052-6234, 1095-7189. DOI: 10.1137/08074008.

[73] P. Kumar, C. W. Oosterlee, and R. P. Dwight. "A Multigrid Multilevel Monte Carlo Method Using High-Order Finite-Volume Scheme for Lognormal Diffusion Problems". In: *International Journal for Uncertainty Quantification* 7.1 (2017). ISSN: 2152-5080, 2152-5099. DOI: 10.1615/Int.J.UncertaintyQuantification.2016018677.

[74] P. Kumar, C. Rodrigo, F. J. Gaspar, and C. W. Oosterlee. "On Local Fourier Analysis of Multigrid Methods for PDEs with Jumping and Random Coefficients". In: *SIAM Journal on Scientific Computing* 41.3 (2019), A1385–A1413. ISSN: 1064-8275. DOI: 10.1137/18M1173769.

[75] R. D. Falgout, S. Friedhoff, T. V. Kolev, S. P. MacLachlan, and J. B. Schroder. "Parallel Time Integration with Multigrid". In: *SIAM Journal on Scientific Computing* 36.6 (2014), pp. C635–C661. ISSN: 1064-8275. DOI: 10.1137/130944230.

[76] M. Fisher. "Background Error Covariance Modelling". In: *ECMWF Annual Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*. Reading, UK, 2003, pp. 45–63. URL: https://www.ecmwf.int/sites/default/files/elibrary/2003/74483-background-error-covariance-modelling_0.pdf.

[77] N. Zagar, E. Andersson, and M. Fisher. "Balanced Tropical Data Assimilation Based on a Study of Equatorial Waves in ECMWF Short-Range Forecast Errors". In: *Quarterly Journal of the Royal Meteorological Society* 131.607 (2005), pp. 987–1011. ISSN: 1477-870X. DOI: 10.1256/qj.04.54.

[78] M. B. Pereira and L. Berre. "The Use of an Ensemble Approach to Study the Background Error Covariances in a Global NWP Model". In: *Monthly Weather Review* 134.9 (2006), pp. 2466–2489. ISSN: 1520-0493, 0027-0644. DOI: 10.1175/MWR3189.1.

[79] M. Bonavita, L. Isaksen, and E. Hólm. "On the Use of EDA Background Error Variances in the ECMWF 4D-Var". In: *Quarterly Journal of the Royal Meteorological Society* 138.667 (2012), pp. 1540–1559. ISSN: 1477-870X. DOI: 10.1002/qj.1899.

[80] A. C. Lorenc. "The Potential of the Ensemble Kalman Filter for NWP—a Comparison with 4D-Var". In: *Quarterly Journal of the Royal Meteorological Society* 129.595 (2003), pp. 3183–3203. ISSN: 1477-870X. DOI: 10.1256/qj.02.132.

[81] M. Buehner. "Ensemble-Derived Stationary and Flow-Dependent Background-Error Covariances: Evaluation in a Quasi-Operational NWP Setting". In: *Quarterly Journal of the Royal Meteorological Society* 131.607 (2005), pp. 1013–1043. ISSN: 1477-870X. DOI: 10.1256/qj.04.15.

[82] R. N. Bannister. "A Review of Operational Methods of Variational and Ensemble-Variational Data Assimilation". In: *Quarterly Journal of the Royal Meteorological Society* 143.703 (2017), pp. 607–633. ISSN: 1477-870X. DOI: 10.1002/qj.2982.

[83] T. Auligné, B. Ménétrier, A. C. Lorenc, and M. Buehner. "Ensemble-Variational Integrated Localized Data Assimilation". In: *Monthly Weather Review* 144.10 (2016), pp. 3677–3696. ISSN: 0027-0644, 1520-0493. DOI: 10.1175/MWR-D-15-0252.1.

[84] C. B. Fandry and L. M. Leslie. "A Two-Layer Quasi-Geostrophic Model of Summer Trough Formation in the Australian Subtropical Easterlies". In: *Journal of the Atmospheric Sciences* 41.5 (1984), pp. 807–818. ISSN: 0022-4928, 1520-0469. URL: https://journals.ametsoc.org/view/journals/atsc/41/5/1520-0469_1984_041_0807_atlqgm_2_0_co_2.xml.

[85] M. Fisher and S. Gürol. "Parallelization in the Time Dimension of Four-Dimensional Variational Data Assimilation". In: *Quarterly Journal of the Royal Meteorological Society* 143.703 (2017), pp. 1136–1147. ISSN: 1477-870X. DOI: 10.1002/qj.2997.

[86] B. Ménétrier, T. Montmerle, Y. Michel, and L. Berre. "Linear Filtering of Sample Covariances for Ensemble-Based Data Assimilation. Part I: Optimality Criteria and Application to Variance Filtering and Covariance Localization". In: *Monthly Weather Review* 143.5 (2015), pp. 1622–1643. ISSN: 1520-0493, 0027-0644. DOI: 10.1175/MWR-D-14-00157.1.

[87] B. Ménétrier, T. Montmerle, Y. Michel, and L. Berre. "Linear Filtering of Sample Covariances for Ensemble-Based Data Assimilation. Part II: Application to a Convective-Scale NWP Model". In: *Monthly Weather Review* 143.5 (2015), pp. 1644–1664. ISSN: 1520-0493, 0027-0644. DOI: 10.1175/MWR-D-14-00156.1.

[88] B. Ménétrier and T. Auligné. "Optimized Localization and Hybridization to Filter Ensemble-Based Covariances". In: *Monthly Weather Review* 143.10 (2015), pp. 3931–3947. ISSN: 1520-0493, 0027-0644. DOI: 10.1175/MWR-D-15-0057.1.

[89] A. Maurais, T. Alsup, B. Peherstorfer, and Y. Marzouk. *Multi-Fidelity Covariance Estimation in the Log-Euclidean Geometry*. 2023. DOI: 10.48550/arXiv.2301.13749. arXiv: 2301.13749.

[90] A. Maurais, T. Alsup, B. Peherstorfer, and Y. Marzouk. *Multifidelity Covariance Estimation via Regression on the Manifold of Symmetric Positive Definite Matrices*. 2023. DOI: 10.48550/arXiv.2307.12438. arXiv: 2307.12438.

[91] Y. Saad. "A Flexible Inner-Outer Preconditioned GMRES Algorithm". In: *SIAM Journal on Scientific Computing* 14.2 (1993), pp. 461–469. ISSN: 1064-8275. DOI: 10.1137/0914028.

[92] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics, 2003. ISBN: 978-0-89871-534-7. DOI: 10.1137/1.9780898718003.

[93] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar. *Neural Operator: Graph Kernel Network for Partial Differential Equations*. 2020. DOI: 10.48550/arXiv.2003.03485. arXiv: 2003.03485.

[94] Z. Li, N. B. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar. "Fourier Neural Operator for Parametric Partial Differential Equations". In: *International Conference on Learning Representations*. 2021. URL: https://openreview.net/forum?id=c8P9NQVtmnO.

[95] A. Stanziola, S. R. Arridge, B. T. Cox, and B. E. Treeby. "A Helmholtz Equation Solver Using Unsupervised Learning: Application to Transcranial Ultrasound". In: *Journal of Computational Physics* 441 (2021), p. 110430. ISSN: 00219991. DOI: 10.1016/j.jcp.2021.110430.

[96] Y.-F. Xiang. "Solution of Large Linear Systems with a Massive Number of Right-Hand Sides and Machine Learning". PhD thesis. Bordeaux, 2022. URL: https://theses.fr/en/2022BORD0383.

[97] M. Shpakovych. *Neural Network Preconditioning of Large Linear Systems*. Technical Report. Inria Centre at the University of Bordeaux, 2023. URL: https://hal.science/hal-04254315. HAL: hal-04254315.

Méthodes hiérarchiques pour des équations aux dérivées partielles déterministes et stochastiques

Les méthodes hiérarchiques sont devenues un outil essentiel pour la simulation numérique efficace de phénomènes physiques avec une fidélité croissante sur des calculateurs à haute performance modernes. Des méthodes bien établies, telles que les méthodes multigrilles et de décomposition de domaine, se sont avérées performantes, de par leur capacité de passage à l'échelle, pour résoudre certains problèmes de grande taille de manière parallèle sur des supercalculateurs. D'autre part, la propagation des incertitudes dans les simulations numériques a reçu une attention croissante ces dernières années. Pour des simulations onéreuses en temps de calcul de problèmes fortement non-linéaires (par rapport à des paramètres d'entrée incertains), des méthodes avancées de quantification des incertitudes doivent être dévelopées. Dans ce manuscrit, nous exposons des idées méthodologiques et algorithmiques pour répondre à des questions spécifiques liées aux méthodes hiérarchiques. Premièrement, nous présentons des approches de décomposition de domaine résilientes pour la résolution numérique d'équations aux dérivées partielles déterministes et stochastiques, ainsi que des stratégies de Monte Carlo accélérées par des modèles de substitution pour la propagation des incertitudes dans des méthodes de décomposition de domaine. Deuxièmement, nous présentons de nouveaux solveurs multigrilles pour des systèmes linéaires issus de discrétisations hybrides d'ordre élevé. Troisièmement, nous proposons des approches hiérarchiques, plus précisément des méthodes multiniveaux et multifidélité, pour la propagation d'incertitudes dans les simulateurs numériques coûteux. Enfin, nous proposons également des idées pour la résolution numérique efficace de suites de systèmes d'équations linéaires, rencontrées typiquement lors de la résolution de problèmes instationnaires et/ou non linéaires, ou lors de l'échantillonnage d'équations aux dérivées partielles stochastiques.

Hierarchical methods for deterministic and stochastic partial differential equations

Hierarchical methods have become essential tools for the efficient numerical simulation of physical phenomena with ever-growing fidelity on modern high-performance supercomputers. Well-established methods, such as multigrid and domain decomposition methods, owing to their scaling capabilities, have proven powerful to solve certain classes of large-scale problems in a parallel fashion on supercomputers. On the other hand, the propagation of uncertainties in numerical simulations has received increasing attention in recent years. For computationally demanding simulations of highly non-linear problems (w.r.t. uncertain input parameters), advanced uncertainty quantification methods need to be designed. In this manuscript, we present methodological and algorithmic ideas to address specific questions related to hierarchical methods. First, we present resilient domain decomposition approaches for the numerical solution of deterministic and stochastic partial differential equations, as well as surrogate-assisted Monte Carlo strategies for the propagation of uncertainties in domain decomposition solvers. Second, we introduce novel multigrid solvers for the linear systems arising from hybridizable high-order discretizations. Third, we propose hierarchical approaches, namely multilevel and multifidelity methods, for the propagation of uncertainties in expensive numerical simulators. Finally, we also propose ideas for the efficient numerical solving of sequences of systems of linear equations that typically arise when solving time-dependent and/or non-linear problems, or when sampling stochastic partial differential equations.